Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health

CAUTION, DRIVERS! CHILDREN PRESENT: TRAFFIC, POLLUTION, AND INFANT HEALTH

Christopher R. Knittel, Douglas L. Miller, and Nicholas J. Sanders*

Abstract—We investigate the effects of automobile congestion on ambient air pollution and local infant mortality rates using data from California spanning 2002 to 2007. Constructing instrumental variables (IV) using the relationship of traffic, weather conditions, and pollutants, we show that particulate matter, even at modern levels, has large marginal effects on weekly infant mortality rates, especially for premature or low birthweight infants. We also find suggestive evidence of large effects for carbon monoxide, though results are imprecise. Finally, we check estimate sensitivity to nonclassical measurement error in local pollution and show that our IV results are robust to such concerns.

I. Introduction

LOCAL air pollution levels decreased dramatically after the passage of the Clean Air Act and its various amendments, which placed strict limits on the concentrations of criteria pollutants.[1] The Clean Air Act Amendments of 1990 (CAAA) helped decrease the concentration of carbon monoxide (CO) by 75% and particulate matter 10 micrometers or smaller (PM10) by 39% from 1990 to 2012.[2] These reductions have been financially costly. The Environmental Protection Agency (EPA) estimates the compliance costs of the CAAA to be $19 billion annually in 2000, increasing to $27 billion by 2010.[3] Over half of these costs are due to the CAAA's National Ambient Air Quality Standards, regulating point and area sources. Regulation of mobile sources accounts for an additional 30%.[4]

We study pollution from mobile sources, investigating the relationship of traffic, pollution, and mortality rates among infants in the state of California. In doing so, we provide the first large-scale study of how road traffic affects local health, with a focus on how regional traffic affects infant mortality rates. We find that higher levels of automobile traffic increase infant mortality and explore the heterogeneity of effects by distance from sources, local weather conditions, and affected subgroups. Effects of traffic on mortality are largely local and, consistent with atmospheric chemistry, greater during periods of lower temperature or higher humidity. We find mortality effects are concentrated among more marginalized infants: those who are premature or of low birthweight. We then investigate the direct role of pollution on infant mortality, using interactions between mobile source emissions and local weather conditions as a source of exogenous variation in pollution in an instrumental variables (IV) setting.

As part of our analysis, we consider how local weather conditions influence pollution levels that similar levels of traffic generate. This interaction between traffic levels and weather conditions is the basis for our instrumental variables analysis of the effect of both PM10 and CO on weekly infant mortality.[5] We gain two valuable benefits of identification by using interactions between weather and traffic as our overall drivers of shifts in local pollution. First, while long-term traffic patterns may correlate with other health-related factors (e.g., economic development) in ways for which our various fixed effects do not control, weekly weather conditions are uncorrelated with these factors over time. Second, we can simultaneously estimate the effects of both CO and PM10 using changes in traffic, where different weather conditions result in different pollution levels by pollutant. We find PM10 has a statistically and economically significant impact on weekly infant mortality—a 1 unit decrease in PM10 saves roughly 10 lives per 100,000 live births, an elasticity of approximately 1. We also find large (though statistically insignificant) negative health effects of CO.

Our work builds on Currie and Neidell (2005, hereafter CN), who examine how California's reductions in carbon monoxide, particulate matter, and ozone during the 1990s reduced weekly infant mortality rates. As an extension, we formally consider the potential role of nonclassical measurement error in assigning local pollution levels and through a Monte Carlo exercise show our IV estimates are robust to such concerns. We also expand on Currie and Walker (2011), who consider reductions in traffic from installation of high-speed toll stations and decreased infant mortality close to toll booths, though with little information on pollution or geographic variation. A major benefit of our analysis is the ability to more directly link traffic to mortality and pollution, as well as expanding the traffic–mortality relationship beyond regions close to toll zones.

Received for publication September 20, 2011. Revision accepted for publication January 9, 2015. Editor: Mark W. Watson.

* Knittel: MIT and NBER; Miller: University of California, Davis, and NBER; Sanders: Cornell University.

Thanks to Janet Currie; Matthew Neidell; seminar participants at UC Davis, UC San Diego, and the University of California Energy Institute; and attendants of the NBER Summer Institute and Toxic Substances Research and Teaching Program Symposium (TSR&TP). C.K. gratefully acknowledges financial support from the University of California Energy Institute and UC Davis Institute of Transportation Studies. N.S. gratefully acknowledges funding from the UC Davis Institute of Governmental Affairs, the UC Davis Institute of Transportation Studies, and the TSR&TP through the Atmospheric Aerosols and Health Lead Campus program.

[1] The term *criteria pollutants* refers to six commonly found air pollutants regulated by developing health-based or environmentally based criteria for allowable levels. The pollutants are particular matter, ground-level ozone, carbon monoxide, sulfur oxides, nitrogen oxides, and lead.

[2] From http://www.epa.gov/air/airtrends/.

[3] Benefits from air quality improvements are more difficult to measure. Estimates often rely on correlations between pollution levels and health outcomes that may not reflect causal relationships. The EPA (1999) estimates a wide range for the potential benefits in 2000—from a low of $16 billion to a high of $140 billion (available at http://www.epa.gov/air/sect812/). This range reflects uncertainty with respect to how specific sources affect air quality and how increasing air quality improves health outcomes.

[4] Available at http://www.epa.gov/air/sect812/.

[5] Knittel, Miller, and Sanders (2011) include ozone as an additional pollutant. We find no consistently economically or statistically significant results. As ozone is a secondary pollutant and not directly created by automobile pollution, here we focus on CO and PM10.

The Review of Economics and Statistics, May 2016, 98(2): 350–366

doi:10.1162/REST_a_00548

We contribute to the growing literature demonstrating the use of applied microeconometric techniques in questions of environmental quality and health. Recent examples of such work include the effect of air pollution on infant mortality and birth outcomes (Chay & Greenstone, 2003a, 2003b; Currie & Neidell, 2005; Currie, Neidell, & Schmieder, 2009; Currie & Walker, 2011; Sanders & Stoecker, 2015); contemporaneous health factors (Chay, Dobkin, & Greenstone, 2003; Neidell, 2004; Currie et al., 2009; Lleras-Muney, 2010; Moretti & Neidell, 2011); and life cycle outcomes (Sanders, 2012). Research on other pollution effects includes climate change (Deschênes & Greenstone, 2007, 2011; Deschênes, Greenstone, & Guryan, 2009; Stoecker, 2010), environmental toxins (Reyes, 2007; Currie & Schmieder, 2009), and radiation (Almond, Edlund & Palme, 2009). Ours is the first panel fixed-effects analysis of the effect of traffic on both ambient pollution and health based on weather conditions. Prior studies of the link between auto emissions and pollution are in laboratory environments or specific limited regions using small numbers of computer-monitored automobiles or roadside emission sensors over a limited driving range (see Bishop and Stedman, 1996; Tiao & Hillmer, 1978).[6] While such studies are informative, there remain large gains to understanding the effects of traffic in a real-world setting.

Section II describes our data sources and data set construction. Section III summarizes the chemistry of driving and air pollution, the physiology of air pollution and infant health, the relevant transportation literature on traffic measurement, and the relevant economics literatures on traffic externalities and air pollution's impacts on infant health. Section IV outlines our empirical methodology, section V presents our main results, explores potentially nonclassical measurement error in pollution assignment, and demonstrates the robustness of our IV results to such concerns. Section VI offers concluding remarks.

II. Data

We perform all analysis at the postal code–week level and aggregate data from each source accordingly.[7]

A. *Pollution and Weather Data*

We obtain pollution data from the California Air Resources Board (CARB) website.[8] The data contain daily pollution measures for carbon monoxide and particulate matter smaller than 10 micrometers. CO is a maximum daily 8-hour values, expressed in parts per million. PM10 data are 24-hour averages measured once every six days, expressed in micrograms per cubic meter of air. We calculate weekly averages of daily values and, to obtain a postal code–level measure, calculate the distance between the postal code geographic centroid and each monitor station based on latitude and longitude location information. We then weight each station by 1 over its distance from the centroid using monitors within 20 miles of a centroid. This introduces a potential source of measurement error in assignment of local pollution. We formally investigate this issue using simulation of autocorrelated measurement error in section VF.

Weather data are from the National Climatic Data Center Global Surface Summary of the Day. We use continuous measures on inches of rainfall, maximum daily temperature, average daily windspeed, and specific humidity and further control for counts of days within the week in which there was rain or fog.[9] To calculate a postal code–level weather variable, in the case of continuous variables we use the weighting method discussed above, using weather stations within 20 miles of a postal code centroid.[10] In the case of count data of rainy or foggy days, we use the local maximum of nearby monitors.

B. *Traffic Data*

Data on traffic come from the Freeway Performance Management System (PeMS), maintained by the University of California, Berkeley Department of Electrical Engineering and Computer Sciences.[11] Using sensors (called "loops") buried beneath freeway lanes, the PeMS records data including average speed and total flow of cars. Measurements occur every 30 seconds, and the PeMS computer system aggregates data to longer periods. Traffic data are available from 1999 onward, though many regions were not continuously available until 2002, leading to our chosen time period of analysis. We focus on regions of California for which there is the greatest number of continuously monitored traffic data: the Sacramento Valley, the Bay Area, and the Los Angeles Basin area (regions 3, 4, 7, 11, and 12 in the PeMS data).

We construct our measure of traffic based on two PeMS data items: total flow of cars and length of sensor region. Our preferred metric for traffic approximates average traffic per section of road. Total flow of cars is the count of all cars that pass over a section of road within a time frame. Sensors represent sections of road, and each sensor (or "loop") is assigned a "length" of relevant road about which it provides information. PeMS measures this length by (a) taking the midpoint between a sensor and the next sensor after it,

[6] Currie and Walker (2011) consider the impact of EZ-Pass toll booth modification on health, however, they have little direct information on traffic flows and pollution.

[7] Ideally, we would like to use richer spatial data to identify mothers' proximity to roadways, but the state of California does not provide infant birth and mortality data with more specific geographic region information.

[8] http://www.arb.ca.gov/aqd/aqdcd/aqdcddld.htm.

[9] We do not make spatial adjustments for the issue of wind direction, which may introduce noise into our first stage. Assuming this noise is random (i.e., wind direction is not associated with factors that drive traffic), the error should not affect the consistency of our IV estimates. Schlenker and Walker (2014) formally consider the role of wind direction in the dispersion of CO pollution from airports.

[10] The Global Surface Summary does not report specific humidity, which is the most relevant for mortality. We calculate specific humidity using dewpoint and air pressure as discussed in Barreca (2008).

[11] https://pems.eecs.berkeley.edu.

TABLE 1.—MEANS AND STANDARD DEVIATIONS FOR POLLUTION AND WEATHER VARIABLES

	CO	PM10	Rain	Maximum Temperature	Windspeed	Humidity	Days with Rain	Days with Fog
2002	1.23	31.63	0.02	72.91	5.53	7.76	1.03	2.87
2003	1.16	29.32	0.03	73.81	5.27	8.25	1.32	2.57
2004	1.01	28.96	0.03	73.56	5.40	7.96	1.15	0.66
2005	0.94	26.09	0.05	73.20	5.21	7.90	1.50	2.90
2006	0.90	28.73	0.03	73.21	5.08	7.65	1.57	0.75
2007	0.81	28.93	0.02	72.99	5.11	7.19	1.12	0.62
Total	1.01	28.94	0.03	73.28	5.27	7.79	1.28	1.73
Overall SD	0.57	14.94	0.08	9.31	1.97	2.12	1.54	2.07
Within SD	0.52	12.89	0.08	8.80	1.36	2.03	1.53	1.96
Between SD	0.23	7.65	0.01	3.04	1.42	0.63	0.23	0.69

Cells report unweighted averages of each variable over postal code–week observations within each year. The geographic coverage is the 719 postal codes used in our full analysis, primarily covering the Sacramento Valley and Southern California. Units for CO are parts per million, and units for PM10 are micrograms per cubic meter of air. Units for rainfall are average inches per day. Unit for wind is average wind speed in miles per hour. Unit for humidity is (100 times) the ratio of water vapor to dry air in air space. Units for days with rain and days with fog range from 0 to 7 per week. "Overall SD" is standard deviation across all postal code–week observations. "Within SD" is the standard deviation after absorbing postal code fixed effects. "Between SD" is the standard deviation of the postal code fixed effects. Authors' calculations from EPA and NOAA data. See section II for further details.

(b) taking the midpoint between a sensor and the last sensor before it, and (c) measuring the distance between these two midpoints. In the case of no additional sensors in one (or both) direction(s), PeMS assigns a maximum distance of 2.5 miles per direction. We multiply by this length to get the traffic density per section of road.[12] Our preferred traffic measure is effectively local car miles driven:

$$\begin{aligned} \textit{Carmiles} &= \textit{Total Flow per Sensor Length} \\ &\quad \times \textit{Sensor Length}. \end{aligned} \tag{1}$$

As an example, consider a loop with a reading of 6,000 cars per hour, with a loop length of 2 miles. If the loop "represents" 2 miles of road, we multiply those 6,000 cars by 2 to illustrate that traffic was 6,000 cars driving for 2 miles. We do this to help continuity in traffic measures across regions with more versus fewer sensors for the same length of road. To obtain a weekly value, we sum hourly values over the week.[13]

To calculate a postal code–level traffic measure, we use traffic sensors up to 20 miles from a postal code centroid. We first consider the role of traffic in both local health and pollution, allowing different marginal impacts of traffic across 5-mile intervals.[14] Based on observed relationships between distance and the marginal effect of additional traffic, we then simplify to a single traffic measure by weighting traffic flows in terms of equivalent traffic flow directly at the postal code centroid.[15] We define an individual postal code traffic measure using sensors $s = 1, \ldots, n$ in week w as

$$\textit{Traffic}_{z,w} = \sum_{s=1}^{n} \textit{Carmiles}_{s,w} \times \textit{weight}. \tag{2}$$

We begin with an agnostic view on the distance weights and use our distance analysis of the first stage to construct a weighting scheme. Given the high traffic volumes in our regions of study, our measure of numbers is large: the mean level of unweighted *carmiles* within 20 miles is 83 million. To make summary statistics and coefficients more easily readable, we divide all weekly totals by 10 million. Table 1 shows means across time and standard deviations for all weather, traffic, and pollution variables. Columns 1 through 4 of table 2 show means and standard deviations for traffic variables with no distance weighting.

PeMS data and imputation. PeMS constantly checks detector loops for problems with reported data, which could mean either missing data or incorrectly reported data. Such data problems are quite common. In any given 24-hour period, a sensor could report a total of 2,880 samples, while the actual daily sampling rate is often 50% to 90% of daily potential.[16] In the case of missing or faulty data, PeMS uses imputation techniques to adjust for the missing information.

Imputation techniques vary based on the amount of replacement needed. In the case of few missing data, PeMS uses linear regression from nearby neighbor loops, such as a loop in another lane on the same freeway, or upstream or downstream of the current loop. If certain loop regions require imputation for longer periods, PeMS uses predictions from similar loops using regression models over a larger geographic area. When no nearby monitors are available, PeMS uses temporal medians from past days at the same time to

[12] We use a balanced panel of available loops across time for consistently of measurement. As PeMS brings more loops online, the "length" of an older sensor loop can change. For consistency, we use the initial sensor length in building our traffic variable.

[13] Knittel et al. (2011) further control for average speed. However, average speed from single-loop detectors (the most commonly used) requires estimation that is highly variable and, when compared to more accurate double-loop detectors, often incorrect. As such, we now omit speed controls. For details on the use of speed calculations in PeMS data, see http://robotics.eecs.berkeley.edu/~varaiya/papers_ps.dir/gfactoritsc.pdf.

[14] Given our focus on the role of traffic in health, we omit regions with no measured traffic activity (though results are robust to their inclusion).

[15] Unlike weighting with pollution and weather, weights here do not reflect a measure of accuracy. In the case of pollution and weather information, sensors represent a sample of ambient conditions near a particular location, and each additional reading is more information regarding the true level. The closer the measurement is to that location, the more accurately we expect it to reflect the true measure, and thus we apply greater weight to that information. With traffic, more loop miles mean more traffic, not just more information on the true traffic level.

[16] From *PeMS Data Extraction Methodology and Execution Technical Memorandum for the Southern California Association of Governments*, prepared June 2006 by Urban Crossroads, http://web.scag.ca.gov/modeling/pdf/Pems_Technical_Memorandum_Final.pdf.

Table 2.—Means and Standard Deviations for Traffic Variables

	Weekly Car-Miles				
	0–5 Miles	5–10 Miles	10–15 Miles	15–20 Miles	0–15 Miles
2002	0.91	2.06	2.67	2.75	1.32
2003	0.90	2.04	2.66	2.74	1.31
2004	0.93	2.09	2.72	2.80	1.34
2005	0.91	2.04	2.66	2.73	1.31
2006	0.91	2.06	2.68	2.74	1.32
2007	0.92	2.07	2.70	2.76	1.33
Total	0.91	2.06	2.68	2.75	1.32
Overall SD	0.91	1.79	2.10	2.16	1.07
Within SD	0.05	0.10	0.12	0.40	0.06
Between SD	0.91	1.79	2.09	2.12	1.07
	5 miles	10 miles	15 miles	20 miles	
Avgerage loops in distance	22	52	69	71	
Postal codes with traffic at distance	594	647	681	713	

Cells report unweighted averages of each variable over postal code–week observations within each year. The geographic coverage is the 719 postal codes used in our full analysis, primarily covering the Sacramento Valley and Southern California. Units for traffic are in total flow of cars (10 million) times length of loop and approximate car-miles traveled within the noted distance. "Overall SD" is standard deviation across all postal code–week observations. "Within SD" is the standard deviation after absorbing postal code fixed effects. "Between SD" is the standard deviation of the postal code fixed effects. Authors' calculations from PeMS data. Average number of loops displays the average traffic loops within a particular distance from a postal code centroid. Postal codes with nonzero traffic shows how many postal codes have at least one loop within the specified distance range. See section II for further details.

estimate expected data for the particular loop. Finally, in the case of long-term failure of a monitor, PeMS may use temporal medians from nearby loops.

How the PeMS imputation influences our estimates depends on the nature of loop failures. If imputations result in higher or lower than correct measures in a manner unchanging with time or changing seasonally, regional month fixed effects remove this source of bias (this would be the case, for example, if a region had sensors that were continuously imputed across our entire period). If failures and required imputations are random, this introduces an additional source of noise that will bias both reduced-form estimates and first-stage estimates toward 0. Potentially problematic is if failures correspond to unusual traffic events—for example, higher or lower traffic flows than usual cause loops to provide a higher percentage of incorrect readings. Even so, if which loops have these problems is random, local imputation data based on neighboring lanes and upstream or downstream loops are likely good proxies for true traffic measurements, especially when considering count data on the total number of cars (which is unlikely to vary drastically across lanes).

Finally, extreme traffic events such as accidents or road closures causing entire sections to result in faulty data mean PeMS would use district-level measures to impute larger sections of localized data. True zero traffic counts viewed as errors means imputation using regional traffic estimates would place nonzero counts. This would artificially decrease the variation in our traffic data. If a loop cluster had unusually high counts that shut down loops or PeMS replaced with regional data, imputation will again mute the signal, making results we find underestimates of true effects.

C. *Birth Data*

Birth data are from the California Department of Public Health Birth Cohort files, where the department links birth and death files if an infant dies within 52 weeks of birth. This allows us to link any infant who dies within the first year of life to his or her birth outcomes and maternal information. We limit our sample to infants with a gestation period of at least 26 weeks (the beginning of the third trimester), which allows us to assign a trimester-level pollution exposure to every infant for all three trimesters.[17] We then convert all birth and death dates to the weekly postal code level. Aside from providing the time of birth and death and the birth mother's postal code of residence, the Birth Cohort files also provide us with various controls to use in the analysis. These include mother's race, education, and age; potentially confounding birth outcomes (low birthweight and premature birth); public insurance coverage; birth order; infant sex; and, in the case of those who died, the age in weeks at death. Table 3 shows means for variables of infants covered in postal codes in our analysis.

III. The Relationships between Traffic, Weather, and Ambient Pollution

A. *Traffic and Pollution*

Research ties both CO and PM10 to automobile traffic. Up to 90% of all CO in the United States comes from automobile fuel combustion, and automobiles increase PM levels through multiple processes.[18] Fuel combustion results in tailpipe emissions of particulates (e.g., formation of nitrogen oxides, volatile organic compounds, and, in the case of diesel engines, diesel soot), a large portion of which are small scale and especially damaging to health. Traffic also generates particulates through the physical act of friction resulting from wheel-to-road contact, brake, tire, and

[17] We drop infants with gestation lengths greater than 42 weeks, as doctors are likely to induce labor by this period and such values are probably reporting or coding errors. Due to the use of traffic as our instrument for pollution, we drop all deaths caused specifically by auto trauma.

[18] http://www.epa.gov/oms/consumer/03-co.pdf.

TABLE 3.—MEANS FOR INFANT DATA

Male	0.513
African American	0.059
Asian	0.101
Hispanic	0.492
Other race	0.070
Mother is HS graduate	0.722
Mother is college graduate	0.293
Twins	0.030
Triplets or more	0.002
Mother age 19–25	0.274
Mother age 26–30	0.269
Mother age 31–35	0.261
Mother age > 35	0.152
Medicaid	0.401
Care first trimester	0.903
Low birthweight	0.064
Premature	0.044
Year	Annual Mortality Rates
2002	0.00282
	(0.0530)
2003	0.00280
	(0.0528)
2004	0.00276
	(0.0525)
2005	0.00285
	(0.0534)
2006	0.00281
	(0.0530)
Total	0.00281
	(0.0529)

Cells report unweighted averages of individual birth-level data. Death is an indicator variable, with means reported as deaths per 1,000 births. All other variables are indicator variables with means reported as proportions. These primarily cover the Sacramento Valley and Southern California. Authors' calculations from California linked Birth-Death Vital Statistics records. See section II for further details.

gear wear.[19] Given the link between combustion engines and the pollutants considered in this analysis, we anticipate that automobile use and traffic levels affect ambient air pollution through three main channels. First and most obvious is that a greater number of cars on the freeway at any given time results in more fuel burned and more tires on the road. Second, traffic congestion can increase the amount of pollution each individual car creates. Efficiency of automobile combustion is directly related to average travel speed and continuity of driving (Davis & Diegel, 2007), and engines have an optimal revolutions per minute (RPM) range in which they obtain the maximum amount of power for any given amount of fuel. Stop-and-go traffic means fluctuations in the engine revolutions per minute and less time within the optimal RPM range.[20] This also means greater amounts of PM10 generated through additional brake and gear wear. Finally, traffic congestion can decrease the average speed of each vehicle on the road. At a given RPM (and engine efficiency), a slower speed implies more time on the road to travel the same distance, and thus more fuel burned (and emissions created) for each mile traveled.[21]

B. *Pollution, Weather, and Mortality*

As a source of identification, we take advantage of the different impacts of traffic on pollution given local weather conditions. For example, particulate matter and carbon monoxide levels are often higher under temperature inversion, an atmospheric condition caused by differences in upper- and ground-level air temperatures which Arceo-Gomez, Hanna and Oliva (forthcoming) use as a shock to local pollution levels.[22] Inversion is particularly problematic in valley areas, as surrounding mountains serve as containment for the inversion weather system, making it even harder for the air to circulate; during the summer, daily inversions are common in California. Humidity, wind, and rain also influence pollution levels. CO has an oxidation rate that changes with humidity (Lee et al., 1995), and high humidity conditions favor some chemical reactions that create particulate matter.[23] Higher wind speeds can disperse pollutants or increase atmospheric chemical reactions, while rain can decrease both gaseous pollutants and particulate matter through a combination of absorption and water entrapment (for a theoretical analysis of this issue, as well as a discussion of empirical findings, see Shukla et al., 2008) and sometimes increase particulate matter by placing particles onto roadways to then be kicked up by automobile tires when conditions dry up.

We control for a rich a set of weather variables in all regressions, both independent and interacted with traffic conditions. A benefit of such weather–pollution relationships is that interactions between traffic and weather allow us to better identify conditions that are more conducive to traffic causing higher levels of specific pollutants. High traffic levels during hot, windy weeks create different amounts of different pollutants than high traffic levels during cold weeks with stagnant air. Including weather interaction variables allows for simultaneous instrumentation for both pollutants of interest despite only one traffic measure. In addition, it provides an additional source of exogenous variation to our estimates. In cases where our fixed effects are insufficient to control for all "bad" variation, weather conditions remain unlikely to correlate with omitted variables of concern.

Weather controls are also important for mortality analysis. Previous work finds a relationship between weather

[19] Automobiles emit pollutants beyond those we observe here. For example, automobile fuel combustion creates carbon dioxide, volatile organic compounds (which contribute to both particulate matter and ozone formation), nitrogen oxides (also related to ozone), and benzene. These pollutants may impact mortality, and due to likely correlation with our measured pollutants, effects may be picked up by one of the two pollutions for which we control. We interpret our results with this caveat in mind.

[20] RPM variation is also a major factor determining the difference between automobile fuel efficiency in freeway versus city driving.

[21] We note that despite a known scientific relationship between traffic and pollution, correlations in reality are more complicated. Most cars are more efficient at RPMs corresponding to speeds of 45 to 60 mph (Davis & Diegel, 2007). If unhindered traffic flow is moving at speeds above the range of highest efficiency, mild amounts of traffic that slightly lower traveling speeds can increase engine efficiency and decrease emissions.

[22] Temperature inversion results when a layer of warmer air settles over a layer of colder air. Such atmospheric conditions often correlate with movements of air pressure systems. The warm air layer prevents ground-level air from circulating, and the stagnant air creates a buildup of ground-level pollution.

[23] See http://uk-air.defra.gov.uk/assets/documents/reports/aqeg/ch2.pdf.

and heightened mortality rates. For example, Deschênes and Greenstone (2011) find that increased temperatures correlate with higher levels of infant mortality. Barreca (2008) finds similar evidence suggesting that both temperature and humidity can have adverse health effects. Failing to control for weather conditions can bias the estimated relationship between ambient pollution and mortality, as extreme pollution events are often strongly correlated with extreme weather events (Samet et al., 1998).[24]

IV. Empirical Methodology

Our model uses an infant week of life as the unit of observation, and the key parameter of interest is the effect of local traffic pollution on the hazard rate of death. We control for a rich set of geographic and time fixed effects, as well as (somewhat aggregated) individual-level controls.

A. *Mortality Hazards, LPMs, and Fixed Effects Models*

Our main specification is a discrete-time hazard, with the unit of observation being a person-week. The outcome of interest is whether said person died in a given week. Time since birth is the key hazard time element determining mortality risk. This closely follows the model used in CN. We control for the baseline hazard by including a flexible spline in age in weeks (with knots at 1, 2, 4, 8, 12, 20, and 32 weeks) and implementing a linear probability model (LPM).[25]

Given the large number of births that survive 52 weeks before leaving the sample, this gives a computationally taxing number of observations. Extensive controls and fixed effects compound this problem. We adopt a simplification that enables us to use information from all observations by collapsing birth data into cells prior to expanding into the person–week frame.[26] We first collapse all observations to mother postal-code-by-birth week-by-total weeks survived cells. For example, one collapsed cell would be all births in postal code z born in week w that lived 52 weeks. We calculate the cell mean for all mother–child covariates, and then expand observations to the cell–week of life level. In all regressions, we use weighting to approximate the uncollapsed model. This loses little variation, as pollution, traffic and weather are all common at the mother postal-code-by-week level, and greatly reduces the computational burden for estimation. In our preferred estimates, the number of observations decreases from over 75 million to approximately 9 million, covering just over 1.4 million births in 719 postal codes.[27]

We include geographic fixed effects (at the postal code level) and flexible time effects allowing each month in time a different baseline impact (e.g., January 2004 is allowed to vary from January 2005). More specifically, in our preferred specification, we include postal-code-by-month of year fixed effects to flexibly control for monthly shifts within each postal code (e.g., seasonal effects), as well as general month-by-year fixed effects to control for state-level changes over time.[28] Given use of the discrete-time hazard model, there are multiple possible definitions of both month and year. The postal code–specific time fixed effect could refer to the time of birth, which is fixed across event weeks, or could refer to time of observation, which allows it to vary across event weeks. Our specification uses the month and year of the event week to generate the fixed effects. This best fits our first stage, where such fixed effects help better identify the effects of weekly traffic variation on both pollution and mortality. However, to allow for seasonality in birth outcomes, we also control for quarter of birth. In all regressions, we include rich controls for weather (cubic functions of maximum temperature, rainfall, humidity, and wind speed, and linear counts of days with rain and days with fog), as well as individual-level controls (collapsed to cell level means as described above) for child's sex, indicator variables for low birthweight and premature birth, and maternal age, education, and race, and public insurance status for delivery. To control for the possible neonatal impacts of mother pollution exposure, we include average trimester pollution exposure.[29] Note that we do not attempt to instrument for prenatal pollution exposure levels.

Our baseline OLS equation is

$$\begin{aligned} Mort_{c,z,a,m,y,w} = {} & \alpha_{z,m,y} + \beta Pollution_{z,w} + \phi Trimester_c \\ & + \delta X_c + \gamma Z_{z,w} + \phi_{m,y} + spline_a \\ & + \varepsilon_{c,z,a,m,y,w}, \end{aligned} \tag{3}$$

where c indicates collapsed child-cell, z is postal code, a is age in weeks, m is month (January–December), y is year, and w is the current week (running from 1 to 260 in our sample, representing weeks since December 31, 2001). $\alpha_{z,m}$ is the postal-code-by-month of year fixed effect, $\phi_{m,y}$ are general month-by-year effects, X_c are individual-level controls (which do not vary by week of life), and $Z_{z,w}$ are postal code–week level weather controls. *Trimester* is a vector of average pollution levels for the first, second, and third trimesters of gestation individually. Although we present this as if there

[24] Knittel et al. (2011) investigates the importance of weather controls of higher orders.

[25] We prefer the LPM to a logit or probit model as it aids with computational implementation (caused by a large number of time and region fixed effects), as well as eases implementation of the instrumental variables specification.

[26] The case-control methodology outlined in CN yields qualitatively similar results. Our preferred method uses all of the data and avoids a problem with case control estimates, which can be sensitive to changes in the size of the control sample chosen.

[27] We obtain similar results using the full individual-level data.

[28] Collapsing data to the week level means month fixed effects are more complicated—some weeks span multiple calendar months. To simplify, we substitute four-week period effects for month effects, resulting in thirteen "months" per year that each contain exactly four weeks.

[29] Trimester pollution exposure is approximated by averaging postal code–level pollution in weeks 1 to 12 before birth, 13 to 24 before birth, and 25 to 36 before birth for trimesters 1, 2, and 3, respectively.

were just one type of pollution, we allow both types to enter simultaneously in our larger models.

B. *Potential Confounders to OLS and the Use of Instrumental Variables*

Our use of an IV helps with a number of identification concerns in the fixed effects model. First, mothers may self-select into geographic regions, and if mothers with higher values for clean air choose to live in cleaner areas and are wealthier or have access to better health care, this will bias OLS estimates upward (Currie, 2011). Second, changes in local economic activity may correlate with both pollution and infant health. Regional growth will increase pollution levels but also correlate with increases in income levels or health care access. This would bias OLS estimates downward. Third, pollution assignment leads to potential bias in the form of measurement error. The majority of papers in the air pollution and health literature, including this one, assign pollution levels to a particular person, living in a particular geographic area (e.g., postal code or county), based on pollution readings from pollution sensors in or near this geographic area. The researcher may not know the person's exact residence (two recent exceptions to this limitation are Currie et al., 2009, and Currie & Walker, 2011), and it is unlikely that the person is stationary over the time period analyzed. In addition, unless one knows the exact model of spatial dispersion of the pollutant in specific conditions, even if the person lived in the assigned location and never moved from this space, one would measure individual pollution with error. Insofar as this measurement error is "classical," OLS estimates will be biased toward 0. If the measurement error correlates with pollution levels, bias may be in either direction. In section VF, we formally consider the role of nonclassical measurement error in our IV estimates. Finally, individuals engage in avoidance behavior when confronted with environmental bads (Neidell, 2009; Moretti & Neidell, 2011; Schlenker & Walker, 2014; Graff Zivin & Neidell, 2009; Graff Zivin, Neidell, & Schlenker, 2011). Such avoidance behavior can mute the estimated true effect of pollution on health.

Shocks to traffic and interactions with weather conditions provide instruments for reducing all such sources of bias. If individuals sort based on average levels of pollution and traffic but not on traffic and weather shocks, our instrument strategy satisfies the exclusion restriction; similarly, weekly variation in such shocks (after conditioning on geographic and time fixed effects) is likely uncorrelated with economic growth.[30] Conditional on a valid instrument, an IV approach helps with measurement error. Assuming individuals do not systematically modify their actions based on random and potentially unobservable traffic shocks, IV estimates help alleviate the potential bias of avoidance behavior.

The key exclusion restriction for our traffic instrument is that (week-to-week) fluctuations in traffic and its interaction with local weather conditions do not directly affect infant mortality through vectors other than pollution. Since IV models continue to control for the fixed effects of the OLS specification, we believe this is a plausible assumption. An additional concern related to the exclusion restriction comes with our use of weather. Stormy weather, for example, can slow down traffic and also directly affects mortality and ambient pollution (see section III). For this reason, we include all weather variables we interact with traffic independently in all regressions.

Our primary instrument is postal code–level car-miles flow interacted with each of our weather variables, based on chemical interactions between automobile emissions and weather discussed in section III. Specifically, we interact our traffic measure with all the weather variables within the model. This captures, for example, the fact that emissions are less likely to stay concentrated in the atmosphere when there is strong wind or rain. We construct estimated standard errors allowing for clustering at the postal code level.

V. Results

In all regressions, the term *observations* refers to the number of expanded hazard weeks, using the weighted model section IV describes. We list the number of births used in each case in relevant table notes.

A. *Traffic, Mortality, and Pollution by Distance*

We begin with an investigation of the reduced-form relationship between traffic and weekly infant mortality:

$$\begin{aligned} Mort_{c,z,a,m,y,w} = {} & \alpha_{z,m,y} + \beta carmiles_{z,w} + \Sigma_{i=1}^{14} \pi_i carmiles_{z,w} \\ & \times weather_{i,z,m,y,w} + \phi Trimester_c + \delta X_c \\ & + \gamma Z_{z,w} + \phi_{m,y} + spline_a + \varepsilon_{c,z,a,m,y,w}. \end{aligned} \tag{4}$$

Here $\Sigma_{i=1}^{14} \pi_i carmiles_{z,w} \cdot weather_{i,z,m,y,w}$ is a series of fourteen total weather interactions with traffic: cubics for temperature, humidity, rainfall, and wind speed and linear functions for days with rain and days with fog. The use of interactions between traffic levels and ambient weather conditions complicates the interpretation of any one traffic coefficient within the model. Instead, we present the reduced form by calculating the derivative of weekly mortality with respect to traffic across the distribution of all interacted weather variables, which provides a point estimate and standard error for the marginal effect of traffic.[31]

Column 1 of table 4 shows the derivative of mortality with respect to car-miles, in 10 millions. We split traffic exposure

[30] Economic growth may lead to additional traffic shocks. For example, economic development may increase the number of cars on the road at any given time, thus increasing the probability of an accident. To some degree, time fixed effects capture this variation.

[31] We accomplish this using the Stata "margins" command after estimating our regressions.

TABLE 4.—EFFECT OF TRAFFIC ON POLLUTION AND MORTALITY BY DISTANCE FROM POSTAL CODE CENTROID

	(1) Mortality	(2) CO	(3) PM10
Cars within 0–5 miles (10 million)	0.0484**	0.0199	2.7958***
	(0.0240)	(0.0332)	(0.7753)
$\beta \times \frac{\text{Std. Dev Traffic}}{\text{Std. Dev Outcome}}$	0.006	0.032	0.170
Cars within 5–10 miles (10 million)	−0.0104	0.0164	1.1785**
	(0.0164)	(0.0199)	(0.4602)
$\beta \times \frac{\text{Std. Dev Traffic}}{\text{Std. Dev Outcome}}$	−0.003	0.051	0.140
Cars within 10–15 miles (10 million)	−0.0014	0.0183	1.6105***
	(0.0126)	(0.0188)	(0.4113)
$\beta \times \frac{\text{Std. Dev Traffic}}{\text{Std. Dev Outcome}}$	0.000	0.067	0.225
Cars within 15–20 miles (10 million)	0.0013	0.0067	0.4076**
	(0.0049)	(0.0080)	(0.2032)
$\beta \times \frac{\text{Std. Dev Traffic}}{\text{Std. Dev Outcome}}$	0.000	0.025	0.058

We use a starting sample of 1,434,613 births covering 719 postal codes, expanded to 75,777,503 observations in a discrete-time hazard model as described in section IV. All regressions control for month-by-year fixed effects, quarter-of-birth fixed effects, and mother postal-code-by-month fixed effects, as well as all weather and birth controls described in section II. Traffic variables show the calculated marginal effect of an additional 10 million weekly cars (adjusted for length of road coverage as we discuss in section IV) within each distance region from a postal code centroid, with no distance weighting. Regressions include interactions between each traffic measure and all weather variables, as we describe in section IV. We calculate marginal effects postregression and show the mean marginal effect across the entire distribution of weather variables with the corresponding calculated standard error. Column 1 shows the effect of each traffic flow measure on weekly hazard rate probability of death, with coefficients multiplied by 1,000 for ease of reading. Column 2 shows the effect of each traffic flow measure on ambient weekly CO levels in parts per million. Column 3 shows the effect of each traffic flow measure on ambient weekly PM10 levels in micrograms per cubic meter. We cluster standard errors at the mother–postal code level.

into 5-mile distance gradients to explore how effects vary by geographic proximity: 0 to 5 miles, 5 to 10 miles, 10 to 15 miles, and 15 to 20 miles. The marginal estimate is the raw effect of an additional 10 million car-miles per week within each distance bin. We find a statistically significant relationship between traffic levels and mortality, but only for traffic within 5 miles of a postal code centroid. For traffic close by, an additional within–postal code standard deviation in car-miles increases the probability of weekly mortality by 0.2% of the mortality rate standard deviation. For all distances beyond 5 miles, the marginal effects are statistically and economically 0.

Columns 2 and 3 of table 4 follow the design of reduced-form effects in column 1 but use ambient pollution as the outcome of interest. This is the analog of our first stage by individual pollutants and allows us to best construct an appropriate distance cutoff and weighting metric for our first stage in the IV analysis, along with developing a better understanding of how localized traffic must be to influence ambient pollution levels. Column 2 suggests some distance trends for CO, where effects for shorter distances are economically significant, though standard errors are large enough that we fail to reject 0 effects for all distance groups. An additional within–postal code standard deviation in car-miles within 5 miles increases ambient CO levels by 0.15% of a pollution standard deviation, though the effect is noisy at this distance and we cannot reject substantially higher values. A standard deviation increase in car-miles within 5 to 10 miles has a positive effect of 0.84% of a pollution standard deviation, followed by a smaller effect of 0.535 10 to 15 miles away. By 15 to 20 miles, the effect is closer to 0.

Column 3 shows more persistent effects for PM10. An additional standard deviation increase in car-miles within 5 miles increases ambient PM10 levels by 0.9% of a standard deviation. The effect is about half that size for cars within 5 to 10 miles, where it holds for 10 to 15 miles as well. By 15 to 20 miles, the link is no longer statistically significant at conventional levels but remains positive.

Based on our findings for effects by distance, we focus our IV analysis on traffic within 15 miles of a postal code centroid (and thus on postal codes with at least one traffic monitor within 15 miles). To simplify analysis, we sum all traffic within 15 miles rather than using the nonparametric methods discussed above. Both PM10 and CO exhibit correlations that suggest the marginal effect of traffic has some level of decay with distance. As our primary metric, we weight traffic by $1 - \sqrt{distance/15}$.[32] This function places higher weight on cars close to the postal code, with a relatively rapid decrease, followed by close to constant weights for distances farther out. For example, we assign a car directly on the postal code centroid a value of 1 (e.g., $1 - \sqrt{0}$), a car 5 miles away a value of approximately 0.4 (e.g., $1 - \sqrt{5}$), and a car 15 miles away a value of approximately 0.25. Column 5 of table 2 shows the average and standard deviations for our weighted traffic measure.

B. *Reduced-Form Effects of Traffic on Mortality by Weather Conditions and Subgroups*

We expand on our reduced-form analysis using two measures of traffic. Based on our finding that reduced-form effects are largely local, we first consider differences by weather conditions and different subgroup populations for postal codes within 5 miles of a traffic sensor. We then repeat the same process but expand out to our 15-mile cutoff and use the single distance-weighted measure of traffic.

Panel A of table 5 shows results for the restricted sample of postal codes within 5 miles of at least one sensor and how they vary by weather conditions. Results from column 1 suggest a standard deviation increase in weighted car-miles increases mortality by approximately 0.2% of a standard deviation. Figure 1 shows the distribution of marginal effects across all weather conditions, with a dashed line indicating the mean, regression-estimated/effects. For effectively the entire mass of the data, there is a positive correlation between higher traffic levels and increased infant mortality. To calculate overall mortality effects, we translate our marginal effects by multiplying the estimated impact on the hazard rate by 52 to gather the full exposure probability in the first year of life. That is, if the additional hazard in any given week (after controlling for age effects and all other covariates) is β, the total additional hazard for an infant who lives 52 weeks is $52 \times \beta$. This gives the marginal effect on the probability of death in the first year of life.

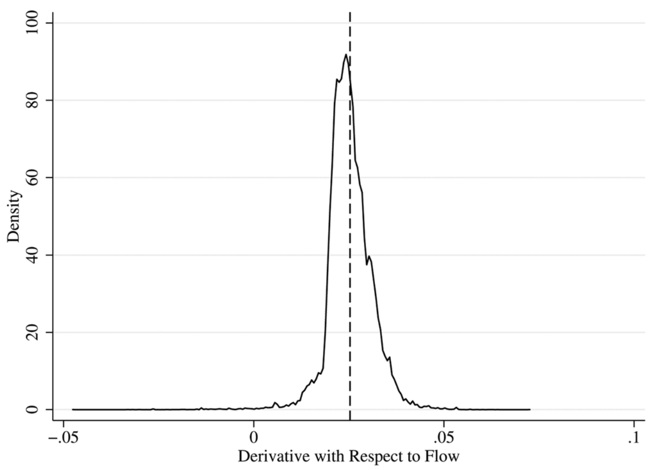
[32] Alternate weighting methods, including $1 - (distance/15)$ and $\frac{1}{1+distance}$ yield similar overall results.

TABLE 5.—VARIATION OF THE REDUCED-FORM IMPACT OF TRAFFIC ON WEEKLY MORTALITY HAZARD BY WEATHER CONDITIONS

	(1)	(2) Temperature	(3)	(4) Humidity	(5)	(6) Wind Speed	(7)	(8) Rainfall	(9)	(10) Foggy Days	(11)
	Overall	25th	75th	25th	75th	25th	75th	25th	75th	0	3
Unweighted weekly traffic within 5 miles (10 million cars) on weekly mortality hazard											
Car-miles	0.0423**	0.0434**	0.0394*	0.0399**	0.0449**	0.0427**	0.0394*	0.0435**	0.0428**	0.0418**	0.0426**
	(0.0201)	(0.0198)	(0.0203)	(0.0203)	(0.0201)	(0.0202)	(0.0201)	(0.0201)	(0.0201)	(0.0202)	(0.0201)
p-value of difference		0.297		0.191		0.381		0.551		0.686	
Distance weighted weekly traffic within 15 miles (10 million cars) on weekly mortality hazard											
Car-miles	0.0253	0.0268	0.0217	0.0259	0.0258	0.0254	0.0242	0.0268	0.0257	0.0239	0.0262
	(0.0192)	(0.0189)	(0.0193)	(0.0193)	(0.0192)	(0.0191)	(0.0193)	(0.0192)	(0.0192)	(0.0191)	(0.0192)
p-value of difference		0.080		0.976		0.718		0.193		0.118	

We restrict analysis to postal codes with at least one traffic sensor within 5 miles (panel A) or 15 miles (panel B), and expand to a discrete-time hazard model as described in section IV. This results in 66,870,281 observations covering 594 postal codes and 1,265,678 births for the 5 mile model and 73,109,698 observations covering 684 postal codes and 1,383,941 births for the 15 mile model. All regressions control for month-by-year fixed effects, quarter-of-birth fixed effects, and mother postal-code-by-month fixed effects, as well as all weather and birth controls described in section II. Outcome is the effect of each traffic flow measure on weekly hazard rate probability of death, with coefficients multiplied by 1,000 for ease of reading. Traffic variables show the calculated marginal effect of an additional 10 million weekly cars (adjusted for length of road coverage as we discuss in section IV) within each distance region from a postal code centroid. Panel A uses no distance weighting. Panel B weights by $1 - \sqrt{\frac{distance}{15}}$ as we describe in section V. Regressions include interactions between each traffic measure and all weather variables, as we describe in section IV. We calculate marginal effects postregression and show the mean marginal effect across the entire distribution of weather variables with the corresponding calculated standard error. Each column shows our evaluated marginal effect over either all data (column 1) or at specific values of interacted weather conditions (columns 2–11). We cluster standard errors at the mother–postal code level.

FIGURE 1.—DENSITY OF REDUCED-FORM IMPACTS: WEIGHTED CAR-MILES WITHIN 15 MILES



Kernel density of reduced-form effect of traffic on infant mortality, based on a model from column 1 of table 5. Each observation in the reduced-form impact of traffic on mortality, based on the particular conditions for that observation. The histogram is the distribution of these impacts. "Car-miles" refers to our unit of traffic intensity—the sum of weekly car-mile counts on each sensor loop times the length of road that loop represents (in 10 millions of car-miles). We weight traffic giving higher weight to more local traffic, where $weight = 1 - \sqrt{\frac{distance}{15}}$.

Columns 2 to 9 show how this effect varies by weather conditions, where we evaluate the marginal effect at the 25th and 75th percentiles of maximum temperature, humidity, wind speed, and rainfall. We also consider the effect by weeks with 0 foggy days versus 3 foggy days, which corresponds approximately to the 25th and 75th percentiles of weekly fog rates. While we find suggestive evidence that certain conditions are more prone to result in higher mortality rates, none result in effects that are statistically different from each other at conventional levels. In general, it appears that colder weeks and more humid weeks result in larger relationships between traffic and mortality, though we note here that the 25th percentile of maximum weekly temperature in California is approximately 67 degrees, not situations of extreme cold.

Panel B repeats this process expanded out to farther postal codes and using our weighted distance metric of traffic within 15 miles. Introducing larger distances makes all reduced-form estimates statistically insignificant at conventional levels, though the average effect suggests an additional standard deviation in distance-weighted cars correlates with a 0.16% of a standard deviation increase in mortality. At the larger distances, the relationship between traffic and weather follows a similar pattern as our earlier local effects, though we now find suggestive differences by rainfall.

We next check impacts of traffic across subgroups. To do so, we collapse data by subgroup cells and interact indicators for each group with all weather and traffic variables. This allows us to estimate the effect of traffic for each subgroup, while additional interactions with all weather variables help estimate differences in responses to traffic rather than interacted traffic–weather. We consider effects for several groups traditionally sensitive to health shocks: African Americans, Hispanics, births covered by Medicaid, births to non–high school graduates, premature births, and children of low birthweight. As with results by weather, we show results restricted to postal codes within 5 miles of a traffic sensor (panel A) and for distance-weighted traffic within 15 miles of a postal code (panel B).

Table 6 shows results by subgroup. We find no consistent detectable differences by race, educational status, or births covered by Medicaid, though for the majority of subgroups, differences are in the anticipated direction: the negative effects of pollution are larger for Hispanics, lower-education mothers, and mothers on Medicaid. Only the estimate for African Americans has an unexpected direction: the marginal effect of traffic is lower for the traditionally more sensitive group. We note, however, that African Americans represent a very small portion of births in California, and the two estimates are essentially the same. The most drastic differences occur for premature infants and infants of low birthweight. The mortality effect is almost entirely localized in these more sensitive infants, and the effect for premature births is the only result statistically significant at conventional levels for the distance of 15 miles. The direct effects of traffic are much larger in more health-sensitive

TABLE 6.—VARIATION OF THE EFFECT OF TRAFFIC ON WEEKLY MORTALITY BY SUBGROUPS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
		Premature		Low Weight		Hispanic		African American		High School Gradsuctes		Medicaid	
	Overall	0	1	0	1	0	1	0	1	0	1	0	1
A. Unweighted weekly traffic within 5 miles (10 million cars) on weekly mortality hazard													
Car-miles	0.0423**	0.0343*	0.0709***	0.0347*	0.0630***	0.0371*	0.0401**	0.0396**	0.0390**	0.0426**	0.0407**	0.0403**	0.0388**
	(0.0201)	(0.0189)	(0.0256)	(0.0188)	(0.0237)	(0.0193)	(0.0193)	(0.0194)	(0.0197)	(0.0197)	(0.0193)	(0.0193)	(0.0194)
p-value of difference		0.025		0.029		0.387		0.924		0.608		0.603	
B. Distance weighted weekly traffic within 15 miles (10 million cars) on weekly mortality hazard													
Car-miles	0.0253	0.0184	0.0485**	0.0206	0.0347	0.0182	0.0247	0.0236	0.0187	0.0264	0.0219	0.0218	0.0221
	(0.0192)	(0.0185)	(0.0247)	(0.0185)	(0.0225)	(0.0187)	(0.0186)	(0.0188)	(0.0192)	(0.0190)	(0.0186)	(0.0187)	(0.0188)
p-value of difference		0.058		0.2685		0.060		0.405		0.185		0.920	

We restrict analysis to postal codes with at least one traffic sensor within 5 miles (panel A) or 15 miles (panel B) and expand to a discrete-time hazard model as described in section IV. This results in 66,870,281 observations covering 594 postal codes and 1,265,678 births for the 5 mile model and 73,109,698 observations covering 684 postal codes and 1,383,941 births for the 15 mile model. All regressions control for month-by-year fixed effects, quarter-of-birth fixed effects, and mother postal-code-by-month fixed effects, as well as all weather and birth controls described in section II. Outcome is the effect of each traffic flow measure on weekly hazard rate probability of death, with coefficients multiplied by 1,000 for ease of reading. Traffic variables show the calculated marginal effect of an additional 10 million weekly cars (adjusted for length of road coverage as we discuss in section IV) within each distance region from a postal code centroid. Panel A uses no distance weighting. Panel B weights by $\sqrt{\frac{distance}{15}}$ as we describe in section V. Regressions include interactions between each traffic measure and all weather variables, as we describe in section IV. We calculate marginal effects postregression and show the mean marginal effect across the entire distribution of weather variables with the corresponding calculated standard error. Each column shows our evaluated marginal effect over either all data (column 1) or for specific birth groups (columns 2–11). We cluster standard errors at the mother–postal code level.

subgroups: 2.0 to 2.6 times larger for premature infants and 1.7 to 1.8 larger for infants of low birthweight.

C. *Variation in Pollution Effects by Weather*

Using our distance-weighted metric of traffic within 15 miles of a postal code, we next consider variations in our first stage under different weather conditions. Column 1 of Table 7 shows the marginal effect of an additional 10 million distance weighted car-miles. For both pollutants, we find an economically and statistically significant relationship between local traffic levels and ambient pollution on average. An additional standard deviation in car-miles increases ambient CO levels by 1% of a standard deviation, and PM by 3% percent of a standard deviation. Columns 2 through 11 of table 7 explore how the marginal effect of traffic on ambient pollution varies under different weather conditions. As with the reduced form, we evaluate the marginal effect of traffic at the 25th and 75th percentiles of the weather variable of interest. Given the high level of statistical precision in our estimates, almost all cross-weather comparisons are statistically different from each other. Interestingly, different weather conditions affect each pollutant differently. Columns 2 and 3 show that traffic on hotter weeks creates relatively more CO and relatively less PM10: higher temperatures in the summer can favor inversions, raising CO, while lower temperatures favor formation of certain atmospheric particles, raising PM10. Humidity shows the opposite effect: higher humidity means the same traffic conditions generate more PM10 and less CO. Effects for both pollutants are smaller on windier weeks and rainier weeks. Finally, while higher fog has little impact on PM10 creation, higher amounts of fog mean greater generation of CO. These varied effects are what allow us to separately identify both CO and PM10 in our upcoming IV regressions—the same traffic levels in varied weather conditions influence each pollutant differently.

As further evidence of the varied relationship of traffic, weather, and pollution, we next plot the estimated marginal effect of weighted traffic in different weather conditions for different pollutants. To do so, we calculate the estimated marginal effect of an additional 10 million car-miles (approximately 1 standard deviation) at every point in our data, and then plot the kernel density of estimated effects. Figure 2 shows the marginal effect distribution for CO across weather conditions for all continuous higher-order variables: temperature, humidity, wind speed, and rainfall. Figure 3 repeats the process for PM10. In figures 4 and 5, we plot two densities for each weather variable: one for all values at or below the 25th percentile of the relevant variable and one for all values at or above the 75th percentile.

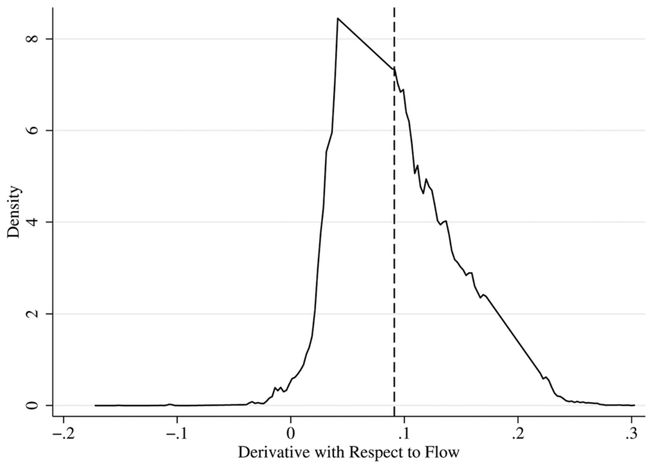
The densities largely align with our marginal estimates, with the exception of the impact of traffic on CO at different levels of humidity and PM10 at different temperatures. Our table finds substantially lower levels of CO on humid days, but graphs show that the densities almost entirely overlap.

TABLE 7.—VARIATION OF THE EFFECT OF TRAFFIC ON AMBIENT POLLUTION BY WEATHER CONDITIONS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
		Temperature		Humidity		Wind Speed		Rainfall		Foggy Days	
	Overall	25th	75th	25th	75th	25th	75th	25th	75th	0	3
A. Distance weighted weekly traffic within 15 miles (10 million cars) on carbon monoxide											
Car-miles	0.0911***	0.0486***	0.1196***	0.1170***	0.0859***	0.0845***	0.0867***	0.0925***	0.0909***	0.0793***	0.0988***
	(0.0114)	(0.0115)	(0.0113)	(0.0113)	(0.0119)	(0.0113)	(0.0120)	(0.0115)	(0.0113)	(0.0112)	(0.0115)
p-value of difference		0.000		0.000		0.570		0.003		0.000	
B. Distance weighted weekly traffic within 15 miles (10 million cars) on particulate matter											
Car-miles	6.5636***	6.8294***	6.4646***	5.5407***	7.6205***	6.6294***	6.5936***	6.6496***	6.5512***	6.5508***	6.5720***
	(0.4870)	(0.4837)	(0.4986)	(0.5093)	(0.4919)	(0.4761)	(0.5023)	(0.4899)	(0.4870)	(0.4793)	(0.4929)
p-value of difference		0.020		0.000		0.769		0.001		0.665	

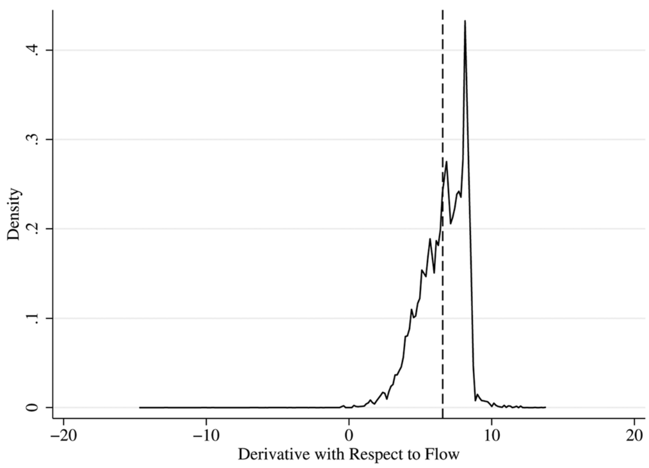
We restrict analysis to postal codes with at least one traffic sensor within 15 miles and expand to a discrete-time hazard model as described in Section IV. This results in 73,109,698 observations covering 684 postal codes and 1,383,941 births. All regressions control for month-by-year fixed effects, quarter-of-birth fixed effects, and mother postal-by-month fixed effects, as well as all weather and birth controls described in section II. Outcome is ambient carbon monoxide in parts per million (panel A) and ambient particulate matter in micrograms per cubic meter (panel B). Traffic variables show the calculated marginal effect of an additional 10 million weekly cars (adjusted for length of road coverage, as we discuss in section IV) within each distance region from a postal code centroid. We weight traffic values using $1-\sqrt{\frac{distance}{15}}$ as we describe in section V. Regressions include interactions between each traffic measure and all weather variables, as we describe in section IV. We calculate marginal effects postregression and show the mean marginal effect across the entire distribution of weather variables with the corresponding calculated standard error. Each column shows our evaluated marginal effect over either all data (column 1) or for specific subgroups (columns 2–5), which we evaluate by interacting all weather and traffic variables with indicators for each subgroup. Tests for differences across race, mother education, and Medicaid status found no statistically significant differences (not shown). We cluster standard errors at the mother–postal code level.

FIGURE 2.—FIRST STAGE FOR CARBON MONOXIDE



Kernel density of estimated impact of traffic on CO based on the first-stage model from table 7. Each observation in the first stage has a predicted impact of traffic on pollution based on the particular weather conditions for that observation. This histogram is the distribution of these impacts.

FIGURE 3.—FIRST STAGE FOR PARTICULATE MATTER



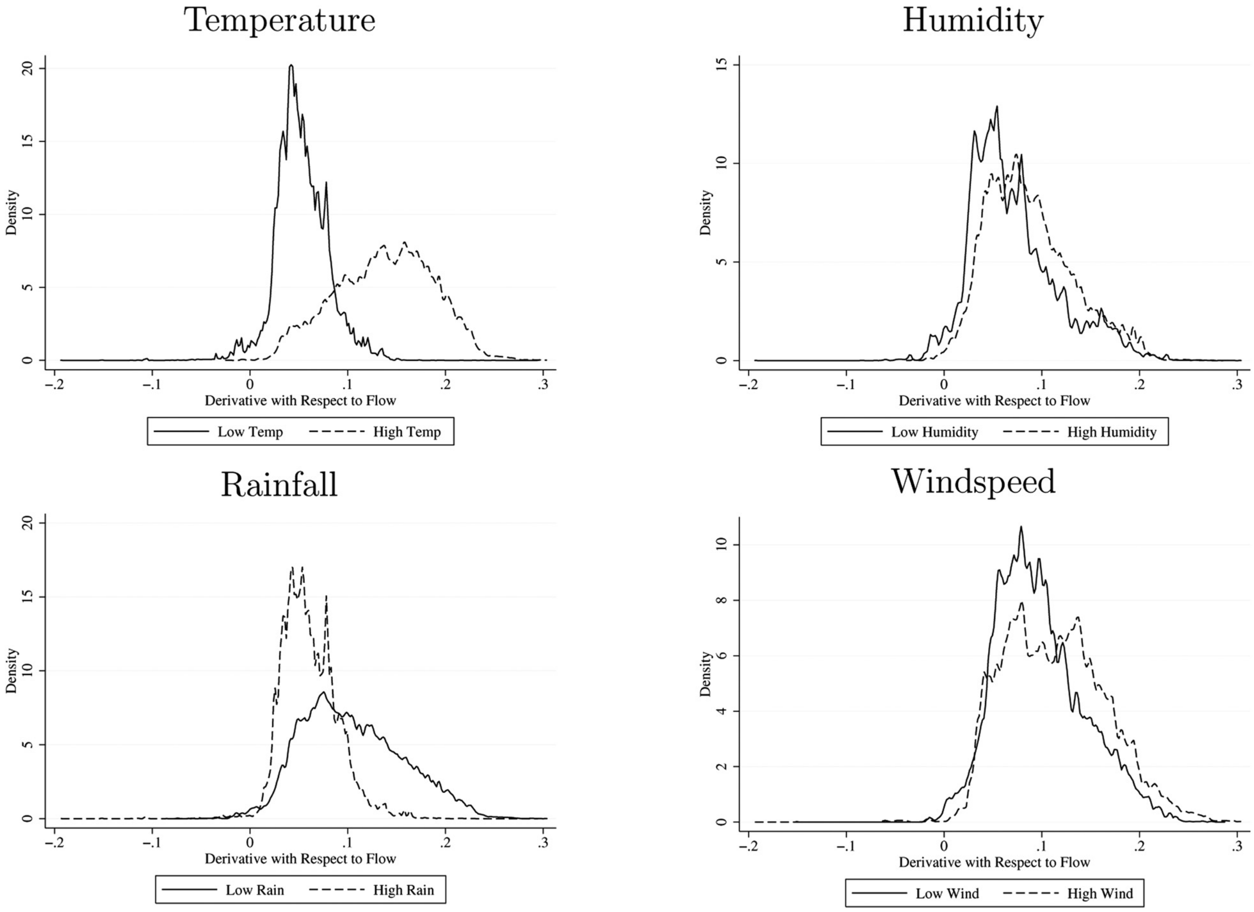
Kernel density of estimated impact of traffic on PM10, based on the first-stage model from table 7. Each observation in the first stage has a predicted impact of traffic on pollution based on the particular weather conditions for that observation. This histogram is the distribution of these impacts.

We also find PM10 is higher on colder days, but densities suggest higher effects in higher temperature. Note, however, that this exercise differs slightly from our analysis in table 7. The table evaluates the marginal effect of traffic holding the given weather variable fixed at the indicated level. The figures show the estimated marginal effect for all observations falling in the range of the indicated variable cutoffs. Thus, in the table case, we have the "all else held constant" result, while in the density case, we cannot observe the effect of any one weather variable in isolation. Comparing PM10 graphs across temperature and humidity suggests this might explain some differences: high-humidity days are also high-temperature days, making the separation of the two effects in our densities difficult.

D. *Pollution and Infant Mortality: OLS and IV Results*

We begin with OLS estimates similar to those of CN, but with a focus on CO and PM10 in the 2002–2007 time period

FIGURE 4.—THE VARIED IMPACT OF TRAFFIC ON POLLUTION BY WEATHER CONDITIONS: CO



Kernel densities of first-stage impacts based on subsets of data from figure 2. Distributions include observations with weather values above the 75th percentile (for high) and below the 25th percentile (for low) of all higher-order weather in the primary analysis. See sections II and VC.

used for our traffic analysis (for an extensive replication of the earlier CN results as well as in-depth discussion of the crosswalk from their model to our preferred estimates, see Knittel et al., 2011). The first column of table 8 includes fixed effects for the month-by-year of event time and mother's postal code-by-month fixed effects; all weather controls; a spline in the child's age; cell-level averages of indicators for child's sex, mother's age, race, and education; the cell-level variable for whether public insurance was used for the delivery; the cell-level average for being of low birthweight (below 2500 grams); and the cell-level average for being classified as premature (more than three weeks early). We multiply all coefficients by 1,000 for ease of reading.[33]

In the 2002–2007 period, CO levels are 40% below those from 1989 to 2000, and average levels of PM10 are 5% lower. Despite these decreases, our estimates are consistent with those from the earlier period, but with greater noise and no statistically significant effects. The point estimates imply conclusions similar to those found in CN–CO correlates positively with infant mortality. A 1-unit decrease in CO saves 13 infant lives, an estimate surprisingly close to the CN result given the substantially lower CO levels in our time period. Our finding for PM10 is effectively 0.

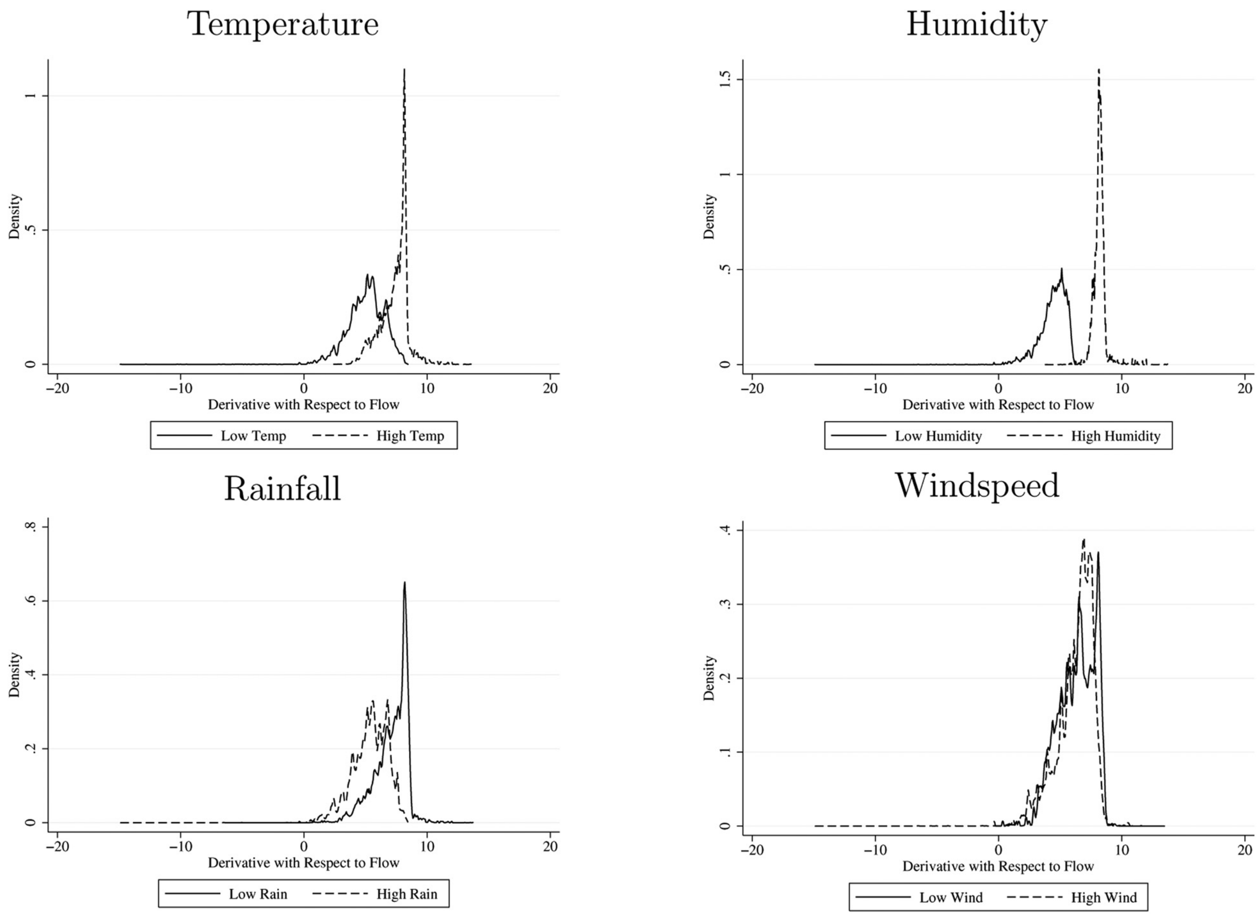
E. Instrumental Variables Estimation

Table 8 reports our main IV results. We include first-stage F statistics below reported coefficients (calculated using the Angrist-Pischke F-test for multiple instruments when relevant) along with calculated marginal effects.[34] While we find positive effects for both pollutants, only PM10 is statistically significant, and it remains so when including both pollutants simultaneously. In the joint pollutant model, a 1-unit decrease in CO results in 82 fewer deaths per 100,000 live births, while a within–postal code standard deviation decrease in CO correlates with 42.94 fewer deaths. We find a 1-unit decrease in PM10 means 10 fewer deaths per

[33] To calculate total mortality effects, we again translate the marginal effect of increased mortality in a given week into an increase in the probability of mortality in the first year of life. Multiplying this probability by 100,000 gives the approximate number of additional deaths similar to those in CN.

[34] We run all IV regressions using the user-generated Stata command xtivreg2 (Shaffer, 2010). As a robustness check, we further adjust our standard errors for two-way clustering across both location (mother postal code) and time (event month-by-year). This decreases our first-stage F-statistics, to around 10 for CO and 3 for PM10. However, the second-stage coefficient remains statistically significant at 10%.

FIGURE 5.—THE VARIED IMPACT OF TRAFFIC ON POLLUTION BY WEATHER CONDITIONS: PM10



Kernel densities of first-stage impacts based on subsets of data from figure 3. Distributions include observations with weather values above the 75th percentile (for high) and below the 25th percentile (for low) of all higher-order weather in the primary analysis. See sections II and VC.

100,000 live births, with a within-postal code standard deviation decrease correlating to 78 fewer deaths per 100,000 live births.

In considering the magnitude of these effects, it is helpful to refer to prior findings on pollution and infant mortality rates. Both CN and Currie et al. (2009) find an effect of approximately 17 avoided deaths (per 100,000 births) per unit of CO in the United States, while Arceo-Gomez et al. (forthcoming) find much larger effects around 166 per 100,000 live births in Mexico. Our results, while noisy, fall between these two. Our more precise PM10 estimates align quite well with prior findings. For example, Chay and Greenstone (2003b) find that a 1 unit drop in total suspended particulates (TSPs) resulted in a drop of 4 to 8 infant deaths per 100,000 live births, while Chay and Greenstone (2003a) found an effect of around 7 to 13 infant deaths per unit. Both used TSPs, which contain both PM10 and larger particulates not included in the PM10 specification. While no direct conversion metric exists, the the World Bank Group (1999) notes a commonly used conversion metric between the two measures is $PM10 = 0.55 \cdot TSP$. Using that conversion metric, the Chay and Greenstone (2003a) results suggest marginal impacts of 7 to 15 and 13 to 23 additional deaths per unit increase of PM10. While CN find no statistically significant effect for PM10, Arceo-Gomez et al. (forthcoming) find around 9 deaths per 100,000 live births per unit, though only in their single pollutant models.

F. *The Role of Weather Interactions and Nonclassical Measurement Error*

A benefit of using interactions between weather and traffic as instrumental variables is the ability to jointly identify the impacts of separate pollutants despite having only one measure of traffic. However, the use of multiple instruments raises the concern of the true source of identification. Are results a product of using enough instruments to get a statistically significant result or by the inclusion of a particular weather effect alone? Both of these issues are of concern. To address this, we repeat our main IV analysis but vary the weather interactions included in the first stage. Table 9 shows the results. Becasuse we begin with fewer than three instruments, we cannot estimate the simultaneous pollutant model, so we instead conduct all analysis in a single-pollutant framework for PM10 where we find statistically significant results in our main model.

TABLE 8.—OLS AND IV ESTIMATES OF POLLUTION ON INFANT MORTALITY

	(1) OLS	(2) IV	(3) IV	(4) IV
Carbon monoxide	0.0025	0.0237		0.0158
	(0.0058)	(0.0249)		(0.0262)
Particulate matter	−0.0001		0.0019**	0.0019**
	(0.0001)		(0.0009)	(0.0009)
First-stage F-statistic				
Carbon monoxide	—	57.02	—	60.11
Particulate matter	—	—	80.51	85.98
Deaths per unit				
Carbon monoxide	13.08	123.43	—	82
Particulate matter	−0.35	—	10	10.14
Deaths per within-postal-code SD				
Carbon monoxide	6.85	64.63	—	42.92
Particulate matter	−4.53	—	128.85	130.68
Deaths per between-postal-code SD				
Carbon monoxide	3.04	28.72	—	19.07
Particulate matter	−2.7	—	76.45	77.58

Regressions use a starting sample of 1,383,941 births, expanded to a discrete-time hazard model for 73,109,698 observations covering 684 postal codes. Column 1 presents OLS results. Columns 2 to 4 present IV results. The instrumental variables are car-miles and car-miles interacted with all included weather variables, as we describe in section V. Columns 2 and 3 include pollution variables individually, and column 4 includes them simultaneously. F-statistics test the hypothesis that the instruments have no predictive power in the first stage. F-statistics on multiple-pollutant regressions are Angrist-Pischke F-statistics for joint significance.

Column 1 shows results using only traffic as an instrument with no additional weather interactions. Column 2 adds an interaction with temperature, column 3 adds an additional interaction with humidity, and so on. The lower panel indicates which weather interactions are included for each column. By column 5, the regressions are equivalent to column 2 in table 8. Effects for PM10 are always positive. And while the estimates using only traffic as an instrument are not statistically significant, they are within a single standard error of the results in the most saturated model. Looking across all specifications, it does not appear that the addition of any single pollutant explains the size or magnitude of our results. Results become much more precisely estimated with the addition of humidity to the interaction set, suggesting there may be a strong link among ambient humidity, traffic, and pollution.

We next formally consider the potential bias caused by error in local assignment of pollution. A motivation for our IV approach is that estimating postal code–week level pollution may result in errors in measurement, which would bias standard fixed effects model estimates of the effects of pollution on health. IV removes the issue if the measurement error is classical. However, this may not hold with our measure of pollution. Errors may correlate with the actual value of the underlying pollution measure, and the variance of the measurement error might vary with distance from pollution monitors (Lleras-Muney, 2010). In this case, bias may remain in the IV.

To investigate the potential effects of nonclassical measurement error in our setting, we perform a Monte Carlo analysis where we model likely error in pollution measurement and examine how such error might alter our findings. We first examine how measurement error in pollution relates to the true pollution level and the distance from nearby pollution measurement stations. We generate an "error" at each actual pollution monitor by first estimating pollution at that monitor using other nearby monitors in a fashion identical to what we do for postal code centroids (using all other monitors within 20 miles, weighted inversely by distance). We then compare the estimated level with the correct level recorded at the monitor. We interpret the measurement at the monitor level as the "truth" and discrepancies in the estimated level from this as the "error." To compare how the expected level of this error (bias) and its variance depend on true pollution level and distance to closest monitor, we regress the error on location and month-by-year fixed effects and take the residuals of this as the object to be explained. We then estimate a linear regression model where the residual error is a function of a spline in true levels of pollution and distance to nearest monitor:[35]

$$\begin{aligned} &\textit{Residual Error} \\ &\quad = [\textit{Spline in Actual Pollution Measure}] \\ &\qquad + [\textit{Spline in Distance to Nearest Monitor}]. \end{aligned} \tag{5}$$

Examination of the spline in "true pollution" suggests the unexplained pollution error is small in magnitude over most of the support of true pollution levels but correlates with true pollution levels for larger-than-average levels of pollution. Examination of the spline in "distance to nearest monitor" suggests no relationship between the bias or variance in mis-measurement and the distance to the nearest monitor (similar results hold when controlling only for distance to closest monitor or only for observed pollution levels). As a whole, we find modest violation of the classical measurement error model along one metric: pollution measurement error correlates with true pollution level. This error is small over most of the range of the data. Panel A of figure 6 shows average true pollution levels (horizontal axis) versus unexplained error (vertical axis), with vertical lines marking the mean (dashed line) and mean plus 2 standard deviations (dotted line) in the data. Panel B shows a similar average by distance from nearest monitor. We conclude that measurement error in estimated pollution data deviates from the classical i.i.d. assumptions, raising the question whether IV estimates will still have the desirable properties of the classical case

We next perform a Monte Carlo simulation that tests for potential bias in the IV estimates. We first construct a data generation process based on the postal-code-by-week data in our sample. We generate the bias and variance of expected errors in pollution assignment from models based on splines in true pollution and distance to the closest monitor, and the lagged error as a predictor. We call this a predicted error. Treating our main estimates as the true parameters and our best estimates of pollution as true pollution levels, for each iteration of the simulation, we create pseudo-pollution

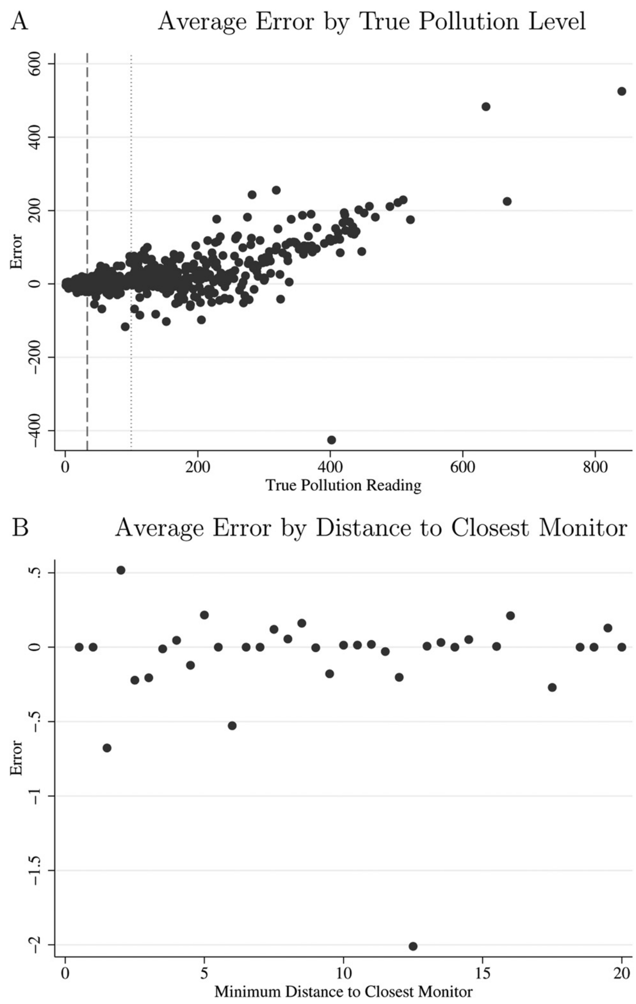
[35] We use linear splines with knot points at mile distances of 3, 6, 9, 12, and 15 miles and pollution points of 150, 300, 450, 600, and 750 micrograms per cubic meter of air.

TABLE 9.—IMPACT OF ADDING WEATHER INTERACTIONS AS ADDITIONAL INSTRUMENTS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PM10	0.003	0.0008	0.0017*	0.0018*	0.0020**	0.0019**	0.0019**
	(0.0022)	(0.0015)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
First-stage F-statistics	330.24	115.95	106.64	85.12	87.77	81.96	80.51
Max temperature		X	X	X	X	X	X
Humidity			X	X	X	X	X
Wind speed				X	X	X	X
Rainfall					X	X	X
Days with rain						X	X
Days with fog							X

Each row cell is a separate IV regression. Column 7 presents results corresponding to column 2 of table 8. Columns 1 to 6 present results based on more parsimonious instrument sets. Each column varies which weather variables we interact with our measure of traffic to create instrumental variables. All columns control for all weather variables.

FIGURE 6.—ERROR IN ASSIGNED POLLUTION BY TRUE POLLUTION LEVEL AND DISTANCE TO CLOSEST MONITOR



"Error" variable represents the difference between an observed pollution monitor level and the predicted level using our distance weighting estimation technique, after adjusting for mother postal code and month-by-year fixed effects. Dashed line indicates the mean of true pollution observations, with the dotted line representing 2 standard deviations above the mean. Panel A shows the mean error by size of observed pollution level. Panel B shows the mean error by distance to closest pollution monitor. See section VF for details.

measures. These pseudo-pollution measures are the true pollution levels plus the predicted error and an additional error component that allows for autocorrelation in errors across time within postal code. Specifically, the pseudo-pollution level in postal code z and period t would be

$$\begin{aligned}&\text{Pseudo-Pollution}_{z,t}\\&\quad = \text{True Pollution}_{z,t} + \text{Predicted Mean Error}_{z,t}\\&\qquad + \delta \times \text{predicted error}_{z,t-1} + \Phi(0,1)\\&\qquad \times \sqrt{(\text{predicted error variance} + \gamma \times \text{predicted error}^2_{z,t-1})},\end{aligned} \tag{6}$$

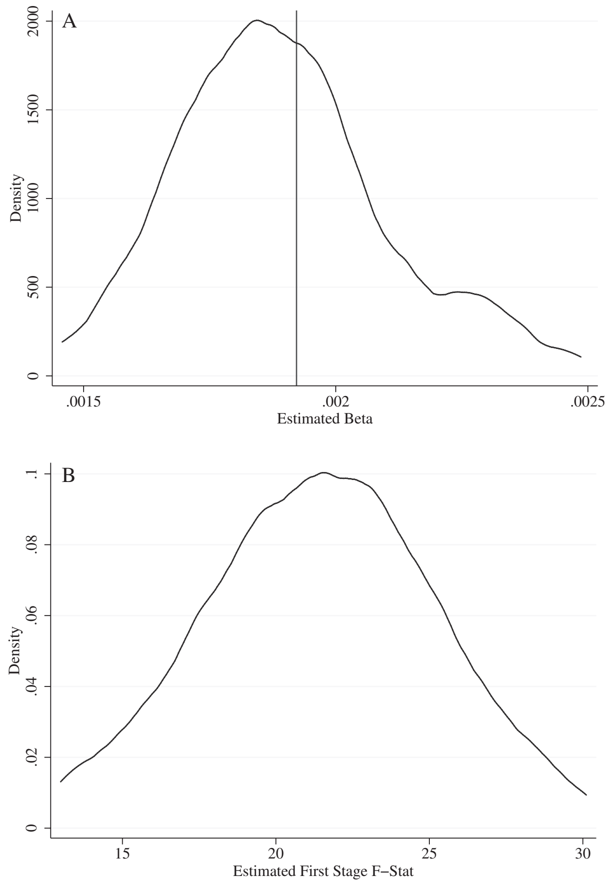
with $\Phi(0, 1)$ being a draw from the standard normal distribution. After generating the pseudo-pollution values, we then estimate our models using IV and save the resulting coefficients. We perform 100 simulations where we replace our "true" pollution with the pseudo-values and compare the distribution of estimated coefficients to our main estimate.

The simulation results indicate that measurement error may have a nonzero but modest effect on our estimates. The estimate for the impact of PM10 with no measurement error is 0.0019 (SE, 0.0009). In 100 Monte Carlo replications with measurement error, the mean estimated coefficient is 0.00186, with a standard deviation of 0.0002 and a range of 0.0014 to 0.0024. This suggests that pollution measurement error slightly increases the noise of our estimates (on the order of 23% of our estimated standard error), but that the bias it might induce is fairly small (about 1% of the true coefficient). Figure 7 shows the density of our simulated estimates, with the vertical line indicating our main results. This small bias appears to be in the direction of attenuation and so is unlikely to lead to spurious findings of effects.

VI. Conclusion

We analyze the impact of local traffic levels on mortality and ambient levels of carbon monoxide and particulate matter. We find a statistically and economically significant link between local traffic levels and infant mortality, where a standard deviation increase in traffic results in a 0.2% of a standard deviation increase in infant deaths. Effects are largest for premature and low birthweight infants. We also consider how effects vary by local weather conditions and find that the largest direct link between traffic and mortality occurs in colder weeks and weeks of greater humidity.

FIGURE 7.—MONTE CARLO RESULTS—DENSITY OF OBSERVED BETAS AND FIRST-STAGE F-STATISTICS



Kernel density shows distribution of results from 100 Monte Carlo simulations of the main effect in column 3 of table 8. Each simulation involves the addition of a random, nonclassical autocorrelated error to the estimated mother postal code–level pollution measure, as section VF describes. Panel A shows the distribution of observed second-stage estimates. Panel B shows the distribution of observed first-stage F-statistics, where we cluster standard errors at the mother postal code. The vertical line on panel A indicates our main result from column 3 of table 8.

Using the relationship between local traffic and regional weather conditions, we next build an instrumental variables model to better understand the direct impact of local pollution on infant mortality. We find suggestive evidence that carbon monoxide contributes to infant deaths even at today's lower levels, though results are noisy and not statistically significant. We also find PM10 has a large and statistically significant effect on infant mortality. In our preferred specification, a 1-unit decrease in PM10 (around 8% of a standard deviation) saves roughly 10 lives per 100,000 births. This represents a decrease in the weekly mortality rate of around 4%. This is consistent with the findings of prior research on ambient particulate matter and suggests that even at today's lower levels, reducing both ambient pollution and traffic congestion has substantial opportunity for health gains. Finally, we explore the role of nonclassical measurement error in our IV estimates using a Monte Carlo simulation where we model potentially autocorrelated error in pollution assignment and show that our main findings are robust to such concerns.

REFERENCES

Almond, Douglas, Lena Edlund, and Mårten Palme, "Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden," *Quarterly Journal of Economics* 124 (2009), 1729–1772.

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

Barreca, Alan, "Climate Change, Humidity, and Mortality in the United States," *Journal of Environmental Economics and Management* 63:1 (2008), 19–34.

Bishop, G. A., and D. H. Stedman, "Measuring the Emissions of Passing Cars," *Accounts of Chemical Research* 29 (1996), 489–495.

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

Chay, Kenneth Y., and Michael Greenstone, "Air Quality, Infant Mortality, and the Clean Air Act of 1970," NBER working paper 10053 (2003a).

——— "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *Quarterly Journal of Economics* 118 (2003b), 1121–1167.

Currie, Janet, "Inequality at Birth: Some Causes and Consequences," *American Economic Review* 101:3 (2011), 1–22.

Currie, Janet, Eric A. Hanushek, E. Megan Kahn, Matthew Neidell, and Steven G. Rivkin, "Does Pollution Increase School Absences?" this REVIEW 91 (2009), 682–694.

Currie, Janet, and Matthew Neidell, "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *Quarterly Journal of Economics* 120 (2005), 1003–1030.

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

Currie, Janet, and Johannes F. Schmieder, "Fetal Exposures to Toxic Releases and Infant Health," *American Economic Review* 99 (2009), 177–183.

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

Davis, Stacy C., and Susan W. Diegel, "Transportation Energy Data Book: Edition 26-2007," ORNL-6978 (Oak Ridge, TN: Department of Energy, 2007).

Deschênes, Olivier, and Michael Greenstone, "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather," *American Economic Review* 97 (2007), 354–385.

——— "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US," *American Economic Journal: Applied Economics* 3 (2011), 152–185.

Deschênes, Olivier, Michael Greenstone, and Jonathan Guryan, "Climate Change and Birth Weight," *American Economic Review* 99 (2009), 211–217.

Graff Zivin, J., and M. Neidell, "Days of Haze: Environmental Information Disclosure and Intertemporal Avoidance Behavior," *Journal of Environmental Economics and Management* 58 (2009), 119–128.

Graff Zivin, Joshua, Matthew Neidell, and Wolfram Schlenker, "Water Quality Violations and Avoidance Behavior: Evidence from Bottled Water Consumption," *American Economic Review* 101 (2011), 448–453.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders, "Caution, Drivers! Children Present. Traffic, Pollution, and Infant Health," NBER working paper 17222 (2011).

Lee, Sung B., Anthony Cocco, Darioush Keyvani, and G. Jordan Maclay, "Humidity Dependence of Carbon Monoxide Oxidation Rate in a Nafion-Based Electrochemical Cell," *Journal of the Electrochemical Society* 142 (1995), 157.

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

Moretti, Enrico, and Matthew Neidell, "Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles," *Journal of Human Resources* 46:1 (2011), 154–175.

Neidell, Matthew J., "Air Pollution, Health, and Socio-Economic Status: The Effect of Outdoor Air Quality on Childhood Asthma," *Journal of Health Economics* 23 (2004): 1209–1236.

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

Reyes, Jessica Wolpaw, "Environmental Policy as Social Policy? The Impact of Childhood Lead Exposure on Crime," *BE Journal of Economic Analysis and Policy* 7 (2007), 51.

Samet, Jonathan, Scott Zeger, J. Kelsall, Jing Xu, and Laurence Kalkstein, "Does Weather Confound or Modify the Association of Particulate Air Pollution with Mortality? An Analysis of the Philadelphia Data, 1973–1980," *Environmental Research* 77 (1998), 9–19.

Sanders, Nicholas J., "What Doesn't Kill You Makes You Weaker: Prenatal Pollution Exposure and Educational Outcomes," *Journal of Human Resources* 47 (2012), 826–850.

Sanders, Nicholas J., and Charles Stoecker, "Where Have All the Young Men Gone? Using Sex Ratios to Measure the Effect of Pollution on Fetal Death Rates" *Journal of Health Economics* 41 (2015), 30–45.

Schlenker, Wolfram, and W. Reed Walker, "Airports, Air Pollution, and Contemporaneous Health," *Review of Economic Studies* (forthcoming).

Shaffer, Mark E., "xtivreg2: Stata Module to Perform Extended IV/2SLS, GMM and AC/HAC, LIML and K-Class Regression for Panel Data Models" (2010), https://ideas.repec.org/c/bocode/s456501.html.

Shukla, J. B., A. K. Misra, S. Sundar, and R. Naresh, "Effect of Rain on Removal of a Gaseous Pollutant and Two Different Particulate Matters from the Atmosphere of a City," *Mathematical and Computer Modelling* 48 (2008), 832–844.

Stoecker, Charles, "Chill Out, Mom: The Long Run Impact of Cold Induced Maternal Stress in Utero," Tufts University School of Medical Health and Tropical Medicine mimeograph (2010).

Tiao, George C., and Steven C. Hillmer, "Statistical Models for Ambient Concentrations of Carbon Monoxide, Lead, and Sulfate Based on the LACS Data," *Environmental Science and Technology* 12 (1978), 820–828.

World Bank Group, *Pollution Prevention and Abatement Handbook* (International Bank for Reconstruction and Development, 1999).