AS THE WIND BLOWS: THE EFFECTS OF LONG-TERM EXPOSURE TO AIR POLLUTION ON MORTALITY*

Michael L. Anderson
University of California, Berkeley, and NBER

22 February 2016

Abstract

There is strong evidence that short-run fluctuations in air pollution negatively impact infant health and contemporaneous adult health, but there is less evidence on the causal link between long-term exposure to air pollution and increased adult mortality. This project estimates the impact of long-term exposure to air pollution on mortality by leveraging quasi-random variation in pollution levels generated by wind patterns near major highways. We combine geocoded data on the residence of every decedent in Los Angeles over three years, high-frequency wind data, and Census Short Form data. Using these data, we estimate the effect of downwind exposure to highway-generated pollutants on the age-specific mortality rate by using bearing to the nearest major highway as an instrument for pollution exposure. We find that doubling the percentage of time spent downwind of a highway increases mortality among individuals 75 and older by 3.6 to 6.8 percent. These estimates are robust and economically significant.

JEL Codes: Q53, I12

Keywords: Externalities, ultrafine particulates

---

* Anderson: mlanderson@berkeley.edu. I thank Reed Walker, Richard Carson, Mark Jacobsen, and Jeremy Magruder for valuable suggestions. Allison Lassiter and Sikandra Christian provided excellent research assistance. Funding for this research was provided under NIA Grant 5P30AG012839 and by the National Institute of Food and Agriculture, U.S. Department of Agriculture, under Project No. CA-B-AEC-7785-H. Any errors in the paper are the author's.

The effect of air pollution on premature mortality is a fundamental parameter for environmental regulation. For example, the United States Environmental Protection Agency (US EPA) estimates that the 1990 Clean Air Act Amendments (CAAA) will generate $12 trillion in gross benefits from 1990 to 2020, with 92 percent of these benefits accruing in the form of avoided mortality (US EPA 2011, Ch. 7, p. 8). In the past decade, researchers have employed quasi-experimental designs with great success to estimate the effects of air pollution on fetal and infant health (Chay and Greenstone 2003; Currie and Neidell 2005; Currie, Neidell, and Schmieder 2009; Jayachandran 2009; Currie and Walker 2011; Knittel, Miller, and Sanders 2015; Arceo-Gomez, Hanna, and Oliva 2015). There is also strong evidence that short-term fluctuations in air pollution negatively impact contemporaneous pediatric and adult health (Ransom and Pope 1995; Pope and Dockery 1999; Friedman et al. 2001; Moretti and Neidell 2011; Schlenker and Walker 2015). In comparison, however, there is a shortage of quasi-experimental evidence linking long-term exposure to air pollution to increased adult mortality. This effect is of great policy interest because the goal of most air quality regulations, such as the CAAA, is to achieve long-term reductions in ambient pollution levels.

Estimating the effects of long-term exposure to air pollution is challenging for two reasons. First, it is difficult to identify quasi-random variation in long-term air pollution levels across geographic areas. Second, even if pollution were randomly assigned, individuals may endogenously migrate in response to pollution (Banzhaf and Walsh 2008). The identifying variation in air pollution thus needs to be cross-sectional in nature (or a very long panel), exogenous, and yet subtle enough not to induce migration.

We exploit quasi-random variation in pollution levels generated by wind patterns near major Los Angeles highways to estimate the effect of long-term exposure to air pollution on mortality rates. The atmospheric sciences literature has established that certain pollutants, and especially ultrafine particles (UFP), are found at elevated levels up to 600 meters downwind of major highways. In contrast, pollution levels rapidly decline within 100 meters on the upwind or parallel wind sides of highways. This pattern suggests the use of location relative to highways as a proxy for pollution exposure.

Our research design compares mortality rates for individuals who live within 600 meters of highways but on different sides, one predominantly upwind and the other predominantly downwind. This comparison should isolate variation in long-term pollution exposure – the median household in our analytic sample has lived at the same address for over two decades – that is uncorrelated with other factors affecting mortality. In particular, after controlling for distance from highway and a fine set of spatial fixed effects, there is little reason to believe that individuals who live downwind of highways differ from individuals who live upwind of highways, unless people move in response to the pollution itself. Such a response seems unlikely because the pollutants in question, UFP, nitrogen oxides, and CO, are measurable with scientific equipment but not readily perceived by the human senses at the concentrations found near highways (and atmospheric research suggests that coarser particles, which are more readily sensed, do not disperse as far). Furthermore, we demonstrate that property values are not lower downwind of highways, which would be the natural consequence of endogenous migration in response to perceived pollution.

We find a statistically and economically significant relationship between downwind exposure in the Los Angeles Basin and mortality rates among the elderly. For individuals over the age of 75 – the most vulnerable group – a one standard deviation increase in share of time spent downwind of a highway increases mortality by 3 to 5 percent. When instrumenting for percentage of time downwind using bearing to the highway, our estimates imply that a one standard deviation increase in time spent downwind of a highway increases mortality by 5 to 6 percent. These effects persist across a range of elderly or near-elderly age groups (e.g., individuals over 65 or over 70) and spatial bandwidths. Our estimates are somewhat larger in magnitude than those from studies that estimate the high-frequency time-series relationship between daily mortality rates and daily particulate levels. They are smaller than existing cross-sectional estimates, but they remain economically significant and imply substantial benefits from regulating UFP and other near-highway pollutants. They represent, to the best of our knowledge, the first quasi-experimental evidence on the effects of long-term exposure to fine or ultrafine particulate pollution on adult mortality.

I. Background

Dozens of studies establish that daily or weekly fluctuations in air pollution have negative impacts on contemporaneous adult health, including mortality (Pope and Dockery 1999). Extrapolating the effects of short-term fluctuations to long-term exposure, however, is problematic for two reasons. First, the effects of exposure may accumulate over time, so that the cumulative effect of long-term exposure is greater than the implied effect of the sum of repeated short-term exposure coefficients. Second, over short time horizons, the "harvesting" effect – the possibility that short-term insults to health "harvest" sick individuals who were about to die anyway – may underlie some of the contemporaneous relationship between pollution shocks and adult mortality. If so, then the effect of cumulative exposure to pollution may be smaller than suggested by short-term estimates. In summary, for adults it is difficult to bound the effects of long-term pollution exposure in either direction using estimates from short-run pollution fluctuations.[1]

A. Particulates and Health

Particulate pollution has been a focus of air quality regulations since the 1970 CAAA. In 1971, the EPA issued CAA standards focusing on total suspended particles (TSPs), or particles of approximately 100 micrometers in diameter or less. In 1987, they revised the standards to focus on PM_{10} (particles 10 micrometers in diameter or less), and in 1997 they issued standards targeting $PM_{2.5}$ (particles 2.5 micrometers in diameter or less). The clear trend in regulation is toward finer particles over time, and the current research focus on the health effects of particulates is on fine particulates ($PM_{2.5}$) and UFP (particles 0.1 micrometers in diameter or less).

The most heavily-cited evidence linking long-term exposure to air pollution and premature adult mortality comes from cross-sectional epidemiological studies. The seminal paper in this series is the "Six City study" (Dockery et al. 1993), which documents a significant relationship between mortality risk and air pollution across six cities. The mortality rate in the most polluted city in that study was 26 percent higher than the mortality rate in the least polluted city, with the strongest association observed for fine particulates ($PM_{2.5}$). This finding was replicated in a follow-up study covering all US metropolitan areas with available pollution data (Pope et al. 2002), and a similar relationship exists for cardiovascular events and $PM_{2.5}$ (Miller et al. 2007). Pope, Ezzati, and Dockery (2009) use two repeated cross sections and demonstrate that long-differenced (20 year) changes in $PM_{2.5}$ correlate significantly with changes in city-level life expectancy. The EPA applies results from this literature when evaluating the CAAA (US EPA 2011), but it is unclear whether the observed relationships reflect a causal effect of air pollution on mortality or whether they reflect the role of unobserved confounding factors that correlate with air pollution levels across cities.

A small number of papers have employed quasi-experimental methods to estimate the effect of long-term pollution exposure on adult mortality. Chay, Greenstone, and Dobkin (2003) use variation in the long-run reduction in TSP pollution induced by the CAAA of 1970. They find that counties with the largest decreases in TSPs (i.e., the most

[1] The issues discussed are less problematic for infants. Because infants are very young, a short-term fluctuation in pollution can represent a large change in total lifetime pollution exposure.

polluted counties prior to 1970) did not experience greater reductions in adult or elderly mortality than counties with smaller decreases in TSPs. However, they urge caution in interpreting these results "due to the imprecision of the estimated effects and evidence of significant problems with the research design" (Chay, Greenstone, and Dobkin 2003, p. 299). Chen et al. (2013) exploit a policy in China that provides coal-fired heat to all cities north of the Huai River. Using a regression discontinuity (RD) design, they estimate that TSPs are 55 percent higher north of the river and that life expectancies are 5.5 years lower. These results imply large effects of air pollution on mortality. The implications for regulation in the United States (US) and other developed countries are unclear, however, because pollution levels are much higher in China.

The other evidence linking particulates and health comes from laboratory or biomarker studies with animals and humans. Elder et al. (2004) and Elder et al. (2007) exposed laboratory rats to UFP levels mimicking urban roadside environments and found negative effects on white blood cell counts and heart rate. Vinzents et al. (2005) and Brauner et al. (2007) document significant relationships between personal exposure to UFP over several hours and oxidative DNA damage in humans. Frampton et al. (2006) exposed human subjects to UFP and found negative effects on blood leukocytes (white blood cells); Brook et al. (2009) exposed human subjects to $PM_{2.5}$ and found adverse effects on blood pressure. Of relevance to our study, both Frampton et al. (2006) and Oberdörster et al. (2009) found that UFP reduced pulmonary diffusing capacity for carbon monoxide (CO), suggesting a negative interaction effect between two of the main pollutants from motor vehicles (UFP and CO).

B. Pollution Dispersion Near Highways

Understanding the dispersal of pollutants from highways is critical for implementing our identification strategy and interpreting our results. Karner, Eisinger, and Niemeier (2010) synthesize results from 41 atmospheric science studies on near-roadway air quality. These studies measure pollutant levels at varying distances from busy highways in the upwind, downwind, and parallel wind directions. Several clear patterns emerge from this meta-analysis that inform our research design.

First, pollutant levels are consistently higher downwind of highways than upwind of highways. This implies that the percent of time spent downwind of highways should affect pollutant exposure. Second, while many pollutants decay to near background levels within 200 meters downwind, several do not. Most significant among these are UFP, which have demonstrated adverse health effects in laboratory studies, nitrogen oxides (NO and NO_2, or

NO_x), and to a lesser degree CO. UFP decay to background levels by 570 to 910 meters downwind, and nitrogen oxides decay to background levels by 550 to 570 meters downwind. Notable pollutants whose plumes do not extend beyond 100 to 200 meters downwind, or whose concentrations do not seem to be strongly affected by wind direction, include coarse and fine particulates (PM_{10} and $PM_{2.5}$) and ozone (a secondary pollutant). In practical terms, by 300 meters the only pollutants with levels that are at least 15 percent higher than background levels are UFP (150 percent higher), NO (70 percent higher), and CO (25 percent higher) (Karner et al. 2010, p. 5337). Dispersion of up to 500 meters is important because the spatial resolution of our data, while high, becomes imprecise for coding at radii of less than 100 meters. Noise is an additional "pollutant" that decays with distance from the highway, but recent research reveals that noise levels do not vary strongly with wind direction and thus are unlikely to affect our research design (Shu, Yang, and Zhu 2014).

An additional study, conducted after the Karner et al. meta-analysis, is particularly relevant to our research design. Quiros et al. (2013) measure UFP concentrations before, during, and after a 36-hour shutdown of the I-405 highway in Los Angeles. This July 2011 event, locally known as "Carmageddon," was scheduled to accommodate a major highway improvement project. During the closure, particle number concentrations – which are determined by UFP counts – were 83 percent lower 50 to 300 meters downwind of I-405 than during comparable non-closure days. There were no substantial trends in particles upwind of the freeway.[2] These results corroborate the effects of downwind exposure on pollution concentrations in the area included in our study.

Elevated outdoor UFP levels may have limited health effects if the particles do not penetrate indoors. Jamriska et al. (1999), Palmgren et al. (2003), and Morawska et al. (2009) study the relationship between outdoor and indoor levels of traffic-generated particle emissions in a variety of contexts. They find that UFPs have high penetration efficiency into buildings unless mitigated with a high efficiency filtration system, which most residential buildings lack.

In summary, the only pollutants that consistently reach levels high enough to generate a meaningful first stage several hundred meters from the highway are UFP, nitrogen

[2] Quiros et al. also compared downwind weekday particle number concentrations in 2011 to concentrations from the same area in 2001 (taken from an earlier study). They found that concentrations fell 60 percent from 2001 to 2011, suggesting that the effects of being downwind from freeways may have declined during this period.

oxides, and CO.[3] These pollutants are either colorless and odorless (UFP and CO) or are found at concentrations too low to be perceptible to the human senses (the odor threshold for NO_2 is 0.12 ppm, which is above the 99.9th percentile of NO_2 measurements at near-highway pollution monitors in our study area; Nagata and Takeuchi 2003, p. 122). It is thus unlikely that individuals will move in response to downwind frequency.

In terms of health impacts, the clearest hazard is UFP, since they are the most elevated relative to background levels and have been shown to have negative impacts in laboratory studies. CO is also dangerous (Currie and Neidell 2005; Currie et al. 2009), and may negatively interact with UFP, but its plume decays much more rapidly. Nitrogen oxides are a criteria pollutant in part because they interact with volatile organic compounds (VOCs) to form ozone. Since the Los Angeles Basin is VOC-limited (South Coast Air Quality Management District 2014), and has been for many years (Milford, Russell, and McRae 1989), additional nitrogen oxides will not increase ground-level ozone concentrations. Nevertheless, there is some evidence that sustained exposure to low levels of nitrogen oxides, like those found in our study area, may have negative health impacts. Complicating inference is the fact that in almost all contexts there is strong colinearity between fine particles and nitrogen oxides (Committee on the Medical Effects of Air Pollutants 2015). A cautious interpretation of our results is that we estimate the reduced form effect of an increase in several near-roadway air pollutants – UFP, NO_2, and CO – on mortality.

A final strand of literature directly related to this research estimates the relationship between roadway proximity and health. Hoek et al. (2002) examine data in the Netherlands and find that the risk of mortality is 41 percent higher for individuals living within 100 meters of major roads or freeways. Gauderman et al. (2007) find that children living within 500 meters of California freeways had depressed lung development relative to children living more than 1,500 meters from freeways. Currie and Walker (2011) exploit a natural experiment arising from the introduction of electronic tolling and find that reductions in traffic congestion near toll plazas reduces the incidence of prematurity and low birth weight among mothers living within 2,000 meters of the plazas. Rosenbloom et al. (2012) find that all-cause mortality among individuals who have previously suffered from heart attacks is 13 to 27 percent higher when living within 1,000 meters of a major roadway.

[3] As noted above, all other pollutants decay to within 15 percent of background levels or less by 300 meters. Since the difference in average downwind frequency between an "upwind" and "downwind" block is only on the order of 15 to 20 percentage points, the actual difference in average pollution exposure between upwind and downwind blocks would only be a maximum of 2 to 3 percent for these other pollutants. These differences in pollution would be too small to generate a detectable effect on mortality in our data, unless the mortality effect were of a clinically implausible magnitude.

II. Data

Our study estimates the effect of downwind exposure from highways in the Los Angeles Basin (the area of Los Angeles County that lies northwest of the Pacific Ocean and southeast of the mountains). The Los Angeles Basin is an ideal study area for several reasons. First, it contains a large population (approximately 5 million in 2000). Second, there are many major highways. Third, there are consistent, predictable wind patterns across the Basin. This is helpful in assigning wind directions to Census Blocks – our unit of analysis – because weather stations are much sparser than Census Blocks. Finally, we have detailed data on Los Angeles real estate transactions. This enables property value based falsification tests.

The data underlying our estimates come from four distinct sources. Our primary outcome is the Census Block age-specific mortality rate. To compute this rate, we combine two data sets. The first is the California Death Address File. These data contain information on every death in California from 1999 to 2001, including the residential address of each decedent. Key variables include age and cause of death. The second data set is the GeoLytics CensusCD 2000 Short Form. These data contain 2000 Census Short Form data, aggregated and geocoded at the Census Block level. Key variables include population by age group, gender, and race. Note that no data beyond these basic demographic variables exist at the Census Block level because the Census Long Form is only distributed to one in six households and is not available below the Block Group level, which is much too coarse a geographic unit for our analysis (United States Census Bureau 2016).

In Los Angeles, each Census Block generally corresponds to a city block and averages 150 to 200 meters on each side. We geocoded the Death Address File addresses for the entire Los Angeles area and assigned each address to a Census Block. We then computed three-year Census Block mortality rates for various census age groups, including ≥65, ≥70, and ≥75 years of age.[4] For each Census Block, we calculated the distance and bearing to the nearest major highway, as defined by ESRI ArcGIS.

Our independent variable of interest is downwind frequency, or the fraction of time spent downwind of a major highway. We define a Census Block as downwind if the wind direction is within 45 degrees of a perpendicular ray running from the highway to the Census

[4] We define the three year mortality rate for the ≥75 years of age group in Census Block *i* as follows. The numerator is the number of deaths in Census Block *i* from 1999 to 2001 among people who would be 75 years or older in 2000. The denominator is the imputed number of people living in Census Block *i* in 1999 who would be 75 years or older in 2000. The imputed number of people living in Census Block *i* in 1999 is the actual number of people living in Census Block *i* on 1 April 2000, plus the number of people who died between 1 January 1999 and 31 March 2000.

Block.[5] If the wind blows in any other direction or if there is no wind, then the Census Block is not downwind. Note that if the wind blows approximately parallel to the highway (i.e., within 45 degrees of the highway's direction), then neither side of the highway is downwind. In Section V.A, we experiment with an alternative definition of downwind that weights exposure by the cosine of the difference in angles between the wind direction and a perpendicular ray from the highway to the Census Block and find similar results. We omit a very small number of populated Census Blocks that are near the intersections of two highways, as the downwind direction for these blocks is ambiguous.

We collected one year of wind data for 20 available Los Angeles Basin weather stations from MesoWest. We matched each Census Block to its nearest weather station and assigned wind directions and wind speeds using this match. The average distance to the nearest weather station in our analytic sample is 4.9 kilometers, with a standard deviation of 2.1 kilometers and a maximum distance of 11.1 kilometers. We verify the accuracy of these data by predicting the measured downwind frequency at Census Blocks within 500 meters of a weather station using data from the next nearest weather station. The correlation coefficient between predicted downwind frequency and actual downwind frequency is 0.87 (N = 64). However, this figure understates the accuracy of our predictions because the average distance to the next nearest weather station is higher than the average distance to the nearest weather station. If we limit the sample to Census Blocks where the next nearest weather station is less than 7.4 kilometers away, the average distance to the next nearest weather station (5.0 km) becomes similar to the average distance to the nearest weather station in our analytic sample (4.9 km). In this restricted sample, the correlation between predicted downwind frequency and actual downwind frequency is 0.96 (N = 32).

The last data set is our data on property sales. We use these data to conduct falsification tests using property values and to make inferences about the frequency at which households in our sample move. These data come from DataQuick and represent the universe of real estate transactions involving single-family homes in Los Angeles County between 1990 and 1998. The data include address, date of transaction, transaction price, and square footage.

Table 1 presents summary statistics for key variables. There are 27,908 Census Blocks in our overall sample (the Los Angeles Basin), but only 9,314 lie in our analytic sample (i.e., between 50 and 600 meters from a major highway). In both samples, the three-

[5] "Wind direction" in this case refers to the direction toward which the wind blows. However, in meteorological data, "wind direction" refers to the direction from which the wind blows.

year mortality rate among individuals 75 and older is approximately 0.15, with two-thirds of that due to cardio-respiratory causes and under 20 percent due to cancer. The average block is downwind of the closest highway approximately 15 percent of the time, and the winds do not blow at all approximately 42 percent of the time. The average block contains approximately 165 individuals, 8 of whom are over the age of 75. Approximately half of all households in both samples own their own homes. The share of black individuals is lower in the analytic sample (11.8 percent) than in the full sample (14.3 percent).

Figure 1 overlays our analytic sample on a map of the Los Angeles Basin. The Census Blocks in our sample are tightly clustered around highways. In a few cases – for example, just below the exact center of the map – the distribution of Census Blocks appears asymmetric, with a much higher density of blocks on one side of the highway. This occurs when one side of the highway is primarily residential, while the other side is primarily industrial or commercial. To ensure that this type of imbalance does not bias our research design, we employ a spatial fixed effects strategy, discussed in Section III, that limits comparisons to areas in which we have residential Census Blocks on both sides of the highway.[6]

A critical question for our research design is how long the average individual in our sample has lived near the highway. If mobility is high in our sample, then the average length of exposure to elevated pollution levels will be short. The Census Short Form does not have a question on how long a household has lived at the current location, but the Census Long Form, which is available at the Census Tract level, does. The median individual over 75 living in one of our analytic sample's Census Tracts has lived at the current location for 25 years, and 78 percent of them have lived at the current location for over 10 years.[7] Thus, the vast majority of "downwind" individuals in our sample have been exposed to elevated pollution levels for over a decade, and many for over two decades.

III. Empirical Strategy

[6] Areas that lack residential Census Blocks on one side of the highway have no variation in downwind exposure within a small spatial radius. Thus, these areas do not contribute to our estimates when employing our spatial fixed effects design. Areas with a few residential Census Blocks on one side of the highway do contribute to our estimates, but the implicit weight they receive is very low because they have little variation in downwind exposure within a small spatial radius.

[7] To calculate these figures, we match each Census Block from the analytic sample to its Census Tract and calculate the statistics across matched Census Tracts, weighting each Census Tract by the number of matched Census Blocks. If we expand our focus to all individuals over age 65, the median individual has lived at the current location for 25 years, and 73 percent of them have lived at the current location for at least 10 years.

Our empirical strategy compares Census Blocks that are close together but differ in downwind exposure from highways. Because downwind exposure changes discontinuously at the highway, and because we limit comparisons to households that are spatially proximate, our strategy shares features with a boundary discontinuity design. In a typical boundary discontinuity design, identification relies on the assumption that housing supply and demand are smooth across the boundary (in this case, the highway). That assumption may not hold for a single highway segment, because highways often form dividing lines between neighborhoods. However, in our case there are at least eight highways and over two dozen highway segments (where a segment refers to a multi-mile stretch of highway that does not intersect any other highways). Thus, our identification relies on the assumption that, if there are discontinuous changes in housing supply or demand at highways, these changes are not consistently related to the prevailing wind directions. In Section V.B, we test for failures in this assumption by examining the relationship between downwind exposure and household characteristics or property values.

To implement our strategy, we trim the sample along the dimension that is orthogonal to the highway. We then generate spatial fixed effects along the dimension that is parallel to the highway, which we refer to as "highway segment fixed effects." We estimate two sets of regressions using these data. First, we estimate ordinary least squares (OLS) regressions of Census Block mortality rates on percentage of time spent downwind of a highway, controlling for distance to the highway and highway segment fixed effects. Later, we estimate two stage least squares (2SLS) regressions in which time spent downwind of a highway is the endogenous regressor and bearing to the highway is the instrument.

Our analytic sample consists of all Census Blocks located between 50 and 600 meters from major highways in the Los Angeles Basin. We set a minimum distance from the highway because our geocoding of residential addresses to Census Blocks and Census Blocks to highways is only accurate to within 50 to 100 meters. This inaccuracy occurs for several reasons. First, the GeoLytics Census Block boundaries are inexact. In theory, they should precisely overlay the road network, which is the primary delineator of Census Blocks in the Los Angeles Basin, but in practice we observe some slippage. Second, Los Angeles highways are wide – often 50 to 75 meters – so their network representation in the ArcGIS shape file is not exact. Third, the mapping of addresses to coordinates is only approximate in many cases. The ArcGIS shape file assigns each road segment an address range, and addresses within that range are linearly interpolated. For example, in a road segment assigned an address range of 101 through 109, the geocoder assumes that the address of 105 lies at the midpoint of that road segment. All of these issues combine to generate measurement error in the assignment of addresses to Census Blocks. Further from the highway, this generates noise in the dependent variable (Census Block mortality rates) but not the independent variable (downwind exposure to the highway), since a Census Block that is far from the highway should have the same downwind exposure as its neighbor. Adjacent to the highway, however, the measurement error affects the independent variable as well, causing attenuation bias. We thus set a minimum distance of 50 meters to the highway in selecting our analytic sample. This minimum distance threshold is analogous to a "donut RD" in the regression discontinuity framework (Barreca et al. 2011).

We take our maximum distance from the highway of 600 meters from the atmospheric sciences literature summarized in Karner et al. (2010). This literature finds elevated UFP levels out to 570 meters when normalizing concentrations against those found upwind of highways.[8] The 600 meter figure lies near the middle of the range of spatial bandwidths used in existing studies of roadway proximity and health (see Section I.B). We test the sensitivity of our results to different maximum (and minimum) distances in Section V.

We generate our highway segment fixed effects after trimming our sample on distance from the highway. Our highway segment fixed effects are similar to the spatial fixed effects (SFE) that have appeared in other spatial analyses (Conley and Udry 2008; Goldstein and Udry 2008; Magruder 2012). The SFE estimator is analogous to a standard fixed effects estimator in that it demeans each observation i relative to other nearby observations. It then estimates the regression $y_i - \bar{y}_i = \beta(x_i - \bar{x}_i)$, where $\bar{y}_i$ and $\bar{x}_i$ represent the mean values for observations within a radius r of observation i. Unlike a standard fixed effects estimator, however, SFE cannot be represented as a set of dummy variables, because the relevant comparison group changes continuously as one moves through space.

Our highway segment fixed effects modify the SFE estimator to demean observation i relative to observations lying within a radius r along the dimension parallel to the highway.[9]

---

[8] Karner et al. report that elevated UFP concentrations persist out to 910 meters downwind of highways when normalized against background concentrations far from highways. However, given our research design, normalizing against upwind levels is more relevant than normalizing against concentrations in areas with no highways.

[9] For Census Block i, we calculate the distance to any point j along the dimension parallel to the highway nearest Block i as $d_{ij} = \sqrt{\left(lat_j - lat_i\right)^2 + \left(lon_j - lon_i\right)^2} \cdot \left|\sin\left[\tan^{-1}\left(\frac{lon_j - lon_i}{lat_j - lat_i}\right) - \theta_i\right]\right|$, where *lat* and *lon* represent latitude and longitude (normalized to meters), and θ_i is the bearing of a perpendicular ray from the highway nearest i to Census Block i (converted to radians). Block j is included in the neighborhood mean for Block i if and only if d_{ij} is less than r (and both blocks lie within 600 meters of the same highway).

We implement highway segment fixed effects rather than standard spatial fixed effects because they allow us to independently control the spatial bandwidth along two orthogonal dimensions: distance from the highway and distance along the highway. For example, suppose that $r = 800$ meters and that observation i lies 400 meters south of an east-west highway. Observation i is compared to all other observations on that highway that are within 800 meters in the east-west direction. This includes observations over 400 meters north of the highway, even though these observations are more than 800 meters away from observation i in two-dimensional space. With standard spatial fixed effects, it is impossible to assess the sensitivity of our results to decreasing the radius of the SFE without also decreasing the bandwidth around the highway. Independent manipulation of both bandwidths is important because our highway segment fixed effects are meant to control omitted variables bias, while our bandwidth around the highway determines the composition of our sample (and potentially the average treatment effect).

After trimming our sample to Census Blocks located between 50 and 600 meters and transforming our data with highway segment fixed effects, we estimate OLS regressions of the form

$$\tilde{y}_i = \beta\widetilde{w}_i + \tilde{x}_i\delta + \tilde{\varepsilon}_i \qquad (1)$$

where y_i represents the three-year mortality rate in Census Block i among individuals 75 and older, w_i represents the fraction of time that Census Block i is downwind of a highway, and x_i represents other covariates. We define the transformation $\tilde{u}_i = u_i - \bar{u}_i$, where $\bar{u}_i$ is the mean of observations lying within r meters of observation i along a line parallel to the highway. We set a default highway segment fixed effect bandwidth of $r = 800$ meters but test our results' robustness to different bandwidths. Covariates in the vector x_i include distance to the highway and weather station fixed effects.

We augment our OLS estimates with 2SLS estimates that employ bearing to highway as an instrument for downwind exposure. 2SLS estimates have two potential advantages over OLS estimates. First, the 2SLS estimates should be less sensitive to the exclusion of spatial fixed effects because bearing to the highway is evenly distributed throughout the Los Angeles Basin. In contrast, even if housing is evenly distributed across both sides of all highways, downwind exposure could be higher in certain areas of Los Angeles simply because winds might blow more consistently in those areas. Second, the 2SLS estimates should reduce the measurement error in downwind frequency that arises because most Census Blocks do not contain weather stations. Because the measurement error will likely attenuate the OLS estimates, we expect – and find – that the 2SLS estimates exceed the OLS estimates in magnitude.

We parameterize our instrument, bearing to the nearest major highway, as a set of seven dummy variables. Each dummy variable represents a 45-degree range (e.g., 22.5 degrees to 67.5 degrees, 67.5 degrees to 112.5 degrees, etc.), and the excluded category is north (337.5 degrees to 22.5 degrees). Our first stage regression is thus

$$\tilde{w}_i = \tilde{z}_i\alpha + \tilde{x}_i\gamma + \tilde{v}_i \quad (2)$$

where z_i represents the set of 45 degree range dummy variables, and w_i, x_i, and the transformation $\tilde{u}_i$ are as defined above. The second stage estimates the equation:

$$\tilde{y}_i = \beta\hat{w}_i + \tilde{x}_i\delta + \tilde{\varepsilon}_i \quad (3)$$

where $\hat{w}_i$ are the fitted values from the first-stage results.

In all regressions (OLS and 2SLS), we compute standard errors that are robust to spatial dependence, following Conley (1999). We employ a uniform kernel and a spatial bandwidth of 3,200 meters (two miles) in computing the standard errors. Their size is insensitive to reasonable variations in this bandwidth or alternative kernel choices.

IV. Results

We begin with a graphical analysis of the relationship between downwind exposure and mortality. Figure 2 presents results from a local linear regression of the three-year mortality rate among individuals 75 and older on the frequency of downwind exposure to a major highway. In this figure, both mortality rates and downwind exposure are residualized with respect to 800 meter highway segment fixed effects; downwind frequency is thus negative for a small number of Census Blocks. Figure 2 reveals that Census Blocks with a high frequency of downwind exposure have higher mortality rates than Census Blocks with a low frequency of downwind exposure. The relationship appears approximately linear, except for modest convexity at low levels and strong concavity at very high levels of downwind exposure, though the number of observations, and thus the precision of the estimates, is low at the extremes.

Figure 3 presents the instrumental variables analog of Figure 2. Figure 3 plots the relationships of two variables with respect to bearing to the highway (the instrument).[10] The first plot – the dashed blue line – is the relationship between downwind frequency and bearing to the highway. This plot is the graphical analog of the first-stage regression. It

[10] As in Figure 2, all variables in Figure 3 are residualized with respect to 800 meter highway segment fixed effects.

reveals that Census Blocks located east and north of highways (i.e., those whose bearing to the highway is west or south) are downwind much more often than those located west or south of highways. The second plot – the solid red line – is the relationship between the three-year mortality rate among individuals 75 and older and bearing to the highway. This plot is the graphical equivalent of the reduced-form regression. It reveals that Census Blocks located east and north of highways have higher mortality rates than those located west or (to a lesser extent) south of highways. The visible correlation between the dashed blue line and the solid red line suggests a relationship between downwind frequency and mortality, consistent with Figure 2.

Our results tables report the coefficient on downwind frequency, which ranges from zero to one. However, the raw coefficient is not directly relevant because a change in downwind frequency from zero to one represents a shift of almost eight standard deviations and is far outside the support of our data. In the text, we thus refer to effects of a one standard deviation (0.13 unit) change in downwind frequency; by coincidence, this is approximately equivalent to doubling downwind frequency from its average level of 0.154 units.

Table 2 presents results from estimating equation (1) via least squares. Column (1) regresses the three-year mortality rate among individuals 75 and older on frequency downwind, plus controls for distance to the highway and weather station fixed effects. It does not transform the data using highway segment fixed effects, instead including flexible controls for latitude and longitude (quintics in latitude and longitude, plus first and second order interactions between latitude and longitude). A one standard deviation (or 0.13 unit) increase in downwind frequency is associated with a 0.5 percentage point (or 3 percent) increase in the all-cause mortality rate. This result is statistically significant ($t = 2.3$). Column (2) transforms the data using highway segment fixed effects and corresponds to equation (1); this is our preferred OLS specification. The effect of a one standard deviation change in downwind frequency increases to 0.8 percentage points (5 percent) and becomes highly significant ($t = 3.6$).

Columns (3) through (8) in Table 2 report results for mortality from specific causes. Columns (3) and (4) report effects on cardio-respiratory related mortality using the same regressions as columns (1) and (2), respectively. Previous epidemiological studies, as well as laboratory studies, suggest that air pollution should have pronounced impacts on cardiovascular health. Columns (3) and (4) reveal that over half the effect on overall mortality is due to deaths from cardio-respiratory diseases, and the specification with highway segment fixed effects achieves statistical significance (t = 3.0). Columns (5) and (6) report effects on lung cancer deaths, while columns (7) and (8) report effects on deaths from other cancers. In all cases, the effects are positive but statistically insignificant.

Table 3 presents results from the first-stage regression of downwind frequency on bearing to the nearest highway. Column (1) estimates the relationship with the default set of controls plus flexible controls for latitude and longitude, and column (2) implements highway segment fixed effects. The coefficients in both columns are similar and confirm the visual relationship in Figure 3; Census Blocks with a highway to the west or south (i.e., located east or north of the highway) are downwind at a higher frequency than Census Blocks with a highway to the east or north. Since we parameterize bearing to the nearest highway as seven indicator variables, the possibility of many weak instruments is a concern. However, the F-statistic on the instruments ranges from 26.6 to 30.6, which is well above the suggested critical values for first-stage F-statistics in Stock, Wright, and Yogo (2002). Furthermore, the partial R^2 for the instruments exceeds 0.55 in both columns, implying that our instruments explain the majority of the variation in downwind frequency.[11]

Table 4 reports 2SLS estimates of the effect of downwind exposure to highways on mortality. Columns (1) and (2) present the effects on overall mortality among 75+ year olds. The first column includes the default controls and flexible functions of latitude and longitude but no highway segment fixed effects, while the second column adds highway segment fixed effects and is our preferred 2SLS specification. The estimated effect on mortality rates of a one standard deviation increase in downwind exposure is 0.8 percentage points (5 percent) without highway segment fixed effects and 0.9 percentage points (6 percent) with highway segment fixed effects. Both estimates are highly significant (t = 2.9 and t = 3.0) and pass overidentification tests (i.e., we fail to reject the hypothesis that all of our instruments estimate the same parameter). The 2SLS estimates are less sensitive than the OLS estimates to the use of highway segment fixed effects, presumably because bearing to the highway is more balanced across space than is downwind frequency.

Columns (3) and (4) report 2SLS estimates of the effects on cardio-respiratory mortality. As with the OLS estimates, the effect on cardio-respiratory mortality accounts for the majority of the overall mortality effect. A one standard deviation increase in downwind frequency raises the cardio-respiratory mortality rate by 0.4 (Column 3) or 0.5 (Column 4) percentage points. Both estimates are marginally significant (t = 1.9 and t = 2.0). Columns

[11] Estimating the effects using LIML, whose median is generally close to the population parameter to be estimated even in cases with many instruments, generates estimates that are nearly identical to the 2SLS estimates.

(5) through (8) report 2SLS estimates of the effects on mortality from lung cancer and other cancers. All point estimates are positive, but most are statistically insignificant. The one exception occurs for lung cancer, which achieves marginal significance in Column (6) (t = 1.9); the point estimate implies that a one standard deviation increase in downwind exposure increases lung cancer mortality by 0.1 percentage points (17 percent).

V. Robustness and Falsification Tests

The estimated effects of downwind exposure on mortality are conditional on choices about the affected population and the appropriate spatial bandwidths. In this section, we explore our estimates' sensitivity to these choices and conduct a series of falsification exercises to test whether the relationship between mortality and downwind exposure could be due to residential sorting.

A. Robustness to Parameter Choices

Our regressions estimate the effect of downwind exposure on mortality rates among a specific population: 75+ year olds living 50 to 600 meters from highways. As we describe in Section III, data limitations dictate the minimum distance from a highway (50 meters), and the results from the atmospheric sciences literature inform the maximum distance from a highway (600 meters). We use a radius of 800 meters for our spatial fixed effects because it corresponds to one-half mile and is close in magnitude to the 600 meter radius that we apply around the highways. We focus on 75+ year olds because previous cross-sectional studies have found that the relationship between air pollution and negative health events increases with age in both proportional and absolute terms (Miller et al. 2007). Finally, for simplicity, we define "downwind" to mean that the wind direction is within 45 degrees of a perpendicular ray running from the highway to the Census Block. Tables 5 and 6 examine how our estimates change with respect to these parameter choices.

Table 5 reports estimates from our preferred OLS and 2SLS specifications for a variety of spatial bandwidths. Each coefficient represents a separate regression. Columns (1) and (2) report effects of downwind exposure on all-cause mortality, and columns (3) and (4) report effects of downwind exposure on cardio-respiratory mortality. The top set of rows reproduces the baseline OLS and 2SLS estimates, taken from Tables 2 and 4, for comparison purposes.

The first set of rows (following the top set) presents results from regressions that change the definition of downwind frequency. The alternative definition of downwind frequency weights exposure by the cosine of the difference in angles between the wind direction and a perpendicular ray from the highway to the Census Block. Formally, the weight is $w = \cos(\theta - 90)$, where θ is the angle between the wind direction and the highway. This implies $w = 1$ when the wind blows perpendicular to the highway, $w = 0.71$ when the wind blows at a 45-degree angle to the highway, and $w = 0$ when the wind blows parallel to the highway. We set a zero lower bound on w so that it does not become negative when a Census Block is upwind. With this alternative definition, Census Blocks receive some downwind exposure even when the wind blows at angles between 0 and 45 degrees to the highway. Using this alternative definition, we find estimates that are slightly smaller in magnitude than our baseline estimates but remain highly significant.

The next two sets of rows present estimates that apply spatial fixed effects with radii of 400 meters (one-quarter mile) and 1,600 meters (one mile). In all columns, the estimates are reasonably close to our baseline estimates, implying that our results are not very sensitive to changes in the radius of our spatial fixed effects. The subsequent two sets of rows present estimates that change the "donut size," or minimum distance from a highway, to 25 meters or 100 meters. Reducing the donut size to 25 meters decreases the OLS (2SLS) effects on all-cause mortality by 15 percent (21 percent). The OLS (2SLS) effects on cardio-respiratory mortality drop by 21 percent (33 percent). Increasing the donut size to 100 meters has the opposite effect, with effect sizes increasing by approximately 20 percent, except in column (3), where they increase by 12 percent. These patterns are consistent with the fact that measurement error in a Census Block's location relative to the highway becomes more severe as the donut size shrinks.

The bottom two sets of rows present estimates that change the maximum distance from a highway to 400 meters or 800 meters. Reducing the maximum distance to 400 meters has the largest impact of any spatial bandwidth modification; the OLS effect on all-cause mortality drops by 23 percent, though it remains statistically significant ($t = 2.4$). The 2SLS effect is less impacted, dropping by 16 percent. The effects on cardio-respiratory mortality are also less impacted, dropping by 6 to 21 percent (though they lose significance due to a decrease in the coefficients and an increase in the standard errors). Increasing the maximum distance to 800 meters has minimal impact on most estimates except the 2SLS effect on cardio-respiratory mortality, which decreases by 28 percent and loses statistical significance.

Table 6 reports estimates from our preferred OLS and 2SLS specifications for different age groups. The sample size as the age groups expand because the likelihood of a Census Block containing a positive number of people in a given age group grows with the size of the age group.[12] Columns (1) through (4) report effects on all-cause mortality and cardio-respiratory mortality for 70+ year olds. The effects are 8 to 35 percent smaller in absolute magnitude than for 75+ year olds, but statistical significance remains unchanged. Columns (5) through (8) report effects on all-cause mortality and cardio-respiratory mortality for 65+ year olds. The effects diminish further but, with the exception of the 2SLS coefficient for cardio-respiratory mortality, remain statistically significant. Columns (9) through (12) report effects for a much younger age group, 50+ year olds. The coefficients are close to zero and reveal no significant effects for this younger group.

Overall, while the estimates do vary with some spatial bandwidths, both the OLS and 2SLS effects on all-cause mortality are always significant. The effects decrease for younger age groups, as expected, but persist when expanding the population to include 70-74 year olds and, in most cases, 65-69 year olds.

B. Falsification Tests

Identification in our study hinges on the assumption that an individual's bearing to the nearest highway is "as good as randomly assigned." There are two ways in which this assumption could fail. One would be if there were discrete changes in housing supply or demand at highways that were unrelated to winds but consistently occurred in the direction of prevailing winds. In practice, this would entail northern and eastern sides of highways being consistently poorer than southern and western sides. The second would be if households moved in response to the wind-driven pollution. This is unlikely since UFP and CO pollution are undetectable to human senses, and any movement in response to illness would attenuate our effects rather than inflate them. In either scenario, however, we would expect demographic characteristics and property values to vary with downwind exposure. Tables 7 and 8 thus estimate the relationships between these characteristics and downwind exposure.

Table 7 presents results from OLS and 2SLS regressions in which the dependent variable is a measure that should be unrelated to downwind exposure if our research design is valid. Columns (1) and (2) estimate OLS and 2SLS regressions in which the dependent variable is the share of households that own their own home. The coefficients are statistically insignificant, and the point estimates imply that downwind Census Blocks have higher rates of home ownership, contrary to what we might expect if residential sorting were occurring.

[12] Restricting the samples to include only Census Blocks that are in our baseline sample for 75+ year olds generates coefficient estimates of similar magnitudes to those reported in Table 6.

In either column, we can reject the hypothesis that a one standard deviation increase in downwind exposure correlates with a greater than 0.5 percentage point (1 percent) decline in homeownership rates. Columns (3) and (4) estimate regressions in which the dependent variable is the share of individuals who are African-American. The coefficients are statistically insignificant, and the point estimates imply that downwind Census Blocks are less likely to contain African-Americans. Columns (5) through (8) estimate regressions in which the dependent variable is the external-cause mortality rate – i.e., deaths from accidents, homicide, or suicide – among 75+ year olds (the fifth and sixth columns) or among all individuals (the seventh and eighth columns). In all cases, the coefficients are statistically insignificant, although in general they are imprecisely estimated relative to the mean because external-cause mortality is a rare outcome. Ideally we would test whether household income and education vary with downwind exposure as well, but these measures are not on the Census Short Form and thus are not available below the Block Group level. As an alternative, we test whether property values vary with downwind exposure.

Table 8 presents results from OLS and 2SLS regressions in which the dependent variables are housing prices or housing characteristics. The unit of observation is a house or condominium sale, and we match each sale to a Census Block from our analytic sample to assign downwind frequency. These regressions represent a strong test of the research design in that any large-scale residential sorting should manifest itself in housing prices. The data for these regressions come from DataQuick (1990 to 2000 sales) or the Los Angeles County Assessor's Office (2006 to 2010 sales). The DataQuick data's date range fits our study period better, but they only include sales of single-family homes and only cover the City of Los Angeles (which does not contain the entire Los Angeles Basin). The Assessor's Office data covers sales of all residential units in the entire Los Angeles Basin, but the date range is somewhat later than our study's data. Given these limitations, we present estimates for each data source separately.

Column (1) of Table 8 reports results from OLS regressions of log price on downwind frequency, as well as our standard controls. Panel A reports estimates on the 1990 to 2000 sales sample (DataQuick), and Panel B reports estimates from the 2006 to 2010 sales sample (Assessor's Office). A one standard deviation (0.13 unit) increase in downwind exposure is associated with a statistically insignificant 0.2 percent *increase* in property values in either data set. Column (2) reports analogous estimates from 2SLS regressions, and the effects are negative but statistically and economically insignificant. For example, the largest coefficient (–0.113 in Column (2) of Panel A) implies that a one standard deviation increase in downwind frequency is associated with a statistically insignificant 1 percent decrease in property values. Columns (3) and (4) estimate the same regressions as Columns (1) and (2) but include log square footage and a cubic in date sold as controls to increase precision. The standard errors fall by 37 to 46 percent, but all coefficients remain statistically and economically insignificant. Columns (5) and (6) estimate OLS and 2SLS regressions in which the dependent variable is square footage to check that square footage is not endogenously determined by downwind frequency. A one standard deviation increase in downwind frequency is associated with a statistically insignificant 13 to 27 square foot increase in house size.

When considering the potential for omitted variables bias it is instructive to directly compare the magnitudes of the coefficients from our property value regressions with those from our mortality rate regressions. The *largest* coefficient in the price regressions in Table 8, –0.140 in Column (4) of Panel A, implies that a one standard deviation increase in downwind frequency is associated with a statistically insignificant 1.8 percent decrease in property values. The comparable mortality rate regression coefficient in Table 4 implies that a one standard deviation increase in downwind frequency causes a 5.8 percent increase in mortality rates. In the Los Angeles Basin we find a cross-sectional elasticity of mortality rates with respect to property values of approximately 0.05. This implies that a 1.8 percent decrease in property values is associated with a 0.09 percent increase in mortality rates, a figure that is almost two orders of magnitude less than our estimated effect of 5.8 percent. It is thus unlikely that our mortality coefficients simply represent neighborhood differences between upwind and downwind Census Blocks, unless these neighborhood differences are somehow not capitalized into property values.

VI. Discussion

Our results imply that living downwind of highways increases mortality rates among the elderly. However, it is difficult to interpret the magnitude of our "reduced-form" estimates without a "first stage" relating downwind frequency to pollution. Estimating this first-stage relationship is challenging because air quality monitors are sparsely located and do not measure UFP, the pollutant we expect our instrument to affect most strongly.[13]

A. "First-Stage" Estimates

[13] The absence of UFP monitoring is not surprising since UFPs are not currently a regulated pollutant.

As a proxy for UFP, and as a relevant pollutant itself, we consider measurements of NO_2. Vehicles are the primary source of NO_2 in Los Angeles, accounting for 85 percent of nitrogen oxide emissions (US EPA 2014). Furthermore, the near-roadway dispersion pattern of NO_2 mimics UFP more closely than other pollutants mimic UFP (Karner et al. 2010). We calculate the downwind frequency of air pollution monitors near highways in the Los Angeles Basin, and estimate a first-stage relationship between downwind frequency and NO_2 concentrations. We compare these first-stage estimates to results from the atmospheric sciences literature and apply them in interpreting our reduced-form results.

Four air pollution monitors in the Los Angeles Basin are close to highways: the West Los Angeles-Veterans Administration (VA) Hospital monitor near Santa Monica, the Los Angeles-Westchester Parkway monitor near Los Angeles International Airport (LAX), the North Long Beach monitor, and the Lynwood monitor. The first two lie southwest of the nearest highways and are thus primarily upwind, and the latter two lie northeast of the nearest highways and are thus primarily downwind. We collected hourly NO_2 measurements from these monitors from 1995 to 2009.[14]

Figure 4 plots average hourly NO_2 concentration against downwind frequency for each monitor. The relationship appears to be approximately linear, although there are only four points of support because downwind frequency is constant within a monitor. Table 9 presents estimates from regressions of hourly NO_2 concentrations on monitor downwind frequency. Column (1) reveals that a one standard deviation (0.13 unit) increase in downwind frequency is associated with a 9.2 part per billion (ppb) increase in NO_2 concentrations, or 33 percent of the mean level. Column (2) adds day-of-sample fixed effects to increase precision and eliminate bias from any imbalance in sample periods across monitors. A one standard deviation increase in downwind frequency is associated with a 7.9 ppb increase in NO_2 concentrations (29 percent of the mean level).

A primary concern in interpreting these estimates is that we cannot control for monitor-specific characteristics. For example, one upwind monitor is near LAX, and one downwind monitor is near the Port of Long Beach. An idealized research design would include highway segment spatial fixed effects to control for any local factors that might affect NO_2 concentrations at a monitor, but we lack sufficient monitors for this design. Stated another way, we would like to cluster at the monitor level, but it is infeasible to do so with only four monitors. As an alternative robustness check, Figure 5 plots coefficients from two sets of

[14] One monitor started collecting data in 2004, and another stopped collecting data in 2008. To ensure that an imbalance in sample periods across monitors does not affect our results, we estimate first-stage regressions with day-of-sample fixed effects.

regressions, each estimated separately by hour of day. The first set of regressions – the solid line – regresses hourly downwind frequency on average downwind frequency (which is fixed within a monitor). The relationship between these two variables is close to zero from midnight until 8 a.m., and then becomes strongly positive from 10 a.m. until 8 p.m. This pattern reveals that most of the downwind exposure at downwind monitors accrues during daylight hours when the winds blow consistently. It also implies that we should expect the relationship between NO_2 levels and average downwind frequency to be stronger during the day than during the night. The second set of regressions – the dashed line – regresses NO_2 concentrations (which vary by hour) on average downwind frequency. This reveals that the relationship between NO_2 concentrations and average downwind frequency is strong during the day, when downwind monitors are actually downwind, and weak during the night, when they are not. These patterns are consistent with the hypothesis that downwind exposure generates the observed differences in NO_2 concentrations across monitors and are inconsistent with the hypothesis that monitor-specific characteristics generate the observed differences in NO_2 concentrations across monitors. If the latter were true, we would expect NO_2 concentrations to be consistently higher throughout the day at downwind monitors, contrary to Figure 5. For completeness, we note that the small NO_2 coefficients during nighttime hours are not the result of an absence of NO_2 during these hours; NO_2 concentrations from midnight to 8 a.m. are nearly identical to the overall average.

Our first-stage estimates are broadly consistent with the results in the atmospheric sciences literature. Karner et al. (2010) report that, across 11 studies, NO_2 concentrations are on average 1.7 to 2.2 times higher than ambient levels on the prevailing downwind side of the highway. Our preferred estimate (Column (2) of Table 9) implies that average NO_2 concentrations at the two downwind monitors are 2.1 times higher than ambient levels.[15] This 2.1-times figure is also consistent with the results from Quiros et al. (2013) following the shutdown of I-405 in Los Angeles. Quiros et al. find that NO levels are approximately twice as high on the downwind side of I-405 when it is open relative to when it is closed.[16]

We combine our first-stage estimates with our reduced-form results to generate a back-of-the-envelope estimate of the elasticity of mortality with respect to pollution. Our first-stage estimates imply that moving from the upwind side to the downwind side increases

[15] On average, the two downwind monitors are downwind 26.8 percent of the time. The intercept for the regression in Column (2) of Table 9 is 15.0, so the implied average NO_2 concentration at the downwind monitors is 15.0 + 61.1*0.268 = 31.4. This figure is 2.1 times higher than the intercept of 15.0 (i.e., a theoretical monitor that is never downwind of the highway).

[16] Quiros et al. take measurements at several distances on the eastern (downwind) side of I-405, but, past 150 meters, the NO concentrations on operational days stabilize at double the concentrations of the closure day.

downwind frequency by 15.5 percentage points and average pollution levels by 43 percent,[17] and our reduced-form estimates imply that a 15.5 percentage point increase in downwind frequency raises mortality rates by 3.6 to 6.8 percent. The "IV" estimate thus suggests an elasticity of mortality rates (among 75+ year olds) with respect to near-roadway pollution in the range of 0.10 to 0.18.

B. Policy Estimates

Two natural questions are how our estimates compare to estimates from the existing literature and what their potential policy implications are. We consider several relevant comparisons from the existing literature: time-series estimates of the effects of short-run exposure; long-differences estimates from Pope et al. (2009); cross-sectional estimates from Dockery et al. (1993) (the "Six City study"); and RD estimates from Chen et al. (2013). Making these comparisons requires some transformation of our results.

Pope and Dockery (1999) summarize an extensive literature estimating the mortality effects of short-run exposure to particulates using daily time-series data. They find consistent estimates of a 0.5 to 1.5 percent increase in mortality rates in response to a 10 $\mu g/m^3$ increase in particulate pollution. Lipfert and Wyzga (1995) translate these effects into elasticities in a meta-analysis and find an average elasticity of daily adult mortality rates to fine particulate pollution of 0.039. Most studies do not separately estimate effects for the elderly and non-elderly, but among those that do, the mortality effects appear concentrated among the elderly. To compare short-term exposure results against ours, we construct a life table using observed mortality rates in Census Tracts within one kilometer of Los Angeles Basin highways. We compare the projected effects of a 10 percent increase in pollution using our estimates with the projected effects of the same increase using estimates from studies examining short-term exposure. Specifically, we compare the effect of a mortality elasticity of 0.14 for 75+ year olds (our average estimate) to the effect of a mortality elasticity of 0.039 for all adults (the average estimate from short-term exposure studies).

Using our life table, and assuming our estimated mortality effects occur only at age 75 and beyond, we compute that a 10 percent change in near-roadway air pollution changes life

[17] On average, the two downwind monitors are downwind 26.8 percent of the time, while the two upwind monitors are downwind 11.3 percent of the time. The intercept for the regression in Column (2) of Table 9 is 15.0, so the implied average NO_2 concentration at the downwind monitors is 15.0 + 61.1*0.268 = 31.4, and the implied average NO_2 concentration at the upwind monitors is 15.0 + 61.1*0.113 = 21.9. The proportional increase in average NO_2 concentration from moving from upwind to downwind is thus 31.4/21.9 = 1.43.

expectancy at birth by 0.05 years.[18] In comparison, a mortality elasticity of 0.039 for all adults implies that a 10 percent change in pollution changes life expectancy at birth by 0.036 years. Thus our estimates are approximately 40 percent larger in magnitude than the implied effects from short-term exposure studies. This pattern suggests that any harvesting issues that may inflate long-run projections from daily exposure studies are more than offset by the negative impacts of cumulative long-run exposure to pollutants.

In contrast to short-term exposure studies, Pope et al. estimate that a 10 percent (2 $\mu g/m^3$) decrease in $PM_{2.5}$ over 20 years in the US increases life expectancy by 0.12 years. Our estimate that a 10 percent change in near-roadway air pollution changes life expectancy at birth by 0.05 years is approximately 60 percent smaller than the estimate from Pope et al. Thus our estimates are larger than those from daily-exposure studies but smaller than those estimated off long-run trends. One concern in comparing effect magnitudes is that restricting mortality effects to 75+ year olds in our simulations may attenuate the predicted impacts if the true effects extend to younger age groups as well. However, if we apply our largest estimate for 65+ year olds in Table 6 to individuals age 65 and above in our life table, we compute that a 10 percent change in near-roadway pollution changes life expectancy at birth by 0.06 years, which is within the range of our predictions above (see Footnote 18). Even when applying the insignificant coefficients for 50+ year olds from Table 6 and adding one standard error to these coefficients, we only project that a 10 percent change in near-roadway pollution changes life expectancy at birth by 0.05 to 0.08 years.

Dockery et al. find an elasticity of mortality rates with respect to fine particle pollution ($PM_{2.5}$) of approximately 0.2 when using cross-sectional data. However, their outcome is the 15-year mortality rate amongst individuals from age 25 to 74 at baseline. To compare our estimates, we compute the effect of an increase in mortality rates amongst 75+ year olds on overall mortality rates for a cohort of 25 to 74 year olds followed over 15 years. The result is an elasticity of mortality rates amongst 25 to 74 year olds with respect to near-roadway pollution of approximately 0.03,[19] which is approximately seven times smaller than the cross-sectional estimate from Dockery et al.

Finally, Chen et al. estimate that a 55 percent increase in TSPs in China reduces life expectancy at birth by 5.5 years. Using our life table, and assuming that mortality effects

[18] Our smallest estimate (OLS with no highway segment FE) generates an effect on life expectancy of 0.035 years, and our largest estimate (2SLS with highway segment FE) generates an effect on life expectancy of 0.060 years.

[19] Our smallest estimate (OLS with no highway segment FE) generates an elasticity of 0.02, and our largest estimate (2SLS with highway segment FE) generates an elasticity of 0.04.

occur only at age 75 and beyond, we compute that a 55 percent increase in near-roadway air pollution reduces life expectancy at birth by 0.2 years.[20] This is approximately 27 times smaller than the estimate from Chen et al. Pollution levels in China, however, are much higher than in the US; average TSP levels in the US were about 60 $\mu g/m^3$ in 1990 (Chay and Greenstone 2003), while Chen et al. report average Chinese TSP levels of 350 to 550 $\mu g/m^3$ (six to nine times higher).

When reconciling our results with the existing literature, several factors are important to consider. First, the relevant particulates differ across papers. We focus on UFP, and to a lesser degree nitrogen oxides and CO. Lipfert and Wyzga, Dockery et al., and Pope et al. focus on coarser $PM_{2.5}$ (and, implicitly, other pollutants that correlate with fine particulates), and Chen et al. focus on still coarser TSPs. Second, Dockery et al. and, to a lesser degree, Pope et al. do not employ quasi-experimental research designs, so their estimates may reflect some degree of selection bias. Third, even when comparing different "long-term exposure" papers, the pollution exposure period may differ. While the median 75-year old in our study has lived at the same location for over 25 years, younger individuals have shorter occupancy durations. The median 45-to-54 year old in our study area, for example, has lived in the same location for only eight years. An eight-year exposure period is roughly comparable to the implicit exposure period in Pope et al. but is shorter than the exposure periods in Dockery et al. and Chen et al. The briefer exposure period for younger individuals could contribute to the null effects we observe on those younger than 65. Finally, the sparsity of pollution monitors and lack of UFP monitoring affects the reliability of our "first-stage" estimates. If we have overestimated the first stage, then we will underestimate the IV coefficient.

To gauge the potential benefits from regulating mobile-source pollution, we consider a policy in which we replace all cars on Los Angeles-area highways with zero-emission vehicles (ZEVs). To calculate the impact of this policy, we construct a counterfactual scenario in which no Census Blocks are ever downwind of a highway. In this scenario, applying our 2SLS estimates to the life table reveals a 0.24-year increase in life expectancy at birth. This increase equates to an additional 372,000 life-years gained across the 1.55 million individuals in our analytic sample. The economic value of this life-expectancy gain totals \$37.2 billion when valuing each life-year at \$100,000 (Neumann, Cohen, and Weinstein 2014). We do not attempt to calculate the exact cost of replacing every car in the Los Angeles Basin with a ZEV over several decades. Nevertheless, we note that there are approximately 2.9 million

[20] Our smallest estimate (OLS with no highway segment FE) generates an effect on life expectancy of 0.14 years, and our largest estimate (2SLS with highway segment FE) generates an effect on life expectancy of 0.28 years.

cars in the Los Angeles Basin, and the value of applying the federal electric vehicle tax credit to all of these vehicles equates to $21.8 billion.[21] In that sense, the local air pollution benefits alone may justify a significant fraction of the current electric vehicle credit, at least in dense urban areas.[22]

VII. Conclusion

We find statistically and economically significant effects of exposure to near-roadway pollution on mortality amongst the elderly. We find no evidence of selection bias or residential sorting – both demographic characteristics and property values appear unrelated to downwind exposure – suggesting that households are generally unaware of the invisible pollution gradient.

When comparing our estimates to estimates from the existing literature, the overall trend that emerges within the United States is that studies leveraging more plausibly exogenous pollution variation appear to find smaller elasticities of mortality with respect to pollution. The largest elasticities arise in cross-sectional studies such as Dockery et al. (1993) and Pope et al. (2002). Pope et al. (2009) employ a long-differencing strategy across cities and find elasticities that are smaller than the cross-sectional studies but larger than the ones reported here. Our estimates are somewhat larger than results from papers examining daily response of mortality rates to fluctuations in air pollution, however, suggesting that in this case the inability to account for cumulative effects outweighs the potential bias from harvesting when making long-run extrapolations from high-frequency studies.

Chen et al. (2013) find a much larger elasticity of life expectancy with respect to pollution – 20 to 30 times larger – than our estimates imply. While higher Chinese pollution levels will directly generate larger elasticities if the dose-response relationship remains constant, this factor alone cannot explain the full difference (since Chinese pollution levels are "only" six to nine times higher than in the US). The larger estimates in Chen et al. thus suggest potential convexity in the relationship between air pollution levels and mortality rates.

---

[21] Los Angeles County contained approximately 5.9 million registered automobiles in 2008 (California Department of Finance 2009), and the Los Angeles Basin contains approximately half the population of Los Angeles County. The federal electric vehicle tax credit is $7,500, so $7,500*2.9 million = $21.8 billion.

[22] Complicating the comparison is the fact that both the costs and benefits evolve dynamically. The benefits figure does not take into account that the "treated" population will include future cohorts not yet born, while the costs figure does not take into account that even low-maintenance electric vehicles will need replacement after two or three decades. The purpose of the comparison is thus not to conduct a precise benefit-cost analysis but to establish that the value of the life-expectancy gains and the electric vehicle credit are of the same order of magnitude.

Our estimates imply that near-highway pollution has economically significant impacts on life expectancy, with a value totaling tens of billions of dollars in the Los Angeles area alone. Given that over 70 percent of the US population lives in urbanized areas (US Census Bureau 2015), the potential nationwide impacts of near-highway pollution are considerably larger. Our results thus suggest significant potential benefits to regulating UFP, which currently are not subject to any standard, and other mobile source pollutants.

References

Arceo-Gomez, Eva O., Rema Hanna, and Paulina Oliva. Forthcoming. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City." *Economic Journal*.

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): 843–63.

Barreca, Alan I., Melanie Guldi, Jason M. Lindo, and Glen R. Waddell. 2011. "Saving Babies? Revisiting the Effect of Very Low Birth Weight Classification." *The Quarterly Journal of Economics* 126(4): 2117–23.

Brauner, Elvira Vaclavik et al. 2007. "Exposure to Ultrafine Particles from Ambient Air and Oxidative Stress-Induced DNA Damage." *Environmental Health Perspectives* 115(8): 1177–82.

Brook, Robert D. et al. 2009. "Insights Into the Mechanisms and Mediators of the Effects of Air Pollution Exposure on Blood Pressure and Vascular Function in Healthy Humans." *Hypertension* 54(3): 659–67.

California Department of Finance. 2009. "2008 California Statistical Abstract."

Chay, Kenneth, Carlos Dobkin, and Michael Greenstone. 2003. "The Clean Air Act of 1970 and Adult Mortality." *Journal of Risk and Uncertainty* 27(3): 279–300.

Chay, Kenneth Y., and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *The Quarterly Journal of Economics* 118(3): 1121–67.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. 2013. "Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences* 110(32): 12936–41.

Committee on the Medical Effects of Air Pollutants. 2015. *Interim Statement on Quantifying the Association of Long-Term Average Concentrations of Nitrogen Dioxide and Mortality*. London: Public Health England. https://www.gov.uk/government/publications/nitrogen-dioxide-interim-view-on-long-term-average-concentrations-and-mortality (February 22, 2016).

Conley, T. G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92(1): 1–45.

Conley, Timothy G, and Christopher R Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100(1): 35–69.

Currie, Janet, and Matthew Neidell. 2005. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *The Quarterly Journal of Economics* 120(3): 1003–30.

Currie, Janet, Matthew Neidell, and Johannes F. Schmieder. 2009. "Air Pollution and Infant Health: Lessons from New Jersey." *Journal of Health Economics* 28(3): 688–703.

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

Dockery, Douglas W. et al. 1993. "An Association between Air Pollution and Mortality in Six U.S. Cities." *New England Journal of Medicine* 329(24): 1753–59.

Elder, Alison et al. 2007. "Effects of On-Road Highway Aerosol Exposures on Autonomic Responses in Aged, Spontaneously Hypertensive Rats." *Inhalation Toxicology* 19(1): 1–12.

Elder, Alison C. P. et al. 2004. "Systemic Effects of Inhaled Ultrafine Particles in Two Compromised, Aged Rat Strains." *Inhalation Toxicology* 16(6-7): 461–71.

Frampton, Mark W. et al. 2006. "Inhalation of Ultrafine Particles Alters Blood Leukocyte Expression of Adhesion Molecules in Humans." *Environmental Health Perspectives* 114(1): 51–58.

Friedman MS et al. 2001. "Impact of Changes in Transportation and Commuting Behaviors during the 1996 Summer Olympic Games in Atlanta on Air Quality and Childhood Asthma." *JAMA* 285(7): 897–905.

Gauderman, W James et al. 2007. "Effect of Exposure to Traffic on Lung Development from 10 to 18 Years of Age: A Cohort Study." *The Lancet* 369(9561): 571–77.

Goldstein, Markus, and Christopher Udry. 2008. "The Profits of Power: Land Rights and Agricultural Investment in Ghana." *Journal of Political Economy* 116(6): 981–1022.

Harley, Robert A. et al. 1993. "Photochemical Modeling of the Southern California Air Quality Study." *Environmental Science & Technology* 27(2): 378–88.

Hoek, Gerard et al. 2002. "Association between Mortality and Indicators of Traffic-Related Air Pollution in the Netherlands: A Cohort Study." *The Lancet* 360(9341): 1203–9.

Jamriska, M., S. Thomas, L. Morawska, and B. A. Clark. 1999. "Relation between Indoor and Outdoor Exposure to Fine Particles near a Busy Arterial Road." *Indoor Air* 9(2): 75–84.

Jayachandran, Seema. 2009. "Air Quality and Early-Life Mortality Evidence from Indonesia's Wildfires." *Journal of Human Resources* 44(4): 916–54.

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): 5334–44.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders. Forthcoming. "Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health." *Review of Economics and Statistics*.

Lipfert, Frederick W., and Ronald E. Wyzga. 1995. "Air Pollution and Mortality: Issues and Uncertainties." *Journal of the Air & Waste Management Association* 45(12): 949–66.

Magruder, Jeremy R. 2012. "High Unemployment Yet Few Small Firms: The Role of Centralized Bargaining in South Africa." *American Economic Journal: Applied Economics* 4(3): 138–66.

Milford, Jana B., Armistead G. Russell, and Gregory J. McRae. 1989. "A New Approach to Photochemical Pollution Control: Implications of Spatial Patterns in Pollutant Responses to Reductions in Nitrogen Oxides and Reactive Organic Gas Emissions." *Environmental Science & Technology* 23(10): 1290–1301.

Miller, Kristin A. et al. 2007. "Long-Term Exposure to Air Pollution and Incidence of Cardiovascular Events in Women." *New England Journal of Medicine* 356(5): 447–58.

Morawska, L. et al. 2009. "Variation in Indoor Particle Number and PM2.5 Concentrations in a Radio Station Surrounded by Busy Roads before and after an Upgrade of the HVAC System." *Building and Environment* 44(1): 76–84.

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.

Nagata, Yoshio, and N. Takeuchi. 2003. "Measurement of Odor Threshold by Triangle Odor Bag Method." *Odor Measurement Review, Office of Odor, Noise and Vibration Environmental Management Bureau, Ministry of the Environment, Government of Japan, Tokyo, Japan*: 118–27.

Neumann, Peter J., Joshua T. Cohen, and Milton C. Weinstein. 2014. "Updating Cost-Effectiveness — The Curious Resilience of the $50,000-per-QALY Threshold." *New England Journal of Medicine* 371(9): 796–97.

Oberdörster, G et al. 2009. *Assessment of Ambient UFP Health Effects: Linking Sources to Exposure and Responses in Extrapulmonary Organs*.

Palmgren, Finn et al. 2003. "Characterisation of Particle Emissions from the Driving Car Fleet and the Contribution to Ambient and Indoor Particle Concentrations." *Physics and Chemistry of the Earth, Parts A/B/C* 28(8): 327–34.

Pope, CA, DW Dockery, and others. 1999. "Epidemiology of Particle Effects." In *Air Pollution and Health*, San Diego: Academic Press, 673–705.

Pope, C. Arden III, Majid Ezzati, and Douglas W. Dockery. 2009. "Fine-Particulate Air Pollution and Life Expectancy in the United States." *New England Journal of Medicine* 360(4): 376–86.

Pope III C, Burnett RT, Thun MJ, and et al. 2002. "Lung Cancer, Cardiopulmonary Mortality, and Long-Term Exposure to Fine Particulate Air Pollution." *JAMA* 287(9): 1132–41.

Quiros, David C. et al. 2013. "Air Quality Impacts of a Scheduled 36-H Closure of a Major Highway." *Atmospheric Environment* 67: 404–14.

Ransom, Michael R, and C. Arden III Pope. 1995. "External Health Costs of a Steel Mill." *Contemporary Economic Policy* 13(2): 86–97.

Rosenbloom, Joshua I. et al. 2012. "Residential Proximity to Major Roadway and 10-Year All-Cause Mortality After Myocardial Infarction." *Circulation* 125(18): 2197–2203.

Schlenker, Wolfram, and W. Reed Walker. Forthcoming. "Airports, Air Pollution, and Contemporaneous Health." *Review of Economics Studies*.

Shu, Shi, Pu Yang, and Yifang Zhu. 2014. "Correlation of Noise Levels and Particulate Matter Concentrations near Two Major Freeways in Los Angeles, California." *Environmental Pollution* 193: 130–37.

South Coast Air Quality Management District. "VOC White Paper." Presented at the Diamond Bar, CA. http://www.aqmd.gov/docs/default-source/Agendas/aqmp/white-paper-working-groups/voc-white-paper-draft-presentation-final-080114.pdf?sfvrsn=2 (February 22, 2016).

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 & Economic Statistics* 20(4): 518–29.

United States Census Bureau. 2015. "2010 Census Urban and Rural Classification and Urban Area Criteria." http://www.census.gov/geo/reference/ua/urban-rural-2010.html (July 2, 2015).

———. 2016. "Census Data API." http://api.census.gov/data/2000/sf3.html (February 5, 2016).

United States Environmental Protection Agency. 2011. "The Benefits and Costs of the Clean Air Act from 1990 to 2020."

———. "State and County Emission Summaries | Air Emission Sources | US EPA." http://www.epa.gov/cgi-bin/broker?_service=data&_debug=0&_program=dataprog.state_1.sas&pol=NOX

&stfips=06 (July 23, 2014).

Vinzents, Peter S. et al. 2005. "Personal Exposure to Ultrafine Particles and Oxidative DNA Damage." *Environmental Health Perspectives* 113(11): 1485–90.

Warneke, Carsten et al. 2013. "Photochemical Aging of Volatile Organic Compounds in the Los Angeles Basin: Weekday-Weekend Effect." *Journal of Geophysical Research: Atmospheres* 118(10): 5018–28.

Figure 1: Analytic Sample Census Blocks

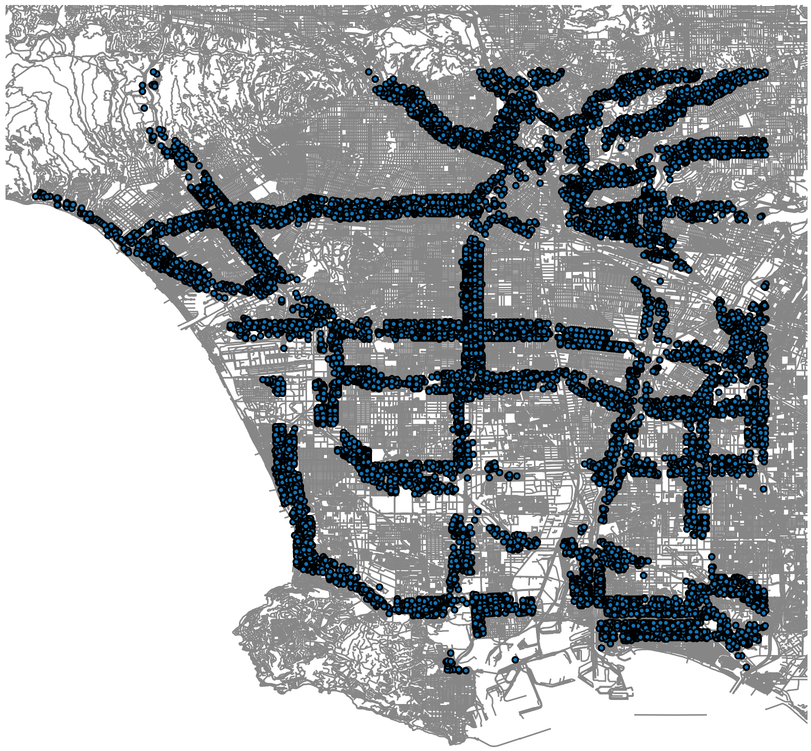

Figure 2

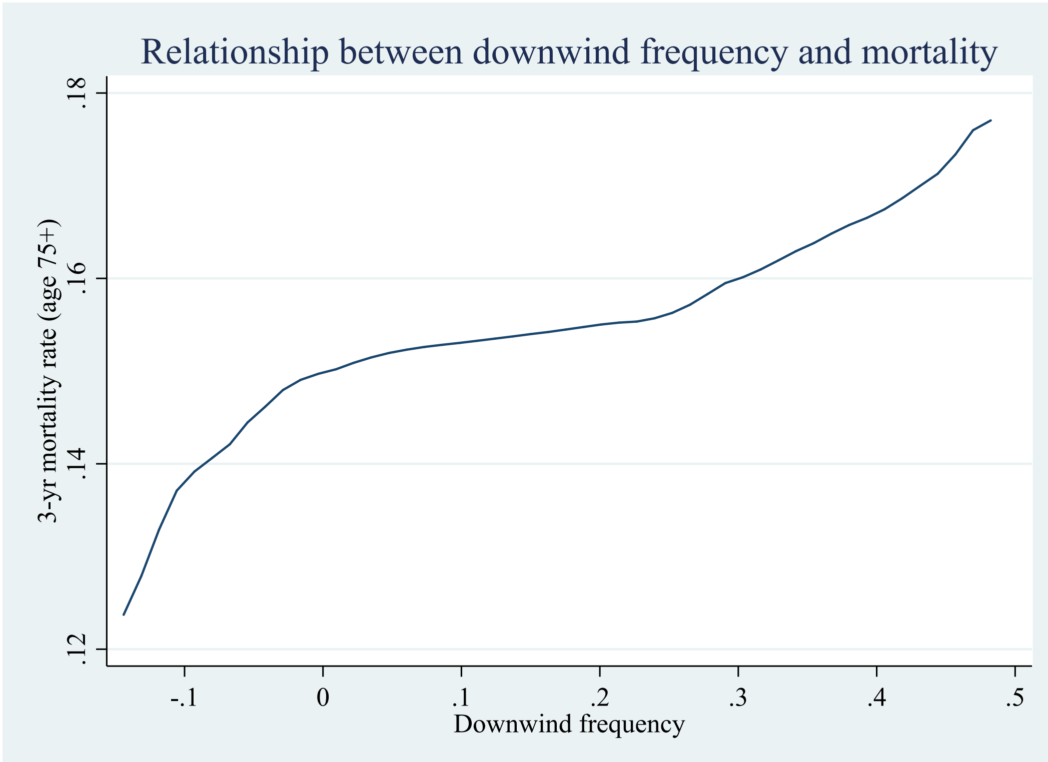

Figure 3

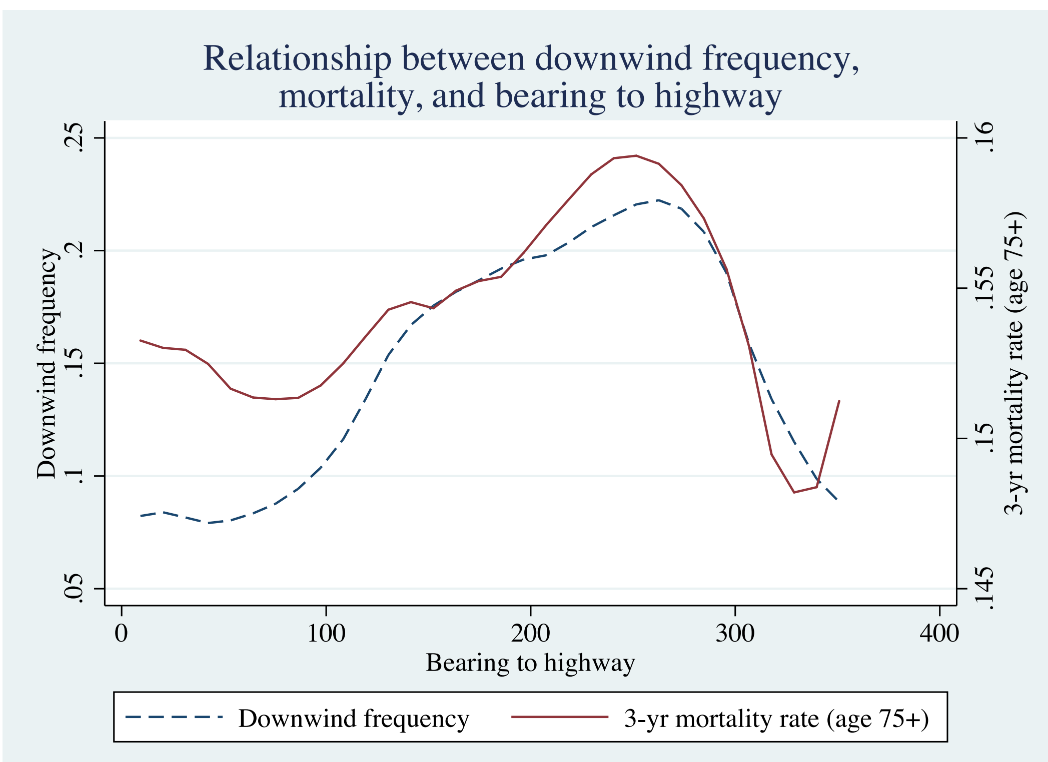

Figure 4

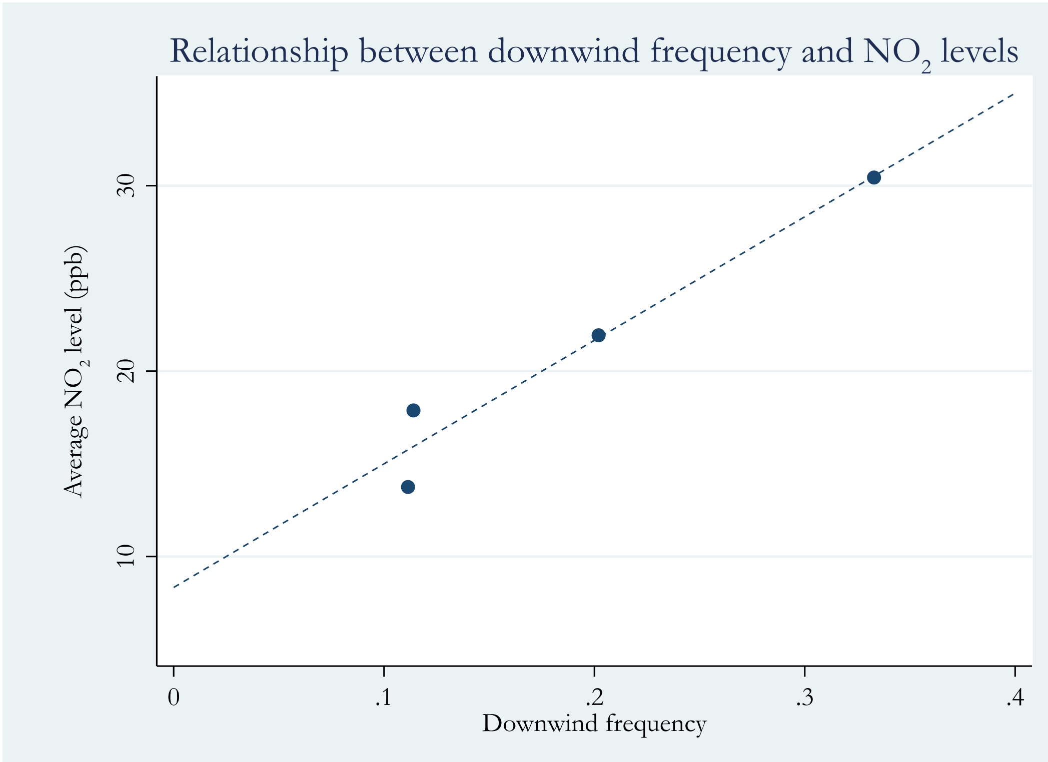

Figure 5

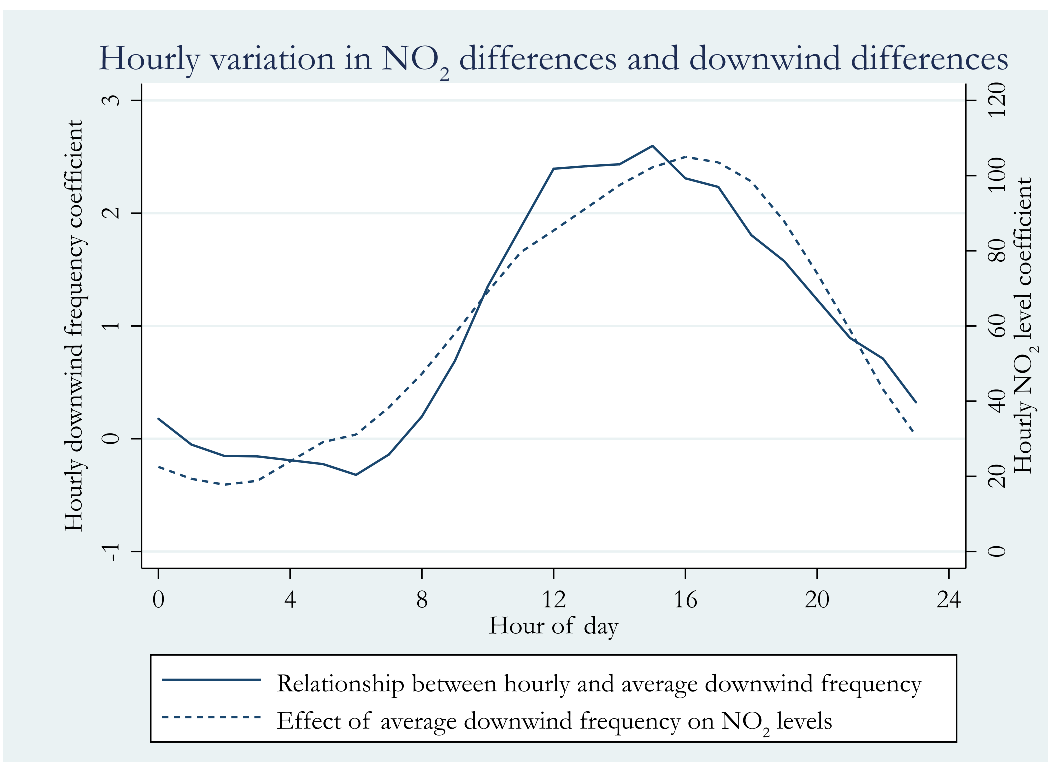

Table 1: Summary statistics

	Full sample			Analytic sample		
	Mean	Range	N	Mean	Range	N
3-yr mortality rate among 75+ year olds from:						
All causes	0.157 (0.178)	0.000–0.857	27,908	0.154 (0.179)	0.000–0.833	9,314
Cardio-respiratory	0.103 (0.145)	0.000–0.800	27,908	0.102 (0.147)	0.000–0.800	9,314
Cancer	0.029 (0.076)	0.000–0.667	27,908	0.029 (0.077)	0.000–0.667	9,314
External causes	0.002 (0.019)	0.000–0.500	27,908	0.002 (0.020)	0.000–0.500	9,314
Other variables						
Frequency downwind of major highway	0.150 (0.130)	0.003–0.490	27,908	0.153 (0.130)	0.003–0.490	9,314
Frequency dead wind	0.423 (0.135)	0.101–0.660	27,908	0.420 (0.131)	0.101–0.660	9,314
Population	164.8 (180.2)	1–6,375	27,908	166.9 (167.9)	1–2,215	9,314
Population aged 75+	8.5 (15.1)	1–542	27,908	8.0 (15.3)	1–542	9,314
Distance to highway (meters)	1,182 (1,035)	0–7,666	27,908	313 (158)	50–600	9,314
Share owner occupied	0.548 (0.323)	0–1	27,869	0.521 (0.317)	0–1	9,301
Share black	0.143 (0.240)	0–1	27,908	0.118 (0.202)	0–1	9,314

Notes: The observation is the Census Block. Parentheses contain standard deviations. The analytic sample is limited to Census Blocks with centroids between 50 and 600 meters from major highways.

Table 2: OLS Effects of Frequency Downwind of Highways

	3-year mortality rate for 75+ year olds from:							
Dependent Variable:	All causes		Cardio-respiratory		Lung cancer		Other cancer	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Frequency downwind	0.036	0.062	0.019	0.042	0.002	0.006	0.007	0.005
	(0.016)	(0.017)	(0.013)	(0.014)	(0.003)	(0.004)	(0.007)	(0.009)
Highway segment FE		Yes		Yes		Yes		Yes
Dependent variable mean	0.154	0.154	0.102	0.102	0.006	0.006	0.022	0.022
N	9,314	9,314	9,314	9,314	9,314	9,314	9,314	9,314

Notes: Each column represents a separate regression of the dependent variable on the percent of time spent downwind of a major highway. The observation is the Census Block, and the sample is limited to Census Blocks with centroids between 50 and 600 meters from major highways. Parentheses contain spatial standard errors with a 3,200 meter bandwidth. All regressions include controls for distance to highway and weather station fixed effects. Regressions without highway segment fixed effects include quintics in latitude and longitude and first and second order interactions between latitude and longitude. Regressions with highway segment fixed effects include highway segments fixed effects with an 800 meter bandwidth.

Table 3: First-stage relationship between bearing to highway and frequency downwind

Dependent Variable:	Frequency downwind	
	(1)	(2)
Highway northeast	0.005	-0.025
	(0.010)	(0.009)
Highway east	0.031	-0.023
	(0.016)	(0.023)
Highway southeast	0.070	0.054
	(0.014)	(0.012)
Highway south	0.108	0.109
	(0.013)	(0.012)
Highway southwest	0.222	0.195
	(0.026)	(0.024)
Highway west	0.243	0.188
	(0.036)	(0.034)
Highway northwest	0.028	0.025
	(0.028)	(0.022)
F-statistic	30.6	26.6
Partial R^2	0.552	0.656
Highway segment FE		Yes
N	9,314	9,314

Notes: Each column represents a separate regression of the frequency downwind on seven indicators summarzing bearing to the nearest major highway. The omitted category is north. The observation is the Census Block, and the sample is limited to Census Blocks with centroids between 50 and 600 meters from major highways. Parentheses contain standard errors clustered on a spatial grid with a width of 0.05 degrees longitude or latitude in each cell. All regressions include controls for distance to highway and weather station fixed effects. Regressions without highway segment fixed effects include quintics in latitude and longitude and first and second order interactions between latitude and longitude. Regressions with highway segment fixed effects include highway segments fixed effects with an 800 meter bandwidth. The *F*-statistic tests the hypothesis that all seven bearing indicators equal zero; the partial R^2 is the R^2 generated by these seven bearing indicators after partialing out controls.

Table 4: 2SLS Effects of Frequency Downwind of Highways

	3-year mortality rate for 75+ year olds from:							
Dependent Variable:	All causes		Cardio-respiratory		Lung cancer		Other cancer	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Frequency downwind	0.064	0.068	0.033	0.036	0.003	0.008	0.014	0.015
	(0.022)	(0.023)	(0.017)	(0.018)	(0.004)	(0.004)	(0.009)	(0.010)
Highway segment FE		Yes		Yes		Yes		Yes
Dependent variable mean	0.154	0.154	0.102	0.102	0.006	0.006	0.022	0.022
Over-ID test *p*-value	0.13	0.12	0.49	0.24	0.29	0.59	0.32	0.48
N	9,314	9,314	9,314	9,314	9,314	9,314	9,314	9,314

Notes: Each column represents a separate 2SLS regression of the dependent variable on the instrumented percent of time spent downwind of a major highway. The observation is the Census Block, and the sample is limited to Census Blocks with centroids between 50 and 600 meters from major highways. The instruments are a set of seven indicator variables summarizing bearing to the nearest major highway. Parentheses contain spatial standard errors with a 3,200 meter bandwidth. All regressions include controls for distance to highway and weather station fixed effects. Regressions without highway segment fixed effects include quintics in latitude and longitude and first and second order interactions between latitude and longitude. Regressions with highway segment fixed effects include highway segments fixed effects with an 800 meter bandwidth. Overidentification test statistics are for Sargan's chi-squared test.

Table 5: Robustness of Effects to Different Spatial Parameters

Dependent Variable:	3-year mortality rate for 75+ year olds from: All causes		Cardio-respiratory		N
	(1)	(2)	(3)	(4)	
Estimation method:	OLS	2SLS	OLS	2SLS	
Baseline estimate	0.062	0.068	0.042	0.036	9,314
	(0.017)	(0.023)	(0.014)	(0.018)	
Modification:					
Cosine-weighted downwind frequency	0.058	0.059	0.037	0.033	9,314
	(0.017)	(0.021)	(0.014)	(0.017)	
400 m hwy segment FE	0.072	0.072	0.052	0.039	9,314
	(0.018)	(0.024)	(0.015)	(0.020)	
1,600 m hwy segment FE	0.054	0.065	0.036	0.037	9,314
	(0.017)	(0.022)	(0.014)	(0.017)	
25 m "donut" around hwy	0.053	0.054	0.033	0.024	9,601
	(0.018)	(0.024)	(0.014)	(0.019)	
100 m "donut" around hwy	0.077	0.083	0.047	0.044	8,461
	(0.021)	(0.026)	(0.016)	(0.020)	
Within 400 m of hwy	0.048	0.057	0.033	0.034	6,240
	(0.020)	(0.027)	(0.017)	(0.022)	
Within 800 m of hwy	0.064	0.062	0.036	0.026	12,230
	(0.015)	(0.020)	(0.012)	(0.016)	

Notes: Each cell represents a separate regression of the dependent variable on the percent of time spent downwind of a major highway (OLS) or instrumented percent of time spent downwind of a major highway (2SLS). The observation is the Census Block, and the sample is limited to Census Blocks with centroids between a minimum of 25/50/100 meters (50 m is the baseline) and a maximum of 400/600/800 meters from major highways (600 m is the baseline). The instruments are a set of seven indicator variables summarizing bearing to the nearest major highway. Parentheses contain spatial standard errors with a 3,200 meter bandwidth. All regressions include controls for distance to highway, weather station fixed effects, and highway segment fixed effects.

Table 6: Effects of Frequency Downwind for Different Age Groups

Dependent Variable:	3-yr mortality rate for 70+ year olds from: All causes		3-yr mortality rate for 70+ year olds from: Cardio-respiratory		3-yr mortality rate for 65+ year olds from: All causes		3-yr mortality rate for 65+ year olds from: Cardio-respiratory		3-yr mortality rate for 50+ year olds from: All causes		3-yr mortality rate for 50+ year olds from: Cardio-respiratory	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Estimation method:	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
Frequency downwind	0.040	0.050	0.030	0.033	0.026	0.032	0.024	0.020	0.005	0.009	0.003	0.002
	(0.015)	(0.020)	(0.011)	(0.016)	(0.011)	(0.016)	(0.009)	(0.012)	(0.006)	(0.009)	(0.005)	(0.006)
Dependent variable mean	0.135	0.135	0.086	0.086	0.115	0.115	0.071	0.071	0.062	0.062	0.036	0.036
N	10,096	10,096	10,096	10,096	10,512	10,512	10,512	10,512	11,040	11,040	11,040	11,040

Notes: Each cell represents a separate regression of the dependent variable on the percent of time spent downwind of a major highway (OLS) or instrumented percent of time spent downwind of a major highway (2SLS). The observation is the Census Block, and the sample is limited to Census Blocks with centroids between a minimum of 50 meters and a maximum of 600 meters from major highways. The instruments are a set of seven indicator variables summarizing bearing to the nearest major highway. Parentheses contain spatial standard errors with a 3,200 meter bandwidth. All regressions include controls for distance to highway, weather station fixed effects, and 800 meter highway segment fixed effects.

Table 7: Effects of Frequency Downwind on Placebo Measures

					External-cause mortality rate among:			
Dependent Variable:	Share owner occupied		Share African-American		75+ year olds		All residents	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation method:	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
Frequency downwind	0.081	0.090	-0.009	-0.035	-0.0001	-0.0004	-0.0013	-0.0020
	(0.071)	(0.074)	(0.028)	(0.036)	(0.0017)	(0.0023)	(0.0009)	(0.0014)
Dependent variable mean	0.521	0.521	0.118	0.118	0.0021	0.0021	0.0013	0.0013
N	9,301	9,301	9,314	9,314	9,314	9,314	9,314	9,314

Notes: Each cell represents a separate regression of the dependent variable on the percent of time spent downwind of a major highway (OLS) or instrumented percent of time spent downwind of a major highway (2SLS). The observation is the Census Block, and the sample is limited to Census Blocks with centroids between a minimum of 50 meters and a maximum of 600 meters from major highways. The instruments are a set of seven indicator variables summarizing bearing to the nearest major highway. Parentheses contain spatial standard errors with a 3,200 meter bandwidth. All regressions include controls for distance to highway, weather station fixed effects, and 800 meter highway segment fixed effects.

Table 8: Effects of Frequency Downwind on Property Values

Dependent Variable:	Log price				Square feet	
	(1)	(2)	(3)	(4)	(5)	(6)
Estimation method:	OLS	2SLS	OLS	2SLS	OLS	2SLS
<u>Panel A: 1990-2000 sales</u>						
Frequency downwind	0.017	-0.113	-0.037	-0.140	208.5	124.7
	(0.078)	(0.173)	(0.049)	(0.094)	(186.2)	(351.3)
Control for log sq ft and date sold			Yes	Yes		
Dependent variable mean	11.961	11.961	11.961	11.961	1,311.2	1,311.2
N	21,456	21,456	21,456	21,456	21,456	21,456
<u>Panel B: 2006-2010 sales</u>						
Frequency downwind	0.017	-0.026	-0.069	-0.091	162.3	97.1
	(0.094)	(0.103)	(0.058)	(0.065)	(138.5)	(144.5)
Control for log sq ft and date sold			Yes	Yes		
Dependent variable mean	13.029	13.029	13.029	13.029	1,385.1	1,385.1
N	21,713	21,713	21,713	21,713	21,713	21,713

Notes: Each cell represents a separate regression of the dependent variable on the percent of time spent downwind of a major highway (OLS) or instrumented percent of time spent downwind of a major highway (2SLS). The observation is a housing sale, and the sample is limited to sales in Census Blocks with centroids between a minimum of 50 meters and a maximum of 600 meters from major highways. The instruments are a set of seven indicator variables summarizing bearing to the nearest major highway. Parentheses contain spatial standard errors with a 3,200 meter bandwidth. All regressions include controls for distance to highway, weather station fixed effects, and 800 meter highway segment fixed effects. Regressions in Columns (3) and (4) include controls for log square footage and a cubic in time of sale (measured at the daily frequency).

Table 9: Relationship Between NO_2 and Frequency Downwind of Highways

Dependent Variable:	NO_2 Concentration (ppb)	
	(1)	(2)
Frequency downwind	70.5	61.1
	(2.1)	(1.6)
Day-of-sample FE		Yes
Dependent variable mean	27.4	27.4
N	400,218	400,218

Notes: Each column represents a separate OLS regression of the dependent variable on the percent of time spent downwind of a major highway. The observation is the hour-by-site. Parentheses contain standard errors clustered by month of sample.