The True Cost of Air Pollution: Evidence from the Housing Market

Daniel M. Sullivan*

March 22, 2017

Abstract

This paper presents evidence that current research significantly underestimates the effects of air pollution. Conventional methods of measuring pollution exposure cannot account for sharp changes in pollution over short distances or the wind-driven dispersion of pollutants. I use a state-of-the-art atmospheric dispersion model, which solves these problems, with a natural experiment to estimate the causal effect of NO_x exposure on house prices in metropolitan Los Angeles. The wind-based estimates are over 10 times larger than conventional estimates and imply that the net social benefit of RECLAIM, the local cap-and-trade program underlying the natural experiment, is roughly \$500 million per year.

*Resources for the Future, Washington, DC (email: sullivan@rff.org). I am especially grateful to David Cutler, Edward Glaeser, Lawrence Katz, and Robert Stavins for their feedback. I also thank Spencer Banzhaf, John Coglianese, Timothy Layton, Jing Li, Jonathan Libgober, Amanda Pallais, Christopher Palmer, Jisung Park, Daniel Pollmann, and James Stock, as well as seminar participants at Harvard, BYU, Notre Dame, University of Wisconsin-Madison, Resources for the Future, and Camp Resources XXIII. Funding from the National Institute on Aging, through Grant Number T32-AG000186 to the National Bureau of Economic Research, is gratefully acknowledged.

House price capitalization is routinely used to measure the social value of local amenities which lack an explicit market. But the case of air pollution presents a puzzle: house prices do not seem to respond very much to air pollution. Smith and Huang (1995) note that improved air quality affects house prices much less than would be expected given the health benefits. More recent studies, including those using quasi-experimental research designs, have not resolved this puzzle.[1] Deepening the confusion is a large literature which finds house prices to be very responsive to many other locational amenities, including school quality (Black 1999; Cellini, Ferreira, and Rothstein 2010); crime risk (Linden and Rockoff 2008; Pope 2008); and local cancer risk (Davis 2004). What is different about air pollution, a disamenity whose negative value has been well established in other contexts?[2]

In this paper, I present evidence that air quality does have a large effect on house prices and that estimates of the impact of pollution exposure can be severely biased when exposure is poorly measured. Unlike other economic variables, there are no large-sample data on air pollution exposure, so commonly used measures of exposure are imprecise and geographically coarse. However, pollution concentrations can change dramatically over short distances. Pollution spikes around highways and polluting firms, particularly in the area just downwind of the pollution source. Coarse measurements of pollution exposure are unable to account for granular changes in exposure or the wind-driven distribution of pollution, resulting in significant measurement error and biased regression estimates.

Moreover, the nature of this measurement error is such that natural experiments do not necessarily remove the resulting bias. For example, in a geographic difference-in-differences research design which uses distance to define treatment status, the assumed treatment and control groups are contaminated because most pollution travels downwind, making many "control" households heavily treated and vice versa. Similarly, interpolations of pollution monitor data, which are often used to measure local exposure, smooth over the many local spikes in pollution caused by firms

1. For example, Chay and Greenstone (2005) report a marginal willingness to pay to reduce pollution in line with Smith and Huang (1995). See Section 1 for further discussion.

2. Neidell (2009) and Moretti and Neidell (2011) find that attendance at outdoor attractions drops precipitously in response to pollution alerts. Qin and Zhu (2015) find that Internet searches in Chinese cities for "emigration" spike on high pollution days.

located between distant monitors. This leaves instruments which are based on firm location or the monitor averages themselves correlated with the measurement error.

I solve these problems by measuring local exposure to air pollution with an atmospheric dispersion model which can account for sharp changes in pollution and the meteorological forces that drive its dispersion. The model, AERMOD, was developed by atmospheric scientists with the American Meteorological Society and the U.S. Environmental Protection Agency (EPA) and uses extensive data on meteorological conditions and pollution sources to describe how a pollutant is dispersed around its source. Because it explicitly considers individual pollution sources, AERMOD captures the sharp changes in pollution exposure that happen around each source. With comprehensive data on houses and administrative data on all polluting firms in greater Los Angeles, I am able to map the behavior of every firm to the air quality of every house in the region.

I use this atmospheric science–based measure of exposure in a quasi-experimental hedonic framework to answer three questions. First, I estimate the effect on house prices of a large decrease in exposure to industrial NO_x emissions and the associated marginal willingness to pay (MWTP) for pollution reduction. Second, I calculate the implied social value of RECLAIM, a cap-and-trade program for NO_x in southern California which forms the basis of my natural experiment. Third, I re-estimate the house price effects using conventional measures of pollution exposure to test whether these measures are indeed biased downward. For my natural experiment I use the California Electricity Crisis of 2000 which precipitated a permit shortage in RECLAIM. This caused permit prices to skyrocket and firms in the area to quickly adopt abatement technology, suddenly and permanently lowering their emissions.

I find that house prices are very responsive to air quality, much more so than previous findings would suggest, and that RECLAIM led to substantial welfare gains. The average house in the sample area gained an additional $7,800 in value due to the Crisis. The implied MWTP to reduce exposure to industrial NO_x emissions is $3,306 per unit of reduced exposure, whereas past estimates have ranged around $200 per unit.[3] This MWTP implies that RECLAIM, whose social value has long

3. It should be noted that past studies have focused on the price effect of particulate matter rather than NO_x, and care should be taken comparing estimates across pollutants. Section 6.2 includes a

been contested, generates a net social benefit of almost $500 million annually.[4]

However, when using conventional methods for measuring pollution exposure, I am unable to detect any effect the Crisis had on house prices. I estimate geographic difference-in-differences models with various treatment radii (as in Currie and Walker 2011; Hanna and Oliva 2015; Currie et al. 2015; and others) and models assuming uniform radial dispersion of pollution (as in Banzhaf and Walsh 2008). I also estimate instrumental variables models, based on the geographic diff-in-diffs, that use interpolations of pollution monitor readings as the endogenous regressor (as in Hanna and Oliva 2015; Schlenker and Walker 2016; and others). None of the estimates are statistically or economically significant, with many being wrongly signed, suggesting that the much larger price effects found with AERMOD are due to methodology and not this particular sample or natural experiment.

1 The Puzzle of Clean Air's Low Valuation

House prices have long been used to measure the marginal willingness to pay (MWTP) for non-market goods. The MWTP for pollution abatement has been measured this way many times, starting with Ridker and Henning (1967).

The current body of literature suggests that house prices do not respond much to pollution, implying a disparity between the MWTP for pollution reductions and the expected health benefits. In their meta-analysis of OLS estimates of MWTP, Smith and Huang (1995) find that the interquartile range of estimates is $0 to $233 per microgram of Total Suspended Particulates (TSP) per cubic meter of air (or μg/m^3 of TSP).[5] They also find that the mean estimate, $259 per μg/m^3, only covers 6–33% of the associated VSL-based mortality cost. More recent instrumental variables estimates have not narrowed this disparity. Chay and Greenstone (2005) use counties' non-compliance with the National Ambient Air Quality Standards (NAAQS), county-level house prices, and average county pollution monitor readings to estimate a MWTP of $191 for a 1 μg/m^3 reduction in TSP, well within Smith

more thorough discussion of this point.

4. Details of this calculation, including social costs considered, are given in Section 7.

5. All dollar values presented here are deflated to 2014 dollars using the all-items CPI unless otherwise noted.

and Huang's interquartile range.[6] Bayer, Keohane, and Timmins (2009) also use county-level data on house prices and pollution. Instrumenting for local pollution with pollution from distant sources, they estimate a MWTP of \$131 per μg/m^3 reduction of PM_{10}.[7]

This ostensible disconnect between air pollution's costs and house prices is unusual because prices readily respond to other location-specific amenities. Cellini, Ferreira, and Rothstein (2010) use house price responses to bond override elections and estimate the average household is willing to spend \$1.50 for a \$1 increase in school capital expenditures. Linden and Rockoff (2008) find that when a registered sex offender moves into a neighborhood, the value of nearby houses drops by about \$7,000, more than the FBI's estimates of victimization costs would suggest. Davis (2004) looks at how prices respond to the appearance of a cancer cluster in Churchill County, Nevada, where the rate of pediatric leukemia suddenly skyrocketed for unknown reasons. The price response there implies the welfare cost of pediatric leukemia is about \$7 million, in line with estimates of the value of a statistical life from Aldy and Viscusi (2008).

Given the proclivity of house prices to capitalize the value of nearby amenities, the muted price response to air pollution is made even more puzzling by households' strong revealed preferences for clean air in other contexts. Neidell (2009) and Moretti and Neidell (2011) find that attendance at outdoor attractions like sporting events drops precipitously in response to smog alerts. Qin and Zhu (2015) find that Internet searches for "emigration" spike in Chinese cities on high pollution days.

2 Econometric Problems Behind the Puzzle

As an economic variable, exposure to air pollution is unusual because, unlike wages or education, there are no large-sample data on individual-level pollution exposure. Instead researchers approximate or infer exposure levels. Two such approximations are particularly common in the economics literature (see Currie et al. 2014). The first is to use a geographic difference-in-differences design where people close to a

6. Taken from the preferred specification in Chay and Greenstone (2005), Table 5A, column 4.

7. This estimate is based on Bayer, Keohane, and Timmins (2009), Table 6, column 2 and assumes costless migration, which is standard in the literature. They also fit a structural model that allows for costly migration, which yields a MWTP estimate of \$352.

pollution source are assumed to be “treated” by the source’s pollution while those slightly farther away are not treated but otherwise a good counterfactual. The second approach is to use the EPA’s network of pollution monitors to proxy for person-, neighborhood-, or county-level exposure, usually by interpolating between monitors.

However, both of these methods suffer from the same problem: They are unable to capture sharp changes in pollution across short distances which biases regression estimates, even those founded on a quasi-experimental research design.

2.1 Bias in Geographic Diff-in-diff Estimates

In a geographic difference-in-differences design, people around a polluting firm or other pollution source are assigned to treatment and control groups based on their proximity to the firm. The econometrician chooses radius r_0 around the firm to define the treatment group and radius $r_1 > r_0$ to define the control group, reducing the problem to a standard diff-in-diff. The difference over time is taken around some shock to the firm’s pollution output, such as a policy change or other exogenous shock (e.g., Currie and Walker 2011; Hanna and Oliva 2015; Schlenker and Walker 2016). For practical reasons, geographic diff-in-diffs are often centered around more routine changes in firm behavior, such as the construction and opening of the firm itself (e.g., Davis 2011; Currie et al. 2015). A key advantage of the geographic diff-in-diff design is that it allows for the estimation of reduced-form effects of policy changes when the exposure data necessary for second-stage estimates is lacking, as is often the case (e.g., Currie and Walker 2011; Davis 2011; Currie et al. 2015).

The use of a geographic diff-in-diff to study air pollution is problematic because air pollution does not disperse from its source uniformly in all directions, nor is it confined to the neighborhood immediately around the firm. Pollution is blown in the direction of prevailing winds, and significant amounts can travel dozens of miles downwind. This wind-driven dispersion contaminates the geographic diff-in-diff’s circular treatment and control groups, with many individuals upwind in the treatment area having little to no treatment and many downwind in the control area being intensely treated.

To derive the resulting bias, consider a model where the true effect of a polluting firm f on arbitrary outcome y_{it} is

$$y_{it} = N_{it}\alpha + X_{it}\beta + \varepsilon_{it} \tag{1}$$

where X_{it} is pollution exposure to i at time t, N_{it} is exposure to non-pollution disamenities created by the firm, and $t \in \{0,1\}$ indexes time, with $t=0$ preceding some shock to the firm's emission rate and $t=1$ following the shock. Let r_{if} be the distance of i from f and assume that r_0 is chosen so that $r_{fi} > r_0$ implies $N_{it} = 0$. The reduced-form geographic diff-in-diff estimation equation is

$$y_{it} = \gamma_1 + \text{post}_t\gamma_2 + C_i\gamma_3 + \left(C_i \times \text{post}_t\right)\gamma_{\text{GD}} + \mu_{it} \tag{2}$$

where $C_i = \mathbf{1}\{r_{if} \leq r_0\}$ is a dummy variable for individuals living in the treatment area and $\text{post}_t = \mathbf{1}\{t=1\}$. If $X_{it} = 0$ for $r_{if} > r_0$, then $\hat{\gamma}_{\text{GD}}$ recovers the average effect of the pollution shock on people living in the treatment area.

However, this key assumption—the control group is not exposed to pollution—is violated if pollution is carried far downwind. The distribution of pollution around its source depends on meteorology and the source's physical characteristics. Exposure can be written $X_{it} = m_{ft} \cdot h(r_{fi}, \theta_{fi}; \boldsymbol{S}_f)$, where m_{ft} is firm f's emissions in time t and h is the probability density function that a unit of emissions ends up at distance r and direction θ relative to the firm. The vector $\boldsymbol{S}_f$ contains variables about the firm's polluting equipment (e.g., height of the smoke stack) and local meteorological conditions like wind speed and direction. If wind speed is high or the smoke stack is tall, a significant amount of pollution can travel well beyond the 1 or 2 miles generally used for the treatment radius r_0.[8]

The resulting bias can be derived from the diff-in-diff estimator:

$$\begin{aligned}\hat{\gamma}_{\text{GD}} =& \mathbb{E}[y_{it} \mid C=1, \text{post}=1] - \mathbb{E}[y_{it} \mid C=1, \text{post}=0] \\ & - \left(\mathbb{E}[y_{it} \mid C=0, \text{post}=1] - \mathbb{E}[y_{it} \mid C=0, \text{post}=0]\right)\end{aligned} \tag{3}$$

We can write the expected value of y_{it} conditional on i's treatment assignment in terms of the average exposure to the treatment group:

$$\mathbb{E}_i\left[y_{it} \mid C\right] = C \cdot \bar{N}_t^C\alpha + \left[C + \varphi(1-C)\right]\bar{X}_t^C\beta \tag{4}$$

8. Currie and Walker (2011) use an r_0 of 2 kilometers (1.24 miles). Davis (2011) uses values of 1 and 2 miles. Currie et al. (2015) use 0.5 and 1 miles. Hanna and Oliva (2015) use 5 kilometers (3.1 miles). Figure 1 compares actual exposure measured by AERMOD to 1- and 2-mile radii.

where $\bar{X}_t^C = \mathbb{E}_i\left[X_{it} \mid C=1\right]$ and $\varphi = \mathbb{E}_i\left[X_{it} \mid C=0\right]/\mathbb{E}_i\left[X_{it} \mid C=1\right]$ is the ratio of average exposure in the control and treatment groups, so average exposure in the control group is $\varphi\bar{X}_t^C$. Substituting Equation (4) into Equation (3) yields

$$\hat{\gamma}_{\text{GD}} = \underbrace{\left(\bar{N}_1^C - \bar{N}_0^C\right)\alpha}_{\text{Non-pollution Effect}} + \underbrace{(1-\varphi)}_{\text{Wind bias}} \cdot \underbrace{\left(\bar{X}_1^C - \bar{X}_0^C\right)\beta}_{\text{Pollution Effect}} \tag{5}$$

The first term captures the firm's non-pollution effects. As β is the coefficient of interest, the ideal research design would hold N_{it} constant over time, eliminating this term.[9] The second term is the pollution effect, multiplied by the contamination factor $(1-\varphi)$.

Thus, even with an ideal natural experiment that holds non-pollution effects constant over time, the estimated pollution effect is biased because the effect on the control group, which is also treated, cancels out some of the effect on the treatment group. The degree of bias is driven by the wind and other factors in $\boldsymbol{S}_f$ that shape the geographic distribution of pollution, h. For example, because φ increases with wind speed, the contamination factor $(1-\varphi)$ and $\hat{\gamma}_{\text{GD}}$ both become more negative as wind speed increases. Furthermore, because h need not be monotonic in distance r, φ need not be less than 1, meaning $\hat{\gamma}_{\text{GD}}$ could have the wrong sign.[10]

It is important to note that the dependence of φ on h (and $\boldsymbol{S}_f$ in particular) implies that the bias will vary by pollution source and location. Because bias increases with wind speed and greater Los Angeles is one of the least windy areas in the United States, any bias found in by study is likely to be a lower bound for bias in more windy areas. Similarly, when pollution is emitted close to the ground, more of it stays close to the source, keeping φ low. This suggests that estimates of the effects of cars (e.g., Currie and Walker 2011) may suffer from less bias. However, even car exhaust gets carried by the wind (Hu et al. 2009), and a low φ does not mitigate any separate biases, such as those introduced by monitor data. In addition, it is possible that estimates of health effects are not as biased as those for houses because houses

9. This is naturally not the case when the shock to the firm is the construction of the firm itself. In such cases, $\bar{N}_1^C > \bar{N}_0^C = 0$. Note also that as the wind gets stronger and $\varphi \to 1$, $\hat{\gamma}_{\text{GD}} \to \alpha\bar{N}_1^C$ and the geographic diff-in-diff recovers the *non*-pollution effects of the firm.

10. An example of the non-monotonicity of exposure with distance is given by Figure 1, which is caused by the height of the firm's smoke stack and the high temperature and buoyancy of the emitted gases. This also illustrates the importance of variables in $\boldsymbol{S}_f$ other than wind direction.

are fixed in space while people are mobile and thus exposed to the varying levels of pollution in their community.

Econometrically, this contamination problem is a common issue in program evaluation (e.g., Miguel and Kremer 2004) and can be fixed by accounting for the average treatment intensity of each group as in two-stage least squares. However, this requires a good measure of treatment intensity. As Section 2.2 argues, data from geographically sparse pollution monitors do not fit this criterion.

2.2 Bias from Pollution Monitor Interpolation

When data on a spatially correlated variable like rainfall is unavailable for all locations of interest, it is possible to exploit this spatial correlation and interpolate the missing values. Given data on the outcome of interest $\{y_i\}_{i=1}^N$, corresponding data on the spatial variable, $\{x_i\}_{i=1}^N$, are needed but unavailable. However, it is often the case that data are available at a small set of monitor locations, $\{x_m\}_{m=1}^M$. If x is correlated across space, so that $\text{cov}(x_i, x_m)$ is large when i and m are close to one another, the monitor data can be used to construct an interpolation $\tilde{x}_i = \sum_m w_{im} x_m$, where the weights w_{im} are determined by the interpolation technique used. Monitor interpolation is the predominant method used in the economics literature to measure pollution exposure, with inverse distance weighting (IDW) being the standard implementation since Neidell (2004) and Currie and Neidell (2005).[11]

Unfortunately, air pollution is poorly suited for interpolation because the correlation between a monitor reading x_m and exposure x_i can be greatly affected by the presence of pollution sources between m and i. Unlike rainfall and other natural phenomena, air pollution is created by many distinct sources like firms and cars. This creates sharp changes in pollution concentrations over very short distances, such as just upwind versus downwind of a factory. This in turn means the correlation between x_i and x_m depends on more than just distance.

Consider $\text{cov}(x_i, x_m)$ when a polluting firm sits between i and m. If the wind blows toward one, the other will see a much smaller portion of the firm's pollution,

11. IDW uses $w_{im} = d_{im}^{-1} / \sum_{m'} d_{im'}^{-1}$ where d_{im} is the distance between i and m. Another common measure of exposure is the average of a county's monitors, essentially a flat interpolation across the county.

creating significant differences between x_i and x_m. Thus, $\mathrm{cov}(x_i, x_m)$ depends on the direction and speed of the wind, the height of the firm's smoke stack, and a host of other information not contained in the monitor data alone, making x_m of little use in predicting the value of x_i unless m and i are very close. This single-firm example scales to a world with many firms and monitors. The monitors do not pick up any variation in pollution when the source is not upwind of the monitor, and interpolating between monitors smooths over most of the local spikes in pollution exposure that exist around firms.

The way independent pollution sources segment the distribution of x across space creates a missing information problem whose severity is proportional to the geographic density of the monitors relative to the density of sources. In the extreme case with many monitors around every firm, there will be few instances where i is separated from all monitors by a firm and sufficient data will exist to accurately describe the distribution of x. This is the approach taken by atmospheric chemists who temporarily lay down dense monitor arrays (e.g., every 100 meters) to study dispersion patterns around a particular source (e.g., Perry et al. 2005). On the other extreme, with only a few monitors and many firms, x_i may be completely unrelated to all monitor readings and values interpolated from the monitors will be no better than noise. Empirically, the situation in the United States is much closer to the latter scenario with few monitors. According to the EPA's AirData summary files, the average county had 1.01 monitors in 2005, with almost two-thirds of counties having zero monitors. In the Los Angeles area specifically, one of the most studied areas for air pollution in the United States, there are hundreds of firms for every pollution monitor and monitors are spatially sparse, as shown in Figure A3.

Because the problem with interpolation is that the monitor data lack sufficient information, no interpolation technique is able to overcome it. This includes Kriging, a more advanced interpolation technique and the best linear predictor of x_i when certain assumptions about the spatial covariance of x_i hold (see Cressie 2015). The strength of Kriging is that it uses the monitor data to explicitly estimate a spatial covariance function for x_i to determine the proper interpolation weights w_{im}. However, as argued above, data from a sparse monitor network cannot accurately describe the spatial covariance of pollution with many discrete pollution sources.

2.2.1 Evidence of Interpolation Bias in Prior Research

The problems with interpolation described above are evidenced in prior literature.

First, the correlation of interpolated exposure and actual exposure appears to be low after controlling for secular temporal correlation. When using interpolated pollution measures, it is common practice to assess the quality of the interpolation using a leave-one-out cross-validation technique. Pollution exposure at each monitor is predicted using all other monitors, $\tilde{x}_m = \sum_{m' \neq m} w_{mm'} x_{m'}$. Then the correlation of the predicted and true values, $\text{corr}(\tilde{x}_m, x_m)$, is used to gauge the quality of the interpolation. These correlations can be quite high, often above 0.9.[12] However, the correlation of x_m and $\tilde{x}_m$ reported is usually unconditional and captures not just spatial but temporal correlation which will be partialed out in regressions. Consider the extreme case with no spatial correlation, where $x_{it} = \delta_t + \varepsilon_{it}$; δ_t is a time shock common to all locations such as regular seasonal variation; and ε_{it} is a zero-mean i.i.d. white noise term. In this case, $\text{corr}(\tilde{x}_{mt}, x_{mt})$ will be non-zero and potentially large, while the correlation conditional on t—the more relevant value for analyses with time controls—will be zero.[13] This is consistent with Karlsson and Ziebarth (2016), who find that the correlations for pollution IDW interpolations fall precipitously with time controls, from 0.6–0.9 to 0.15–0.4, while weather variables, which are smoother over space, do not exhibit this problem.[14]

Second, the smoothing over of local spikes in pollution around pollution sources should lead to non-classical measurement error in interpolated values, with $\tilde{x}_{it}$ being too low for larger x_{it}. Write $\tilde{x}_{it} = x_{it} + \eta_{it}$ where η is the interpolation error. If the interpolation smooths over variation in x, it will be true that $\text{Var}(\tilde{x}_{it}) < \text{Var}(x_{it})$, which implies $\text{cov}(x_{it}, \eta_{it}) < 0$.[15] Knittel, Miller, and Sanders (2016) plot $\hat{\eta}_{mt}$ and x_{mt} from the cross-validation exercise of their IDW interpolation and find that $\tilde{x}_{mt}$ is indeed increasingly too low for higher values of x_{mt}.[16] They also report that the

12. Currie and Neidell (2005) report cross-validation correlations of 0.92 for ozone, 0.77 for PM_{10}, and 0.78 for CO.

13. See Appendix A for derivations.

14. See Table F1 in Karlsson and Ziebarth (2016).

15. This follows from $\text{Var}(\tilde{x}_{it}) - \text{Var}(x_{it}) < 0$ and the definition of $\tilde{x}_{it}$.

16. Lleras-Muney (2010) also presents evidence of non-classical measurement error in Kriging interpolation, showing that the Kriging standard error increases with x_{mt}. However, she does not report whether the measurement error is increasingly positive or negative.

magnitude and sign of the interpolation error is uncorrelated with the distance to the nearest monitor. Karlsson and Ziebarth (2016) complete a similar exercise for temperature, which is smoother over space than pollution, and find that $\hat{\eta}_{mt}$ and x_{mt} are uncorrelated.

2.2.2 Interpolation bias persists in quasi-experimental designs

While attenuation due to measurement error is often resolved by using instrumental variables, this is only true if the measurement error and the instrument are uncorrelated. Let z be an instrument such that $\mathrm{cov}(x,z) \neq 0$ and $\mathrm{cov}(y,z) = 0$ and let $\eta = \tilde{x} - x$ again be the interpolation error. From the canonical probability limit of the IV estimator, we get

$$\operatorname{plim} \hat{\beta}_{\mathrm{IV}} = \beta \cdot \frac{\mathrm{cov}(x,z)}{\mathrm{cov}(x,z) + \mathrm{cov}(\eta,z)} = \beta \cdot \frac{\mathrm{cov}(x,z)}{\mathrm{cov}(\tilde{x},z)} \tag{6}$$

Note that the asymptotic bias could be positive or negative depending on the joint distribution of (x,z,η) which will vary across research designs.

First, consider the case of a geographic diff-in-diff, which assigns treatment status to those near pollution sources. As discussed above, if pollution sources significantly outnumber monitors, then the pollution spikes caused by many sources will be smoothed over, causing the measurement error to spike near the source. If the treatment variable is an indicator for "near the source", then the treatment variable is clearly correlated with the measurement error and $\hat{\beta}_{\mathrm{IV}}$ is inconsistent. However, the signs of the covariances in Equation (6), which determine the sign of the asymptotic bias, depend on the joint distribution of (x,z,η) which will vary from case to case with the number and location of monitors, the spatial distribution of the study population, and other factors.

Next, consider county-level studies using the Clean Air Act (CAA) as a natural experiment. In these studies, $\tilde{x}_{it}$ is usually the average of a county's monitors and is assumed to represent average exposure in the county. The CAA established limits (National Ambient Air Quality Standards or NAAQS) on county-level pollution as measured by the county's average monitor readings, making the regulatory metric identical to $\tilde{x}_{it}$. If a county's $\tilde{x}_{it}$ exceeds the NAAQS it is in "non-attainment" and local regulators are given additional authority to limit local emissions to lower $\tilde{x}_{it}$.

Thus, the onset of the NAAQS resulted in exogenous changes in local pollution as non-attainment counties suddenly faced additional regulatory pressure while the remainder did not.[17]

Such a research design is likely biased downward because the instrument is more closely related to $\tilde{x}$ than x because regulators specifically target $\tilde{x}$ rather than x. Monitors are not sited within a county to form a representative sample of population exposure, and there is evidence that local regulators strategically site monitors to reduce the likelihood of their county violating the NAAQS (Grainger, Schreiber, and Chang 2016).[18] In addition, Auffhammer, Bento, and Lowe (2009) find that regulators put more effort into reducing pollution levels at problematic monitors within a county, resulting in uneven treatment across monitors and the county. This means that the CAA policy shock affects $\tilde{x}$ more than x which, as Equation (6) shows, leads to downward biased estimates.

3 Measuring Exposure with a Dispersion Model

Atmospheric dispersion models solve the problems described above by explicitly accounting for the sudden changes in pollution exposure around firms and the way pollution is distributed by meteorological forces.

A dispersion model uses data on a polluting firm and the meteorology around it to predict the impact of the firm's pollution on air quality at nearby locations. Recall from Section 2.1 that exposure at location i to firm f's pollution can be written $x_{ift} = m_{ft} \cdot h(r_{fi}, \theta_{fi}; \boldsymbol{S}_f)$, where h is a probability density function over locations (r, θ) for pollution emitted by f. This distribution over space depends on $\boldsymbol{S}_f$, the firm's characteristics (e.g., stack height) and surrounding meteorology. An atmospheric dispersion model is a model of h developed by atmospheric chemists and validated with controlled experiments.[19] With knowledge of h and data on m_{ft} and $\boldsymbol{S}_f$, x_{ift} can be calculated for any arbitrary location (r_{if}, θ_{if}), as can total

17. See, e.g., Chay, Dobkin, and Greenstone (2003) and Chay and Greenstone (2003, 2005).

18. A related problem with monitor averages is that the relationship between the monitors and the true exposure distribution will change over time because monitors are fixed in space while people and firms vary their behavior and location over time.

19. Validation experiments are conducted by placing a dense network of several dozen monitors around a firm, releasing a rare, non-reactive tracer chemical, then comparing model predictions to monitor readings. For example, see Perry et al. (2005).

exposure, $x_{it} = \sum_f x_{ift}$. Most importantly, by explicitly accounting for the local distribution of pollution around every firm, exposure based on a dispersion model does not suffer interpolation's missing information problem.

In this paper, I use AERMOD, the EPA's legally preferred model for short-range applications. This preference is based on the model's high accuracy as established by peer-reviewed field tests (Perry et al. 2005).[20] To account for meteorological conditions, AERMOD uses data on temperature, mean and standard deviation of wind speed and wind direction at multiple elevations; the standard deviation of vertical wind speed; the convectively and mechanically driven mixing heights; and other parameters.[21] AERMOD also accounts for each smoke stack's height and diameter, the temperature and velocity of the gas exiting the stack, and the rate at which the pollutant in question is emitted from the stack (mass per unit time). Given these data, the model outputs the concentration of pollution at a location in micrograms per cubic meter of air ($\mu g/m^3$).

Calculating location-specific exposure using AERMOD and plotting it for the analysis sample in metro Los Angeles makes the problems described in Section 2 more apparent.[22] Figure 1 shows that ignoring the complex distribution of pollution around a firm causes geographic diff-in-diffs to have contaminated control samples and to miss the exposure effects for large portions of the population. The figure shows the average exposure to NO_x emitted by the Scatterwood Generation Station in Los Angeles in 1999, with circles drawn at 1 mile and 2 miles to represent the geographic diff-in-diff treatment and control radii described in Section 2.1. Pollution exposure is significantly higher to the northeast, the direction of prevailing winds, with high concentrations at 5 and even 10 miles downwind, well beyond the 2-mile control restriction. Furthermore, the area with the lowest exposure in the 2-mile sample area is actually in the "treatment" area, right next to the firm.[23]

Figure 2 shows how quickly pollution levels can change over short distances,

20. Regulatory preference is stated in 40 CFR pt. 51, app. W (2004). See Cimorelli et al. (2005) for a rigorous development of the model itself.

21. A full list of the variables used is found in the AERMOD user manual or Cimorelli et al. (2005).

22. Section 5 describes these data and how I implement the AERMOD model.

23. This is because hot gases are buoyant and can travel considerable horizontal distance before reaching the ground, especially when released from a tall smoke stack.

undermining the usefulness of monitor data. It plots exposure to NO_x from all major firms across the sample area in metro Los Angeles, as well as the locations of pollution monitors. This map shows a great deal of variation in pollution, with far more spikes in local exposure than monitors available to measure them.

Figure 3 further highlights the over-smoothing problem that results from interpolation by taking the exposure values in Figure 2 and interpolating between the marked monitor locations. Panel A uses inverse distance weighting (IDW) and Panel B uses the more advanced Kriging procedure.[24] In both cases, most of the spatial heterogeneity is gone, and areas that differ by an order of magnitude in Figure 2 are assigned the same exposure by the interpolations.

Together, these figures help explain the difficulties in measuring pollution exposure and help explain some contradictory results in the current literature regarding the importance of wind. Of the economics papers to address the question of wind and industrial pollution, only one, Hanna and Oliva (2015), finds that wind significantly alters their estimates, and then only in certain specifications.[25] This is likely due to the complexity of atmospheric dynamics which include many factors beyond wind direction and which affect not just the direction but the distance pollution travels from its source.

To further validate the peer-reviewed AERMOD model, I compare AERMOD's predictions against contemporaneous monitor readings in Figure 4. Panel A plots the AERMOD-predicted exposure to NO_x over time at the northern monitor in the

24. The inverse distance weighting used here imposes zero weight on monitors farther than 15 km from the point being interpolated. Such a restriction is commonly used in the literature to prevent interpolated values from being based exclusively on far away monitors (see, e.g., Hanna and Oliva 2015). The Kriging procedure used here is simple Kriging with an exponential variogram.

25. Hanna and Oliva (2015) look at how labor supply in Mexico City responded to a drop in pollution after the closure of a large refinery. They include the local elevation and a linear measure of degrees downwind in some specifications. Davis (2011) estimates the effect of plant openings on nearby house values and includes dummy variables for "upwind" and "downwind" in a robustness check. Contrary to expectations, he finds that houses upwind of plants have slightly lower prices. Schlenker and Walker (2016) measure the change in daily hospital visits due to changes in airport traffic and incorporate wind speed and direction into one of their models, with no substantive difference in results. Luechinger (2014) compares county-level infant health before and after the mandated desulfurization of power plants in Germany. He calls a county "downwind" of the power plant if it falls in the same 30-degree arc as the prevailing wind direction and includes downwind dummies in all his specifications.

sample area (see Figure 2) along with the actual monitor readings from that monitor. Panel B plots the same for the southern monitor. The plotted values are averages from the fourth quarter of each year because the AERMOD and monitor readings are most comparable at this time due to the decreased number of atmospheric chemical reactions during this time of year; these reactions are discussed in more detail below.[26] Figure 4 shows a strong similarity in AERMOD and monitor patterns over time. What differences do exist are likely due to atmospheric chemistry, other sources of NO_x like cars, or limitations of the meteorological data discussed in Section 5.

A final caveat about this measure of exposure is that it does not account for chemical transformations of the emitted NO_x. Pollutants often react with other chemicals in the atmosphere after being emitted. In particular, NO_x can combine with free oxygen to form ozone which is not emitted directly by polluters and is only present at ground level as a product of NO_x-based reactions. Though AERMOD and other models are capable of modeling this chemical process, it requires high-quality data on pre-existing levels of many other pollutants.[27] Because of the lack of such data, I am unable to confidently model the NO_x–ozone process. This means AERMOD predicts "exposure to NO_x emissions", which potentially includes ozone, rather than "exposure to NO_x." While this makes interpreting AERMOD more difficult from a biochemical point of view, this actually makes it a more comprehensive and policy-relevant metric because NO_x emissions are the object of regulation at firms.

26. It should also be noted that each variable is measured in different units. Because firm-level monitoring tracks mass of NO_x emitted (total grams of NO and NO_2) AERMOD measures local exposure in units of mass per volume of air (µg/m^3). In contrast, monitors measure the number of NO and NO_2 molecules relative to other molecules in the air (parts per million). It is generally possible to convert between these two units using the ideal gas law. However, RECLAIM's monitoring systems do not differentiate between NO and NO_2 and the relative ratio of these chemicals is crucial to converting between µg/m^3 and ppm due to their different molecular masses. Given this limitation of the data, and the fact that the NO/NO_2 mix varies both across firms and across time within firms, it is best to compare the AERMOD predictions and monitor readings as is.

27. While UV light is a main part of the NO_x–ozone reactions, they also depend on a class of chemicals called volatile organic compounds, or VOC's. The rate of NO_x–ozone conversion also depends on the relative ratios of NO, NO_2, and ozone. See Sillman (1999).

4 Theory and Research Design

4.1 House Prices and Willingness to Pay

I use hedonic valuation to test whether households value clean air. When choosing a place to live, households weigh a location's amenities, $\boldsymbol{g}$, against the bundled price of those amenities, $P(\boldsymbol{g})$. Rosen (1974) noted that utility-maximizing agents will choose a bundle of amenities and prices $(P(\boldsymbol{g}^*), \boldsymbol{g}^*)$ so that their marginal willingness to pay for each $g_k \in \boldsymbol{g}$ is equal to the corresponding marginal price, P_{g_k}.[28] Estimating average MWTP, which is difficult to do directly, can thus be accomplished by estimating P_{g_k}.

Using capitalization effects to estimate marginal prices and MWTP requires some assumptions. First, in order to identify P_{g_k} using intertemporal variation in house prices, the shape of P, which is endogenously determined in equilibrium, must be constant over the sample period (Kuminoff and Pope 2014). While this assumption is less palatable for longer sample periods and low-frequency data, it is likely to hold when using a short sample period and high-frequency data. Second, agents choose $(P(\boldsymbol{g}^*), \boldsymbol{g}^*)$ endogenously, potentially creating an omitted variables problem (Bartik 1987; Epple 1987). Any attempt to identify P_{g_k} must address this and satisfy the identification assumptions specific to the chosen research design. I address this problem by using the California Electricity Crisis of 2000 as a natural experiment and outline the necessary assumptions below.

4.2 Electricity Crisis as Natural Experiment

In 1994, the South Coast Air Quality Management District (SCAQMD), which regulates air pollution in Los Angeles, Orange, San Bernardino, and Riverside Counties, instituted a cap-and-trade program for NO_x emissions called RECLAIM.[29] Firms were given an initial allocation of year-specific RECLAIM Trading Credits (RTCs) for each upcoming year. Every year, firms must surrender one RTC for every pound of NO_x emitted. Excess RTCs can be sold to other firms but not banked for

28. There are a number of theoretical frameworks that can be used to estimate MWTP. See Palmquist (2005) and Kuminoff, Smith, and Timmins (2013) for summaries of valuation using hedonic pricing and equilibrium sorting models.

29. For additional details about RECLAIM, see Fowlie, Holland, and Mansur (2012).

future years. To ease firms' transition into the program, SCAQMD set the aggregate number of RTCs to be high at first and gradually decrease. It was anticipated that without firm adjustment total emissions would exceed total RTCs around 1999.

However, firms did not adequately adjust to the decreasing RTC cap. To avoid exceeding the cap, some firms would need to lower emissions by decreasing production or installing abatement equipment to remove NO_x from their emitted smoke. But RTC prices were so low there was little short-run incentive to abate. Some firms even canceled orders they had placed for abatement equipment prior to RECLAIM. SCAQMD reported in mid-1998 that abatement installations were lagging behind what was necessary to avoid the coming "cross-over point" when emissions would exceed permits. Firm managers later said they believed other "companies were reducing their emissions or were going to begin installing [abatement equipment], and as a result believed that they would be able to buy credits. . . [and] that long-term RTC prices would continue to stay low or would at least gradually rise to the cross-over point" (EPA 2002, p. 24).

This failure to anticipate increased RTC prices caused the cap-and-trade market to nearly collapse at the onset of the California Electricity Crisis in mid-2000. The heart of the Crisis was that existing electricity generators struggled to meet demand.[30] To prevent rolling blackouts, many electricity producers significantly increased generation and, as a result, their NO_x emissions. This caused the RTC cap to finally bind which in turn caused a dramatic spike in RTC prices, from $2,800 per ton in 1999 to $62,000 by the end of 2000 (see Figure A1).

Firms not generating electricity responded by finally installing abatement equipment, leading to a permanent decrease in the average firm's emissions of almost 40%. This sudden drop is shown by the solid line in Figure 5 which plots the annual average of firm emissions scaled by own-firm sample maximum to give each firm equal weight. The dashed lines show that emissions from electricity generators also fell to roughly 50% of pre-Crisis levels once the Crisis subsided.

The permanence of these pollution reductions, despite the temporary nature

30. The exact causes of the Crisis, such as the deregulation of wholesale electricity markets and market manipulation by certain actors, remain a source of debate. See Borenstein (2002) and Weare (2003), especially Section 3.

of the Crisis, is due to the permanence of the RECLAIM cap-and-trade market. RECLAIM's permanently binding emissions cap pushed firms to find permanent abatement solutions. Had firms fully anticipated the eventual binding of the cap, the Crisis may not have caused a sharp change in emissions behavior. Instead, the Crisis synchronized the long-term adaptation to the cap.

The sudden, permanent drop in emissions that followed the Crisis can be used to construct a set of instruments for local residents' exposure to firms' pollution. When faced with high RTC prices, high-emission firms had a larger incentive to cut emissions so the Crisis should have had a larger effect on houses downwind of these firms. A house's pre-Crisis exposure to emissions can thus be used to gauge its exposure to the effects of the Crisis relative to other houses. Using aermod_{it}, the AERMOD-predicted exposure to house i in time t, I define pre-Crisis exposure aermod_pre_i as the average exposure across all 8 quarters in 1995 and 1996, the first two years of firm-level emissions data. With aermod_pre_i as a measure of treatment exposure, a variable intensity diff-in-diff instrument can be constructed: $\text{aermod\_pre}_i \times \text{post}_t$ where $\text{post}_t = \mathbf{1}\{t \geq 2001\}$ is an indicator variable for post-Crisis years. The corresponding event study instruments, $\text{aermod\_pre}_i \times \delta_y$ where δ_y is a dummy variable for year y, capture the differential effects of the Crisis on house i in year y relative to the omitted year. These can be used to test the common trends assumption underlying the diff-in-diff.

The identification assumption behind these instruments is that there are no coincidental changes in house prices or non-industrial pollution exposure that are correlated with the instruments, conditional on the other covariates. For example, the housing bubble might have induced more appreciation in poorer neighborhoods which may be relatively more polluted before the Crisis due to residential sorting. Fortunately, we can explicitly control for time trends in such risk variables, and the build up of the bubble was not a discrete event like the Crisis was, so this assumption can be assessed using the event study. Another potential problem is that the instruments might be correlated with changes in NO_x from cars. This would bias second-stage estimates upward if industrial exposure were correlated with automobile exposure *and* the Crisis also caused a sudden and permanent drop in car usage in the area. The former condition is unlikely given the large area that firms affect, while highways rarely have a significant impact beyond 500 meters (Karner, Eisinger, and Niemeier 2010). Furthermore, traffic data show that no significant change in driving patterns coincided with the Crisis.[31]

4.3 Estimation Strategy

The marginal price of pollution exposure can be estimated using the following model:

$$\ln p_{it} = \beta \cdot \text{aermod}_{it} + \alpha_i + \delta_t + \sum_k \gamma_{1k} \cdot w_{ik} \cdot t + \sum_k \gamma_{2k} \cdot w_{ik} \cdot t^2 + \varepsilon_{it} \tag{7}$$

where p_{it} is the price of house i in quarter t; aermod_{it} is exposure to industrial NO_x emissions; α_i are house fixed effects; δ_t are time (quarter-year) effects; $(\gamma_{1k}, \gamma_{2k})$ are coefficients on quadratic time trends for local geographies, defined by a 10-km grid, and local economic conditions that might affect house prices (discussed below); and ε_{it} is the usual residual term. I estimate this equation using two-stage least squares (2SLS), with the primary specification using $\text{aermod\_pre}_i \times \text{post}_t$ to instrument for aermod_{it} as detailed in the previous section.

The additional controls included in Equation (7) account for a number of factors that may confound estimates of β, such as amenities not included in the available data and differential trends across local housing markets. The house fixed effects, α_i, capture of all time-invariant characteristics about the house like square footage, number of bedrooms, proximity to the beach, etc. The time effects, δ_t, account for general trends in the housing market over time, as well as seasonal trends within each year (e.g., if houses consistently sell for more during the summer). The local geographic trends allow different parts of the metropolitan area to have different secular trends.[32]

31. Unreported regressions show traffic patterns had no significant break from trend through the period of the Crisis. I use data from the California Department of Transportation's Freeway Performance Management System (PeMS) for the Bay Area (region 11), 1999–2005, because data for Los Angeles only go back to 2001.

32. Given the large size of the sample region, it would be natural for local trends to be defined by cities, which have economically meaningful boundaries (unlike zip codes) and are generally small but not so small as to be computationally burdensome (unlike tracts and zip codes). Unfortunately, many houses do not have a city listed in the data, and the cities of Los Angeles and Long Beach cover a large portion of the sample region while also having a great deal of within-city heterogeneity. To overcome these issues, I use a 10-km grid which is aligned to preserve as many city boundaries as possible. This grid results in 17 different areas that get their own quadratic time trend.

The local trends in socio-economic variables are specifically targeted at concerns related to the housing bubble, which differentially impacted neighborhoods with poor credit. Mian and Sufi (2009) find that zip codes with lower incomes and credit scores were affected more by the expansion of sub-prime credit. If these areas also experienced relatively bigger air quality improvements thanks to the Crisis, the coefficient on aermod_{it} could pick up any increase in house prices due to the expansion of sub-prime credit. To prevent this, I interact the following variables with quadratic time trends: the average loan-to-value ratio for houses sold in the house's census tract in 2000; the average predicted interest rate for mortgages taken out in the house's census tract in 2000; and the median household income in the house's census block group in 2000.[33] The predicted mortgage interest rate data was calculated by DataQuick using proprietary methods and is included in the house data described in Section 5.

I restrict the analysis to the period 1997–2005. RECLAIM's first full year of emissions data collection was 1995, and data from 1995 and 1996 are used to construct aermod_pre_i. Following Fowlie, Holland, and Mansur (2012), I set the last sample year to 2005. This avoids the peak and collapse of the housing bubble.

I restrict the region of analysis to the southwest part of SCAQMD territory, roughly between Santa Monica and Huntington Beach (see Figure 2), to minimize measurement error due to geography. Most major polluters are in this region, and locations farther away from the pollution sources are likely to have less actual exposure from the firms and a lower signal-to-noise ratio in aermod_{it}. Predicting the pollution distribution is also more complicated farther inland because of the San Gabriel and Santa Ana Mountains, which can act like a dam, collecting pollution blown by the prevailing winds from the south and west.

5 Data

Housing data come from county registrar and assessors' offices via DataQuick, Inc. The data include most sales in California since 1990. Data for each property includes square footage, lot size, number of bedrooms and bathrooms, and the year

33. The first two variables are averaged at the tract because the sample of transacted of houses in many block groups is very small.

of construction. Each sale includes the value of all loans taken against the property, as well as interest rates as estimated by DataQuick using proprietary methods. The median income of each house's Census block group is taken from the 2000 Census.

Sales outside normal market transactions are dropped since they may not accurately reflect the market's valuation of the house. Specifically, all transactions must be arms-length, non-distressed sales (i.e., no foreclosure sales or short sales) with a price of at least $10,000. Extremely high-value sales (the top 0.1%) are dropped. I also drop sales that occur within 90 days of a previous transaction, as many of these are duplications. The sample is also restricted to homes built before 1995 to preclude direct sales from developers to consumers. Table A1 shows summary statistics for houses in the sample, including sale price, property characteristics, number of times sold, etc. House prices are deflated to real 2014 dollars using the all-items CPI.

Most of the firm data come from SCAQMD via public records requests (SCAQMD 2015a). These data include each firm's name, address, SCAQMD-assigned ID number, the mass of NO_x the firm emitted every quarter from 1995 to 2005, and all relevant RTC data, including initial allocation of RTCs, the quantity, price, and vintage of exchanged RTCs. Firms' operating addresses were geocoded to get latitude and longitude to represent the location of the firm's smoke stacks. Firms' SIC info is taken from Fowlie, Holland, and Mansur (2012). Data on firms' physical characteristics (smoke stack height and diameter, and temperature and velocity of gas exiting the smoke stack) come from the National Emissions Inventory (NEI).[34] Firms were matched to the NEI using SCAQMD ID number, and firm name and address. Full details of the construction of the firm-level data are given in Appendix B. Table A2 gives summary statistics by 4-digit SIC on emissions, smoke stack parameters, average distance to the nearest meteorological station, and the number of firms in each industry group.

Data on local meteorological conditions come from SCAQMD and were gathered by 27 meteorological stations throughout the region.[35] The data include hourly observations for temperature, wind speed and direction, and other variables described

34. Regulators often collect these data specifically to run atmospheric dispersion models like AERMOD, but the data collected by SCAQMD could not be made available (SCAQMD 2015b).

35. The data are accessible via the SCAQMD website: http://www.aqmd.gov/home/library/air-quality-data-studies/meteorological-data/data-for-aermod

in Section 3. Each station provides at least three years of data between 2006 and 2012. While these stations were not in operation at the time of the Crisis, wind patterns at the given locations are very stable over time.

Air pollution monitor data come from the California Air Resources Board (CARB) and include hourly readings for NO_x and ozone in parts per million (ppm). I aggregate the hourly measures to daily and then monthly averages following Schlenker and Walker (2016). I exclude monitors that did not operate for the entire 1997–2005 sample period. The location of each meteorology and pollution monitor is shown in Figure A3.

I use AERMOD to construct a measure of a house's exposure from all industrial sources. Software implementing AERMOD is available on the EPA's website.[36] As discussed in Section 3, house i's exposure to NO_x emissions from firm f at time t can be written $x_{ift} = \text{NOx}_{ft} \cdot h(r_{fi}, \theta_{fi}; \boldsymbol{S}_f)$, where $\boldsymbol{S}_f$ contains information on the firm's smoke stacks and its surrounding meteorology. Meteorological data for $\boldsymbol{S}_f$ is taken from the meteorology monitor closest to the firm. Given these data and a house's location, AERMOD outputs aermod_{ift}, the house's exposure to the firm's emissions. The house's total exposure to industrial NO_x emissions is $\text{aermod}_{it} = \sum_f \text{aermod}_{ift}$.

Because AERMOD loops over all firms, houses, and hours of meteorological data, it is computationally intensive for such a large sample and so I impose several restrictions on the model for feasibility. First, I only calculate exposure to houses that are within 20 kilometers of a given firm and set exposure outside this radius to zero. Second, I use one year of meteorological data, 2009, which is also the only year during which all of the meteorological stations described in Section 5 were operating. Third, I construct an arbitrary 100-meter grid by rounding each house's UTM coordinates to the nearest 100 meters and calculate the exposure value at the center of each grid square. Houses are then assigned exposure values according to the grid square they occupy.

36. Fortran source code and executables for Windows are available at http://www.epa.gov/scram001/dispersion_prefrec.htm. I use AERMOD version 13350 compiled using Intel Fortran Compiler 15.0 for Linux and run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University.

6 Results

6.1 Event Study of the Crisis's Effects

Figure 6 plots the effects of the Crisis over time on both house prices and pollution exposure. This provides a visual test of the common trends assumption and the credibility of the natural experiment. It plots the estimated $\hat{\pi}_y$ coefficients from the equation

$$Y_{it} = \sum_{y \neq 2000} \left(\text{aermod\_pre}_i \times \delta_y\right) \pi_y + \alpha_i + \delta_t + \sum_k \gamma_{1k} \cdot w_{ik} \cdot t + \sum_k \gamma_{2k} \cdot w_{ik} \cdot t^2 + \varepsilon_{it}$$

where Y_{it} is either $\ln p_{it}$ (the reduced form) or aermod_{it} (the first stage) and all other controls are the same as in Equation (7). Each π_y captures the effect of the Crisis on house prices or pollution exposure in year y relative to the omitted year, 2000. With a valid natural experiment, we should see no effect before the Crisis ($\hat{\pi}_y \approx 0$ for $y < 2000$) with a sharp change immediately following it.

Figure 6 is strong evidence that the Crisis is a valid natural experiment, with little effect on exposure and prices before the exogenous shock of the Crisis and sharp effects immediately following the shock. The effect on house-level pollution exposure over time (the dashed line) unsurprisingly mimics the behavior of firm emissions shown in Figure 5, with a flat profile before the Crisis, a sudden drop right after the Crisis, and a slight negative trend going forward as firms complete their abatement solutions. The effect on house prices (the solid line) is a mirror image of the exposure effect, showing a flat profile before the Crisis followed by a sudden jump in value of houses with improved air quality. This suggests that the instrument based on aermod_pre_i is indeed capturing the effects of the Crisis-induced reduction in exposure rather than other secular changes. For example, if the instrument were instead picking up secular trends like the beginning of the housing bubble, Figure 6 would instead show a smooth, exponential-like increase in prices.

6.2 Instrumental Variables Estimates

Table 1 presents estimates of the causal effect of pollution exposure on house prices. The reduced-form estimates in columns 1 and 2 show that the Crisis-induced pollution reduction significantly increased house prices. Column 1 is the preferred specification based on Equation (7) with house-level fixed effects, year-quarter effects, and trends in local geographic and demographic characteristics. The coefficient of 0.0033 implies that every unit of initial exposure (i.e., treatment intensity in the Crisis) increased the sale price of a home by 0.33%. This coefficient is also precisely estimated, with a t-statistic of 3.94 (p-value less than 0.0001). Multiplying this estimate by the average treatment intensity, 5.331, gives the effect of the Crisis on the average house's value: 1.8% or $7,790 for the average home sold in 2000.

Column 2 is a robustness check for the preferred specification in column 1 which relaxes the house-level fixed effects in favor of block group fixed effects and explicit controls for house quality: interior square feet, lot size, number of bedrooms, and number of bathrooms. While unable to control for all time-invariant house characteristics, this specification allows for the inclusion of properties only sold once during the sample period and for the estimation of the aermod_pre_i main effect which is otherwise subsumed by the house-level effects. The estimate of the Crisis's effect in Column 2, 0.0033, is essentially the same as in column 1, and the coefficient on aermod_pre_i, -0.0027, is negative, confirming that properties initially exposed to more pollution were worth less.[37]

Column 3 presents the first-stage estimate that corresponds to the reduced form presented in column 1. The estimated coefficient of -0.4420 implies that for every unit of exposure to NO_x emissions in 1995–1996 (the basis for aermod_pre_i), roughly 43% of that exposure was removed by the Crisis and RECLAIM. This is consistent with the firm-level behavior shown in Figure 5, which shows a decrease in firm-level emissions of a similar magnitude. This relationship between *firm*-level emissions and *house*-level exposure is non-trivial because it depends on the geographic distributions of firms and houses, the differential behavior of firms, and meteorology. For example, a decrease in average emissions could be driven by firms far from population centers. This estimate shows that the exposure to houses did change significantly due to the Crisis and RECLAIM.

Column 4 presents two-stage least squares (2SLS) estimates of the causal effect

37. The correlation of initial pollution exposure and neighborhood characteristics and how neighborhoods changed demographically in response to the air quality improvement following the Crisis is the focus of Sullivan (2017).

of pollution exposure, with aermod_{it} as the endogenous regressor and $\text{aermod\_pre}_i \times \text{post}_t$ as the excluded instrument. The estimate of -0.0074 is again precisely estimated (t-stat 3.1, p-value 0.002) and implies that an additional unit of exposure (μg/m^3) to NO_x emissions decreases a house's value by 0.74%. Using the average sale price of homes in 2000, this translates to a MWTP to reduce exposure of $3,306 per unit.

Columns 5 and 6 are robustness checks on the preferred 2SLS estimate and show that it is robust to both the choice of instruments and the IV method used. Both columns replace the variable-intensity diff-in-diff instrument with the event study instruments used in Figure 6: $\text{aermod\_pre}_i \times \delta_y$ with year effects δ_y. Column 6 also uses LIML instead of 2SLS. These point estimates and standard errors are both remarkably close to the their counterparts in the preferred specification in column 4, suggesting that neither the choice of instruments nor the choice of IV method is driving the results.

Table 1 also provides evidence that the estimates do not suffer from significant weak-instrument bias. The partial F-statistics for both sets of instruments are large, 6,323 and 923, well above the usual rule-of-thumb of 10.[38] The LIML estimates in column 6 provide further evidence against weak instruments because the LIML estimator is median-unbiased and thus more reliable than 2SLS when instruments are weak (Stock, Wright, and Yogo 2002). The similarity of the LIML estimate to the 2SLS results in column 5 does not raise any concern about weak instruments.

Table A3 shows the robustness of the preferred reduced-form and second-stage estimates to alternative estimates of the standard errors and the inclusion of additional instruments. Columns 1 through 5 present estimates of the standard errors for columns 1 and 4 of Table 1 when clustering by Census block group, by tract, and when using the spatial HAC method of Conley (1999) and Kelejian and Prucha (2007) with a triangle kernel with bandwidths 0.25 miles, 0.5 miles, 1 mile, and 2 miles.[39] Column 6 adds the instrument "uniform_pre×post" where "uniform" is

38. This follows the common practice since Stock and Yogo (2002) and Stock, Wright, and Yogo (2002). However, the usual rules of thumb from Stock, Wright, and Yogo assume spherical error terms, so I follow Coglianese et al. (2017) and assume spherical errors when calculating partial F statistics.

39. For comparison, the median tract in the sample is roughly 0.75 miles across, while a spatial HAC with a 2-mile bandwidth allows arbitrary correlation within an area 4 miles across.

a measure of NO_x exposure assuming uniform dispersion within 2 km of the firm (see Section 6.3) and column 7 adds "Within 2 km of Firm" interacted with post. These instruments should capture any time variant effects of living near polluting firms that are not distributed by the wind which may be inflating the coefficients on pollution exposure. However, Table A3 shows that neither of these variables has any significant impact.

6.3 Comparison to Standard Methods

Section 2 argues that conventional methods of measuring pollution's impact will be biased due to the wind. I test this by re-estimating the effect of the Crisis using these conventional methods instead of AERMOD.

First, I follow Currie et al. (2015) and estimate a geographic diff-in-diff using a model similar to Equation (7) but where each house-firm pair is treated separately, effectively pooling the many firm-level diff-in-diffs:

$$\ln p_{ift} = \text{near}_{if} \times \text{post}_t \cdot \beta + \alpha_{if} + \boldsymbol{X}_{it}\boldsymbol{\Gamma} + \varepsilon_{ift} \tag{8}$$

where α_{if} are now house-firm effects instead of house effects; $\boldsymbol{X}_{it}$ includes the same controls as Equation (7); and near_{if} is a dummy variable for whether house i is within the chosen treatment radius of firm f. I estimate this model with treatment and control radii of 1 and 2 miles and again with 2 and 4 miles.

The resulting reduced-form estimates presented in columns 1 and 4 of Table 2 and are small, imprecise, and have different signs. For the 1-mile treatment, the average effect of the Crisis on log price is 0.0040, less than one fourth the size of the average effect found using AERMOD, 0.018. This estimate is also imprecise, with a standard error of 0.0050. The 2-mile estimate in column 4 implies that treated houses *lost* value after Crisis and is also imprecisely estimated.

The derivation of geographic diff-in-diff bias in Section 2 predicts that the first-stage and reduced-form estimates should have the same bias and that, with a good measure of exposure, the second-stage estimate should be unbiased though potentially noisy. I test this using the firm-specific exposure measure aermod_{ift} as the endogenous regressor. For the 1-mile treatment radius, the biases appear to be roughly equal. The reduced-form effect is 23.0% of the average reduced-form effect found in Table 1, column 1, while the first stage effect is 23.4% of its AERMOD-based equivalent. Consequently, the second stage coefficient, -0.0077, is very similar to the estimates in Table 1. However, this estimate is very imprecise and it is difficult to draw a strong conclusion about the estimates' similarity. For the 2-mile treatment, the reduced-form and first-stage estimates are wrong-signed, making comparison difficult. Ignoring signs, the ratios of reduced-form and first-stage effects are 0.07 and 0.01, respectively.

I also estimate geographic diff-in-diffs using interpolated NO_x and ozone from pollution monitors and present the results in Table A4.[40] None of the reduced-form or second-stage estimates is statistically significant, with many having the wrong sign and changing dramatically with treatment radius and choice of instruments.

The second conventional research design uses radial kernel densities to map firm emissions to local exposure. This is similar to the approach taken by Banzhaf and Walsh (2008), who use the equivalent of a uniform kernel with a 1 mile (1600 meter) bandwidth. I use a triangle kernel with 5-km bandwidth and a uniform kernel with 2-km bandwidth as the proxy for the spatial distribution h instead of AERMOD.[41] In both cases, the sample is restricted to houses within 5 km of a firm. The kernel approach should theoretically be an improvement over the geographic diff-in-diff because it can account for neighboring firms' overlapping treatment areas. Once again, the estimation equation is Equation (7), except that the exposure measure and instruments are constructed using the relevant kernel density instead of AERMOD.

The estimates, presented in Table 3, are generally small, imprecise, and sometimes wrong-signed. Columns 1 and 4 show the reduced-form estimates for the triangle and uniform kernels, respectively. The triangle estimate has the wrong sign while the uniform estimate is imprecise and small, implying an average treatment effect of only 0.01 percent. The first-stage estimates in columns 2 and 5 are very similar to their AERMOD counterpart due to the mechanical relationship between firm emissions and these exposure variables. Columns 3 and 6 show the second-stage

40. As before, the interpolation is calculated using inverse distance weighting using monitors with full NO_x and ozone coverage during the sample period that are no more than 15 km from the point being interpolated.

41. To make the unit-less kernel-based variables comparable to the AERMOD measure, I re-scale them so that their sample average is the same magnitude as the sample average of aermod_{it}.

results, which mirror the reduced-form results.

6.3.1 Summary and Comparison to Prior Research

Table 4 presents the average treatment effects and/or MWTP implied by the estimates in Tables 1 to 3 along side previously discussed estimates from the literature. This allows for a more direct comparison across the various methods. The first column lists the model or paper that generated the estimate; the second column lists the estimated effect of the Crisis on average house prices for models from this paper; and the third column lists the estimated MWTP for a 1 $\mu g/m^3$ reduction in pollution. Note that there are no MWTP estimates from the geographic diff-in-diff models because these models provide no measure of pollution exposure and thus no second-stage estimates.

Panel A gives the preferred AERMOD-based values from Table 1. Panel B gives the values for the geographic diff-in-diffs (Table 2) and the kernel-based dispersion (Table 3), which are generally small, statistically insignificant, and unstable across specification. With the geographic diff-in-diff, moving from a 1- to 2-mile treatment radius flips the sign of the estimate. The same happens in the kernel-based estimates when switching kernels. Taking the most positive estimates from each column, the largest conventional estimates for average price effect and MWTP are 23% and 4% of the AERMOD estimates, respectively.

Panel C gives the values from past literature. These estimates are comparable in magnitude to the most positive MWTP estimate from Panel B, though unlike Panel B, they are statistically significant with p-values no greater than 0.05. The estimates are also very close to one another, regardless of whether OLS or instrumental variables methods were used.

While Table 4 allows for a more apples-to-apples comparison of the various point estimates, care should be taken when considering estimates from prior literature because they represent the MWTP to reduce different pollutants. The pollutant of interest in Smith and Huang (1995) and Chay and Greenstone (2005) is TSP; in Bayer, Keohane, and Timmins (2009) it is PM_{10}; and in this paper it is NO_x emissions, which primarily take the form of NO_x or ozone. The market response to these pollutants could differ by the pollutants' toxicity and salience. However, the relative toxicity of NO_x emissions and particulate matter suggest that the biological harm of particulate matter is at least that of NO_x emissions, if not dramatically greater (see, e.g., Muller and Mendelsohn 2009). In addition, when NO_x emissions take the form of ozone, they are significantly less visible than most particulate matter. It is also true pollutants are highly correlated, and a decrease in one pollutant is likely being accompanied by a decrease in other pollutants, causing estimates of the effect of specific pollutants to capture some of the effects of other pollutants. However, the abatement solutions used by firms in this sample specifically target NO_x emissions through filtration, as opposed to, e.g., more efficient fuel usage that would decrease all pollutants. These facts suggest that the MWTP for NO_x estimated here is likely a lower bound for the MWTP to reduce particulate matter. Nevertheless, comparing estimates for different pollutants is difficult and a more reliable way to test for bias in conventional methodology is to compare methods using the same sample, such as the AERMOD estimates in Panel A and the conventional estimates in Panel B.

7 Conclusion

An accurate estimate of the social value of clean air is critical for setting efficient air quality policy. This paper presents evidence that inaccurate measures of air pollution exposure can lead to severely biased estimates of pollution's effects, even when a natural experiment is used. When using the atmospheric dispersion model AERMOD to measure exposure, I find that the California Electricity Crisis of 2000 significantly lowered exposure to NO_x emissions in metro Los Angeles and caused houses with improved air quality to increase in value by 1.8% on average. This price increase implies a MWTP to reduce exposure to NO_x emissions of \$3,306 per $\mu g/m^3$. When using conventional measures of pollution exposure, I find no statistically or economically significant effect.

A significantly higher social value of clean air has sweeping implications for air quality policy. For instance, the RECLAIM cap-and-trade program has long been questioned on cost-benefit grounds. However, the MWTP above implies that reducing emissions in SCAQMD from 1995 levels to the 2005 RTC cap is worth roughly \$524 million annually, far more than the estimated annual abatement costs of $39 million.[42] EPA's recent multi-year effort to tighten ozone standards is another example of a policy that was incorrectly undervalued and thus met stiff resistance on cost-benefit grounds.[43] More generally, the social welfare calculus for power generation is more likely to favor cleaner sources like solar and nuclear over coal. By extension, the co-benefits of reducing carbon emissions are also greater.

42. Abatement costs based on SCAQMD (2000) and do not consider other costs like worker displacement. SCAQMD asks firms to report how many jobs are lost or gained due to RECLAIM every year. Through 1999, firms reported a total net employment change of -109 workers which they attributed to RECLAIM (SCAQMD 2000). See Fowlie, Holland, and Mansur (2012) for summary of debate on RECLAIM.

43. See, e.g., "Obama Asks EPA to Pull Ozone Rule," *Wall Street Journal*, September 3, 2011; "EPA Sets New Ozone Standard, Disappointing All Sides," *New York Times*, October 1, 2015.

References

Aldy, Jospeh E., and W. Kip Viscusi. 2008. "Adjusting the Value of a Statistical Life for Age and Cohort Effects." *Review of Economics and Statistics* 90 (3): 573–581.

Auffhammer, Maximilian, Antonio M. Bento, and Scott E. Lowe. 2009. "Measuring the effects of the Clean Air Act Amendments on ambient PM_{10} concentrations: The critical importance of a spatially disaggregated analysis." *Journal of Environmental Economics and Management* 58 (1): 15–26.

Banzhaf, H. Spencer, and Randall P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism." *American Economic Review* 98 (3): 843–863.

Bartik, Timothy J. 1987. "The Estimation of Demand Parameters in Hedonic Price Models." *Journal of Political Economy* 95 (1): 81–88.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. "Migration and hedonic valuation: The case of air quality." *Journal of Environmental Economics and Management* 58:1–14.

Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics* 114 (2): 577–599.

Borenstein, Severin. 2002. "The Trouble with Electricity Markets: Understanding California's Restructuring Disaster." *Journal of Economic Perspectives* 16 (1): 191–211.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics* 125 (1): 215–261.

Chay, Kenneth Y., Carlos Dobkin, and Michael Greenstone. 2003. "The Clean Air Act of 1970 and Adult Mortality." *Journal of Risk and Uncertainty* 27:279–300.

Chay, Kenneth Y., and Michael Greenstone. 2003. "Air Quality, Infant Mortality, and the Clean Air Act of 1970." NBER Working Paper #10053.

———. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy* 113 (2): 376–424.

Cimorelli, Alan J., Steven G. Perry, Akula Venkatram, Jeffrey C. Weil, Robert J. Paine, Robert B. Wilson, Russell F. Lee, Warren D. Peters, and Roger W. Brode. 2005. "AERMOD: A dispersion model for industrial source applications. Part I: General model formulation and boundary layer characterization." *Journal of Applied Meteorology* 44 (5): 682–693.

Coglianese, John, Lucas W. Davis, Lutz Kilian, and James H. Stock. 2017. "Anticipation, Tax Avoidance, and the Price Elasticity of Gasoline Demand." *Journal of Applied Econometrics* 32 (1): 1–15.

Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of Econometrics* 92:1–45.

Cressie, Noel. 2015. *Statistics for spatial data.* John Wiley & Sons.

Currie, Janet, Lucas Davis, Michael Greenstone, and Reed Walker. 2015. "Do Housing Prices Reflect Environmental Health Risks? Evidence from More than 1600 Toxic Plant Openings and Closings." *American Economic Review* 105 (2): 678–709.

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell. 2014. "What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution?" *Annual Review of Resource Economics* 6 (1): 217–247.

Currie, Janet, and Matthew Neidell. 2005. "Air pollution and infant health: What can we learn from California's recent experience?" *Quarterly Journal of Economics* 120 (3): 1003–1030.

Currie, Janet, and Reed Walker. 2011. "Traffic Congestion and Infant Health: Evidence from EZPass." *American Economic Journal: Applied Economics* 3 (1): 65–90.

Davis, Lucas. 2004. "The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster." *American Economic Review* 94 (5): 1693–1704.

———. 2011. "The Effect of Power Plants on Local Housing Values and Rents." *The Review of Economics and Statistics* 93 (4): 1391–1402.

Epple, Dennis. 1987. "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products." *Journal of Political Economy* 95 (1): 59–80.

Fowlie, Meredith, Stephen P. Holland, and Erin T. Mansur. 2012. "What Do Emissions Markets Deliver and to Whom? Evidence from Southern California's NOx Trading Program." *American Economic Review* 102 (2): 965–993.

Grainger, Corbett, Andrew Schreiber, and Wonjun Chang. 2016. "How States Comply with Federal Regulations: Strategic Ambient Pollution Monitoring." Working paper. http://www.aae.wisc.edu/cagrainger/monitorpaper.pdf.

Hanna, Rema, and Paulina Oliva. 2015. "The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City." *Journal of Public Economics* 122:68–79.

Hu, Shishan, Scott Fruin, Kathleen Kozawa, Steve Mara, Suzanne E. Paulson, and Arthur M. Winer. 2009. "A wide area of air of air pollutant impact downwind of a freeway during pre-sunrise hours." *Atmospheric Environment* 43 (16): 2541–49.

Karlsson, Martin, and Nicolas R. Ziebarth. 2016. "Population Health Effects and Health-Related Costs of Heat Events: Comprehensive Evidence from Germany." Working Paper. http://www.human.cornell.edu/pam/people/upload/Sunshine.pdf.

Karner, Alex A., Douglas S. Eisinger, and Deb A. Niemeier. 2010. "Near-Roadway Air Quality: Synthesizing the Findings from Real-World Data." *Environmental Science & Technology* 44 (14): 5334–5344.

Kelejian, Harry H., and Ingmar R. Prucha. 2007. "HAC estimation in a spatial framework." *Journal of Econometrics* 140:131–154.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders. 2016. "Caution, drivers! Children present: Traffic, pollution, and infant health." *Review of Economics and Statistics* 98 (2): 350–366.

Kuminoff, Nicolai V., and Jaren C. Pope. 2014. "Do "Capitalization Effects" for Public Goods Reveal the Public's Willingness to Pay?" *International Economic Review* 55 (4): 1227–50.

Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins. 2013. "The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets." *Journal of Economic Literature* 51 (4): 1007–1062.

Linden, Leigh, and Jonah E. Rockoff. 2008. "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws." *American Economic Review* 98 (3): 1103–1127.

Lleras-Muney, Adriana. 2010. "The Needs of the Army: Using Compulsory Relocation of the Military to Estmate the Effect of Air Pollutants on Children's Health." *Journal of Human Resources* 45 (3): 549–590.

Luechinger, Simon. 2014. "Air pollution and infant mortality: A natural experiment from power plant desulfurization." *Journal of Health Economics* 37:219–231.

Mian, Atif, and Amir Sufi. 2009. "The Consequences of Mortgage Credit Expansion: Evidence from the US Mortgage Default Crisis." *Quarterly Journal of Economics* 124 (4): 1449–1496.

Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72 (1): 159–217.

Moretti, Enrico, and Matthew Neidell. 2011. "Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles." *Journal of Human Resources* 46 (1): 154–75.

Muller, Nicholas, and Robert Mendelsohn. 2009. "Efficient Pollution Regulation: Getting the Prices Right." *American Economic Review* 99 (5): 1714–1739.

Neidell, Matthew. 2009. "Information, Avoidance Behavior, and Health: The Effect of Ozone on Asthma Hospitalizations." *Journal of Human Resources* 44 (2): 450–478.

Neidell, Matthew J. 2004. "Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma." *Journal of Health Economics* 23 (6): 1209–1236.

Palmquist, Raymond B. 2005. "Property Value Models." Chap. 16 in *Handbook of Environmental Economics,* edited by Karl-Göran Mäler and Jeffery R. Vincent, 2:763–819. Amsterdam: North Holland.

Perry, Steven G., Alan J. Cimorelli, Robert J. Paine, Roger W. Brode, Jeffrey C. Weil, Akula Venkatram, Robert B. Wilson, Russell F. Lee, and Warren D. Peters. 2005. "AERMOD: A dispersion model for industrial source applications. Part II: Model performance against 17 field study databases." *Journal of Applied Meteorology* 44 (5): 694–708.

Pope, Jaren C. 2008. "Fear of crime and housing prices: Household reactions to sex offender registries." *Journal of Urban Economics* 64:601–614.

Qin, Yu, and Hongjia Zhu. 2015. "Run Away? Air Pollution and Emigration Interests in China." Working Paper.

Ridker, Ronald G., and John A. Henning. 1967. "The Determinants of Residential Property Values with Special Reference to Air Pollution." *Review of Economics and Statistics* 49 (2): 256–257.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82 (1): 34–55.

Schlenker, Wolfram, and W. Reed Walker. 2016. "Airports, Air Pollution, and Contemporaneous Health." *Review of Economic Studies* 83 (2): 768–809.

Sillman, Sanford. 1999. "The relation between ozone, NO_x, and hydrocarbons in urban and polluted rural environments." *Atmospheric Environment* 33:1821–1845.

Smith, V. Kerry, and Ju-Chin Huang. 1995. "Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models." *Journal of Political Economy* 103 (1): 209–227.

South Coast Air Quality Management District. 2000. *Review of RECLAIM Findings.* Internal Report, October 20.

———. 2015a. Public Records Requests #80085 and #80086.

———. 2015b. Correspondence in reply to Public Records Request #80089.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business and Economic Statistics* 20 (4): 518–529.

Stock, James H., and Motohiro Yogo. 2002. "Testing for Weak Instruments in Linear IV Regression." NBER Technical Working Paper #284.

Sullivan, Daniel M. 2017. "Residential Sorting and the Incidence of Air Quality Improvements." Working Paper.

U.S. Environmental Protection Agency. 2002. *An Evaluation of the South Coast Air Quality Management District's Regional Clean Air Incentives Market - Lessons in Environmental Markets and Innovation.* Washington, DC. https://www3.epa.gov/region9/air/reclaim/reclaim-report.pdf.

U.S. National Archives and Records Administration. 2004. *Code of Federal Regulations.* Title 40. Protection of Environment.

Weare, Christopher. 2003. *The California Electricity Crisis: Causes and Policy Options.* San Fransisco, CA: Public Policy Institute of California.

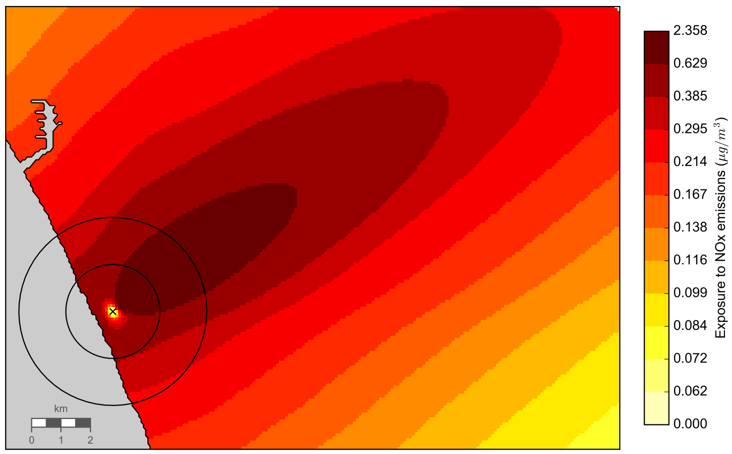

Figure 1: Exposure to NO_x from a Single Firm, 1999

Notes: Colors show average exposure to NO_x emitted by the Scatterwood Generating Stations, Los Angeles, in 1999. Exposure is calculated using AERMOD as described in Section 5. Black "X" marks the location of the firm. Circles mark area within 1 and 2 miles from the firm.

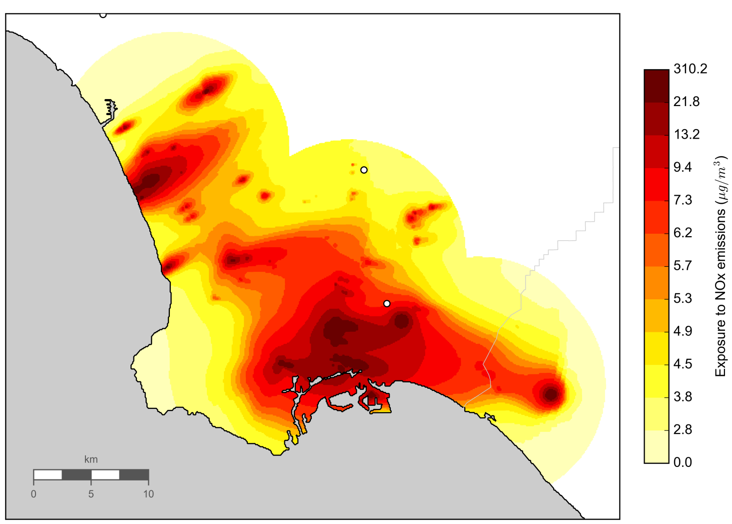

Figure 2: Exposure to Industrial NO_x Emissions, 1999

Notes: Colors show average exposure to NO_x emissions from industrial sources in 1999. White circles mark the location of pollution monitors for NO_x in operation 1997–2005.

(a) Inverse Distance Weighting (b) Kriging

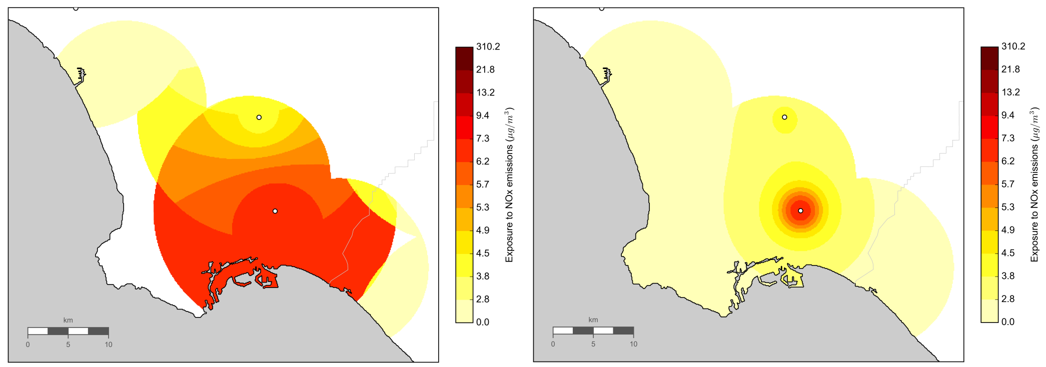

Figure 3: NO_x Exposure as Interpolated from Monitor Locations, 1999

Notes: Figures plot interpolations under the hypothetical that Figure 2 represents true exposure to NO_x emissions but data is only available at monitor locations marked by white dots. These monitors are actual NO_x monitors in operation during sample period (1997–2005) that would be used for interpolation. Color scale for exposure intensity is the same as in Figure 2. Panel (a) plots values interpolated via inverse distance weighting (IDW) with the restriction that monitors are not used (given zero weight) if they are farther than 15 km from the point being interpolated. Panel (b) plots values interpolated via simple Kriging using an exponential variogram.

(a) Northern Monitor (b) Southern Monitor

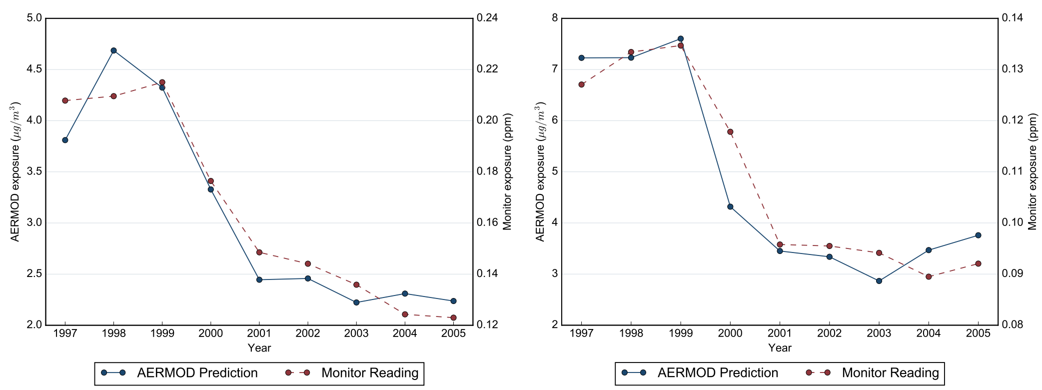

Figure 4: AERMOD and Pollution Monitor Readings Over Time

Notes: Figures plot exposure to NO_x as predicted by AERMOD (solid lines) at the two monitor locations shown in Figure 2, as well as the actual monitor readings for each location (dashed lines). Plotted values are the average from the fourth quarter to minimize measurement issues due to atmospheric chemistry (see Section 3).

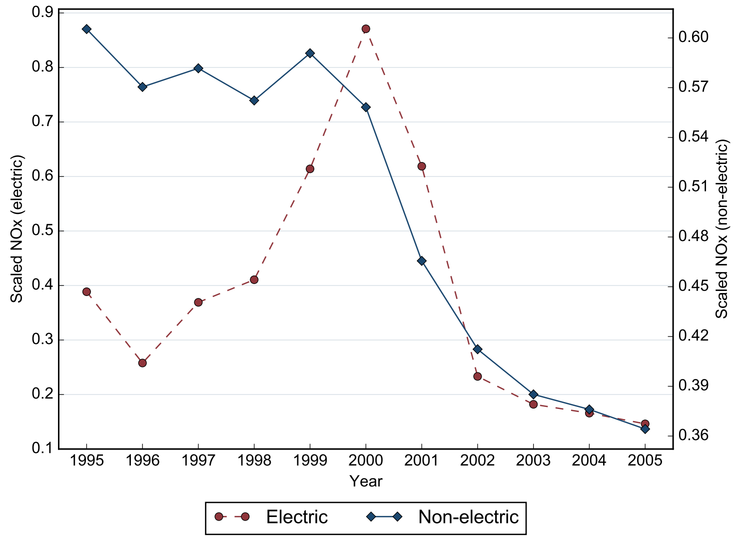

Figure 5: Scaled Firm Emissions of NO_x by Firm Type

Notes: Firm emissions are scaled by firm's own maximum emissions. Sample is restricted to firms that operated in at least 8 quarters.

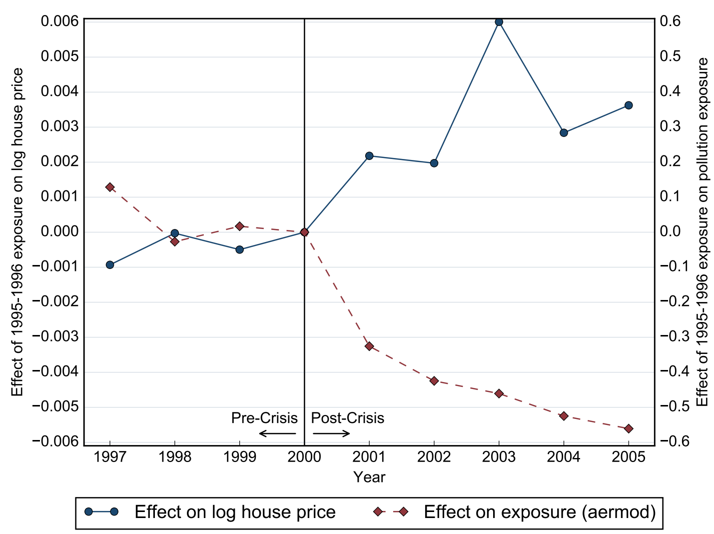

Figure 6: Crisis's Effect on Pollution Exposure and House Prices

Notes: Plotted points are coefficients from a regression of the specified outcome on aermod_pre interacted with year dummies (year 2000 omitted). Sample and other controls as in Table 1, column 1. Average value of aermod_pre is 5.331.

Table 1: Effect on House Prices of AERMOD-measured Pollution Exposure

	(1) ln Price	(2) ln Price	(3) Aermod	(4) ln Price	(5) ln Price	(6) ln Price
Aermod				-0.0074***	-0.0073***	-0.0073***
				(0.0024)	(0.0023)	(0.0024)
Aermod_pre×post	0.0033***	0.0033***	-0.4420***			
	(0.0008)	(0.0005)	(0.0764)			
Aermod_pre		-0.0027**				
		(0.0012)				
Fixed Effects	BG	House	House	House	House	House
Method	OLS	OLS	OLS	2SLS	2SLS	LIML
IV set				Post	Annual	Annual
κ				1	1	1.0003
1st Stage F-stat				6506	951	951
R^2	0.948	0.865	0.910			
N	41,783	118,565	41,783	41,783	41,783	41,783

Notes: Controls include listed fixed effects, year-quarter effects and quadratic time trends by local geography and year 2000 SES variables (see Section 4.3). "Post" IV is aermod_pre × post, "Annual" IV is aermod_pre interacted with year dummies. First-stage F-stat assumes homoskedasticity. Column 2 also includes controls for lot size, bedrooms, bathrooms, interior square feet. Sample average of aermod_pre is 5.331. Standard errors, clustered at 100-m grid, in parentheses: *** p < 0.01, ** p < 0.05, * p < 0.1.

Table 2: Geographic Diff-in-diff Estimates of Crisis's Effect on House Prices

	(1)	(2)	(3)	(4)	(5)	(6)
	0–1 vs. 1–2 miles			0–2 vs. 2–4 miles		
	ln Price	Aermod	ln Price	ln Price	Aermod	ln Price
Near×post	0.0040	-0.5125***		-0.0016	0.0225	
	(0.0050)	(0.0578)		(0.0022)	(0.0221)	
Aermod			-0.0077			-0.0730
			(0.0097)			(0.1205)
Method	OLS	OLS	2SLS	OLS	OLS	2SLS
R^2	0.9454	0.9085		0.9417	0.9095	
N	92,901	92,901	92,901	431,634	431,634	431,634

Notes: Unit of observation is house-firm-quarter. Near=1 for houses closer to firm, e.g., 0–x miles as specified. Controls include house-firm effects and other controls as in Table 1, column 1. Standard errors, clustered by 100-m grid, in parentheses: *** p < 0.01, ** p < 0.05, * p < 0.1.

Table 3: Effect on House Prices of Kernel-measured Pollution Exposure

	(1) ln Price	(2) Triangle	(3) ln Price	(4) ln Price	(5) Uniform	(6) ln Price
Triangle_pre×post	-0.0002	-0.3830***				
	(0.0007)	(0.0112)				
Triangle			0.0004			
			(0.0019)			
Uniform_pre×post				0.0001	-0.4065***	
				(0.0003)	(0.0215)	
Uniform						-0.0003
						(0.0008)
Method	OLS	OLS	2SLS	OLS	OLS	2SLS
R^2	0.948	0.932		0.948	0.905	

Notes: N=41,783. Sample averages of triangle_pre and uniform_pre are 2.303 and 1.681, respectively. Controls as in Table 1, column 1. Standard errors, clustered by 100-m grid, in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Comparison of Capitalization Estimates Across Models

	Model/Paper	Crisis's Effect on Avg. Price	MWTP
	Panel A. Wind-based method		
(1)	Aermod	\$7,860***	\$3,306***
	Panel B. Conventional methods		
(2)	Geo diff-in-diff (1 mile)	\$1,787	
(3)	Geo diff-in-diff (2 miles)	−\$715	
(4)	Triangle kernel	−\$206	−\$179
(5)	Uniform kernel	\$45	\$134
	Panel C. Prior research on other pollutants		
(6)	Smith and Huang (1995)		\$260**
(7)	Chay and Greenstone (2005)		\$191**
(8)	Bayer, Keohane, and Timmins (2009)		\$130***

Notes: For estimates from other papers, the authors' preferred, most comparable estimate is used. Source for Row (1): Table 1, cols 1 & 4. (2): Table 2, col 1; (3): Table 2, col 4; (4): Table 3, cols 1 & 3; (5): Table 3, cols 4 & 6; (6): Smith and Huang (1995), abstract, meta-analysis mean MWTP for TSP; (7): Chay and Greenstone (2005), Table 5A, col 4, MWTP for TSP; (8): Bayer, Keohane, and Timmins (2009), Table 6, col 2, MWTP for PM_{10}.

Appendix

A Monitor cross-validation with no spatial correlation

Let $x_{it} = \delta_t + \varepsilon_{it}$, $\mathrm{Var}(\delta_t) = \sigma_\delta^2$, and $\mathrm{Var}(\varepsilon_{it}) = \sigma_\varepsilon^2$ where ε_{it} is mean-zero and i.i.d. Also let $\tilde{x}_{mt} = \sum_{m' \neq m} w_{mm'} x_{m't}$ where the interpolation weights $w_{mm'}$ are constructed such that $\sum_{m'} w_{mm'} = 1$.

The conventional cross-validation correlation is

$$\mathrm{corr}(\tilde{x}_{mt}, x_{mt}) = \frac{1}{\left[\left(1 + \frac{\sigma_\varepsilon^2}{\sigma_\delta^2} \sum_{m'} w_{mm'}^2\right)\left(1 + \frac{\sigma_\varepsilon^2}{\sigma_\delta^2}\right)\right]^{\frac{1}{2}}} > 0$$

Note that $\mathrm{corr}(\tilde{x}_{mt}, x_{mt}) \to 1$ as $\sigma_\varepsilon^2/\sigma_\delta^2 \to 0$. With large within-year variation but little cross-year variation in wind patterns or firm behavior, this ratio of variances could be very low, leading to a large but erroneous cross-validation correlation.

For the correlation conditional on time effects, we have

$$\begin{aligned}
\mathrm{cov}\left(\tilde{x}_{mt}, x_{mt} \,|\, \delta_t\right) &= \mathbb{E}\left[\left(\tilde{x}_{mt} - \mathbb{E}\left[\tilde{x}_{mt} \mid \delta_t\right]\right)\left(x_{mt} - \mathbb{E}\left[x_{mt} \mid \delta_t\right]\right)\Big|\, \delta_t\right] \\
&= \mathbb{E}\left[\left(\sum_{m' \neq m} w_{mm'} \varepsilon_{m't}\right)\varepsilon_{mt}\,\middle|\, \delta_t\right] = \sum_{m' \neq m} w_{mm'} \mathbb{E}[\varepsilon_{m't}\varepsilon_{mt}] = 0
\end{aligned}$$

B Firm Data Construction

B.1 Geocoding

The accurate geocoding of pollution sources is critical when analyzing the effect these sources have on the surrounding population. Administrative records on the latitude and longitude of each smoke stack operated by the firm would be the ideal data. Regulators often collect this data for the explicit purpose of dispersion modeling, and though SCAQMD does collect this data, they are unavailable for public use (SCAQMD 2015b). In lieu of direct geographic data for each smoke stack, I follow the literature and geocode the firms' street addresses, taking care to use the actual operating address of the firm and not corporate or mailing addresses which are often listed in databases. For large firms and firms that match to interpolated street addresses instead of parcel centroids, I double-checked the coordinates using satellite photos from Google Maps to make sure the geographic point that represents the firm is reasonably close to the actual smoke stacks.[44]

B.2 Facility ID

SCAQMD assigns each facility an ID number; however, a facility may have more than one ID number in the data, both over time and cross-sectionally. This is primarily a concern when matching firms to the NEI, as described in Appendix B.3.

A facility's ID can change under a number of circumstances: the facility is sold, changes its name, or some part of its address changes. For the most part, these changes occur for superficial reasons, e.g., a zip code or street suffix is changed. To construct unique facility ID's, I flagged every pair of facilities less than 400 meters apart and visually inspected satellite photos and emissions data for every cluster of neighboring facilities. First, firms were merged if they occupied the same or neighboring parcels and shared breaks in their time series of emissions. For example, Facility A emits 25 tons per quarter from 1994 to 1999Q3 and then is missing from the data, while Facility B, located at the same parcel of land as A, enters the data in 1999Q4 and begins emitting 25 tons per quarter. Facilities were also merged if they had similar names and occupied the same or neighboring parcels of land. These merges were verified by checking whether or not the firms appeared separately in the NEI.

B.3 Stack Data from the NEI

Data for each firm's smoke stacks is taken from the National Emissions Inventory (NEI) from 1999 and 2002. Besides the smoke stack parameters, the NEI also has data on firm's name, address, SIC, and the equipment's SSC, and the estimated

44. This is potentially important because the firm's "store-front" address right on the street is often at the edge of the property, far away from the smoke stacks. Using unchecked street addresses can introduce significant errors (1–2 km) for firms that occupy large parcels of land.

emissions by pollutant for each stack.[45] It also includes the ID number assigned to the facility by state-level regulators. For SCAQMD firms, this "state ID" consists of a county code, an air basin code, an air district code, and the SCAQMD-assigned facility ID. Using this reconstructed ID, I was able to match most facilities in the SCAQMD emissions data to the NEI using either their own facility ID or an ID from a facility I had previously matched to it as described in section B.2. I used the 2002 NEI data whenever possible, falling back to the 1999 database when necessary. For facilities whose ID's did not match either dataset, I tried to match them using firm address and name. Firms that still did not match were almost all small firms that had ceased to exist before the NEI 1999 data was collected. These firms should have little impact on the overall results and were dropped. For matched facilities, I verified that individual stacks were not duplicated.

Many of the stack parameters in the NEI are flagged as imputed values. The imputation process was not well documented, so I re-imputed them using the median stack parameters from all non-imputed stacks in the SIC and SCC group. Finally, when passing the stack parameters to AERMOD, I weighted each stack according to its reported emissions in the NEI.

45. The Source Classification Codes (SCC) for point pollution sources are a hierarchical index used by the EPA that categorize pollution-generating equipment by combustion type, fuel type, and size. It is analogous to the hierarchical SIC and NAICS industry codes.

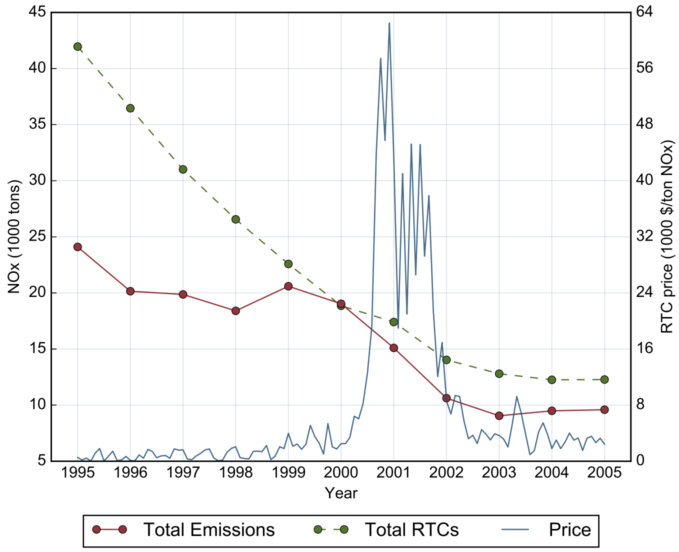

Figure A1: Emissions, Permits, and Permit Price under RECLAIM

Notes: "Total RTCs" is the number of RTCs expiring in the calendar year. "Price" is the average of all arms-length transactions in a month across all RTC vintages.

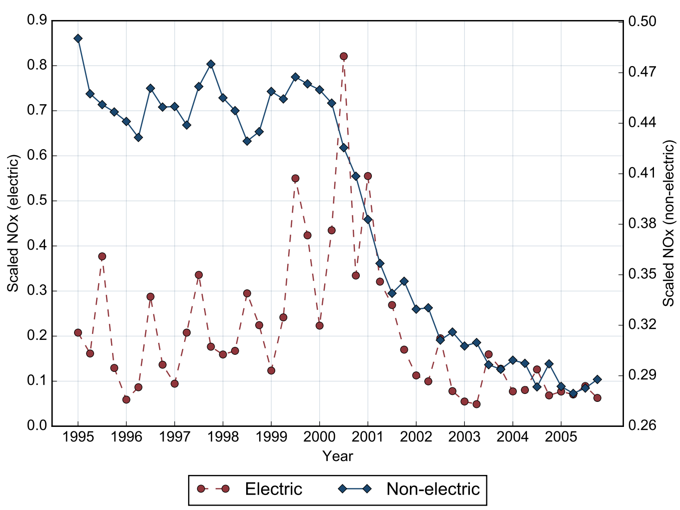

Figure A2: Scaled Firm Emissions of NO_x by Firm Type, Quarterly

Notes: Firm emissions are scaled by firm's own maximum emissions. Sample is restricted to firms that operated in at least 8 quarters.

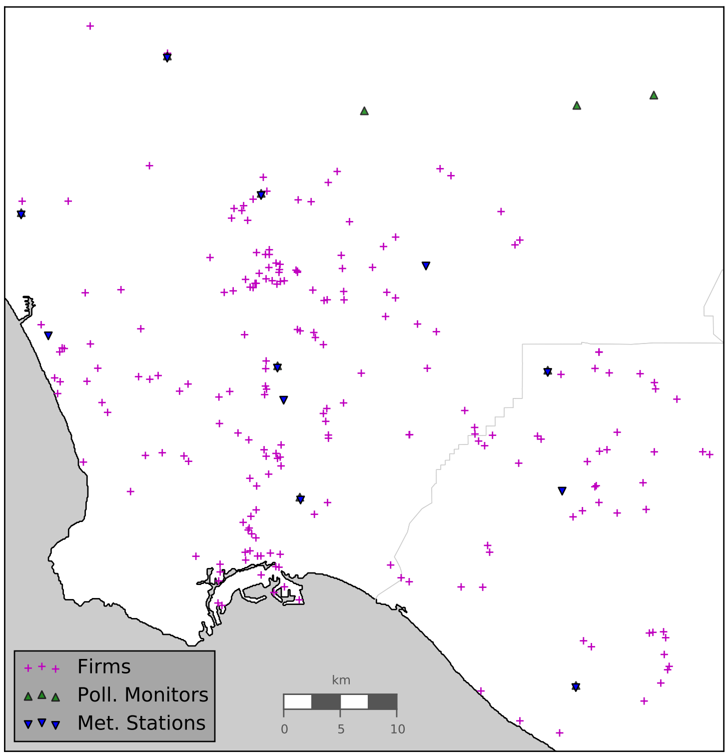

Figure A3: Monitoring Station and Firm Locations

Notes: Firms and meteorology stations are restricted to those that contribute to the main regression sample. Pollution monitors restricted to those with constant NO_x coverage over 1997–2005.

Table A1: House Summary Statistics

	Never Sold	Sold Once		Repeat Sales	
		Pre	Post	Pre	Post
Sale Price		394,839	541,228	420,912	603,347
		(284,955)	(357,514)	(304,854)	(396,748)
Lot Size	6,544	6,617	6,381	6,245	6,010
	(6,662)	(7,173)	(6,793)	(5,567)	(4,926)
Square Feet	1,537	1,611	1,534	1,574	1,492
	(651)	(722)	(690)	(710)	(656)
Year Built	1950	1952	1950	1951	1950
	(15.24)	(15.61)	(15.77)	(16.97)	(16.79)
Bedrooms					
1	0.01	0.01	0.01	0.01	0.02
2	0.23	0.22	0.24	0.25	0.27
3	0.48	0.48	0.49	0.49	0.49
4	0.22	0.23	0.21	0.21	0.19
5+	0.05	0.05	0.05	0.04	0.03
Bathrooms					
1	0.34	0.29	0.33	0.31	0.35
2	0.47	0.47	0.46	0.45	0.45
3	0.13	0.16	0.13	0.15	0.13
4+	0.03	0.04	0.04	0.05	0.04
Sold in Quarter					
1		0.19	0.22	0.20	0.21
2		0.28	0.27	0.29	0.28
3		0.28	0.28	0.28	0.27
4		0.25	0.24	0.24	0.23
Times Sold				2.14	
				(0.38)	
Total Properties	275,218	84,041		19,545	

Notes: Summary statistics from regression sample as described in Section 5. Table lists sample means with standard deviations given in parentheses.

Table A2: Firm Summary Statistics by Industry

	Petroleum Refining	Electric Services	Other Industries	Crude Petroleum and Natrual Gas	Steam and Air-Conditioning Supply	Secondary Smelting and Refining	Other Industrial Inorganic Chemicals	Gypsum Products	All Firms
Mean Emissions (tons)									
1998	524.8	212.6	16.8	31.1	38.8	45.5	39.7	33.7	62.0
2002	380.9	56.1	11.1	8.9	5.7	24.8	37.0	8.8	33.2
Median Emissions (tons)									
1998	332.4	120.0	4.8	5.7	14.5	41.0	34.5	28.7	7.2
2002	255.6	42.2	2.9	1.4	3.7	22.4	43.6	9.3	4.1
Industry Share of Total Emissions (percent)									
1998	42.6	28.8	18.9	3.6	2.1	1.6	1.4	0.9	100.0
2002	56.1	13.8	23.9	1.9	0.6	1.6	1.8	0.4	100.0
Mean Smoke Stack Characteristics									
Height (m)	25.1	37.4	12.3	7.1	19.1	10.6	28.4	19.6	14.9
Diameter (m)	1.3	3.6	0.8	0.4	0.9	0.7	0.9	1.2	1.0
Velocity (m/s)	8.6	20.1	10.8	14.2	12.5	12.5	11.9	9.5	11.7
Temp. (°C)	292.8	231.0	223.0	351.5	191.6	120.9	251.0	271.2	233.7
Mean Dist. to Weather Monitor (km)	7.0	7.5	6.3	6.2	6.8	5.2	6.1	5.9	6.4
No. of Firms	9	15	150	14	6	4	4	3	205

Notes: Sample of firms is those within 20 km of sample area shown in Figure 2.

Table A3: Robustness of House Price Estimates using AERMOD-measured Exposure

	Cluster by		Conley Std. Error w/ Bandwidth				Additional Instruments	
	Block Group	Tract	¼ mile	½ mile	1 mile	2 miles		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Panel A. Reduced Form					
Aermod_pre×post	0.0033^{***}	0.0033^{***}	0.0033^{***}	0.0033^{***}	0.0033^{***}	0.0033^{***}	0.0038^{***}	0.0033^{***}
	(0.0009)	(0.0010)	(0.0009)	(0.0010)	(0.0011)	(0.0011)	(0.0009)	(0.0009)
	[0.000]	[0.002]	[0.000]	[0.002]	[0.002]	[0.003]		
Uniform_pre×post							-0.0004	
							(0.0004)	
"Near Firm"×post								-0.0016
								(0.0058)
			Panel B. 2SLS					
Aermod	-0.0074^{***}	-0.0074^{*}	-0.0074^{***}	-0.0074^{**}	-0.0074^{*}	-0.0074^{*}	-0.0076^{***}	-0.0072^{***}
	(0.0029)	(0.0041)	(0.0028)	(0.0034)	(0.0038)	(0.0041)	(0.0025)	(0.0023)
	[0.010]	[0.072]	[0.009]	[0.030]	[0.053]	[0.074]		

Notes: Each column is a variation of the preferred specifications. Panel A corresponds to the reduced form in Table 1, column 1, and Panel B corresponds to the 2SLS second stage in Table 1, column 4. Standard errors (in parentheses) and p-values (in brackets) are calculated as follows. Columns 1 and 2 use standard errors clustered by Census block group and tract, respectively. Columns 3–6 use SHAC or Conely standard errors with a triangle kernel of bandwidth of ¼ mile, ½ mile, 1 mile, and 2 miles, respectively. Columns 7 and 8 use standard errors clustered at 100-meter grid as in Table 1. The median tract in the sample is roughly 0.75 miles across. A spatial HAC with a 2-mile bandwidth allows arbitrary correlation within an area 4 miles across.

Table A4: Price Effects with Geographic Diff-in-diff and Interpolation

	Panel A. 1-mile treatment, 2-mile control						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln Price	NO_x	ln Price	ln Price	Ozone	ln Price	ln Price
Near×post	0.0037	0.3602			-0.0710		
	(0.0059)	(0.3975)			(0.0835)		
NO_x			0.0102	-0.0079			
			(0.0197)	(0.0049)			
Ozone						-0.0517	-0.0400
						(0.1041)	(0.0265)
Method	OLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
IV Set			Post	Annual		Post	Annual
1st Stage F-stat			1.2	2.8		1.1	2.4
	Panel B. 2-mile treatment, 4-mile control						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln Price	NO_x	ln Price	ln Price	Ozone	ln Price	ln Price
Near×post	-0.0034	-0.0566			0.0490		
	(0.0026)	(0.1474)			(0.0333)		
NO_x			0.0605	-0.0096			
			(0.1619)	(0.0066)			
Ozone						-0.0698	-0.0031
						(0.0726)	(0.0073)
Method	OLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
IV Set			Post	Annual		Post	Annual
1st Stage F-stat			0.2	2.1		3.4	22.6
	Panel C. 3-mile treatment, 6-mile control						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	ln Price	NO_x	ln Price	ln Price	Ozone	ln Price	ln Price
Near×post	-0.0014	-0.0601			0.1772***		
	(0.0018)	(0.0945)			(0.0212)		
NO_x			0.0229	0.0001			
			(0.0475)	(0.0092)			
Ozone						-0.0078	0.0048
						(0.0103)	(0.0053)
Method	OLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
IV Set			Post	Annual		Post	Annual
1st Stage F-stat			0.7	1.4		117.5	52.0

Notes: N for each panel is 76,757; 367,872; and 896,398, respectively. Unit of observation is house-firm-quarter. NO_x and ozone exposure interpolated from monitors using inverse distance weighting. Near=1 for houses within specified treatment radius. Sample restricted to houses within specified control radius. IV Set "Post" is Near×post. IV Set "Annual" is Near times year dummies. 1st Stage F-stat assumes spherical errors. Controls include house-firm effects, year-quarter effects, and quadratic time trends by local geography and year 2000 SES variables.