CESifo Working Papers
www.cesifo.org/wp

What Forces Dictate the Design of Pollution Monitoring Networks?

Nicholas Z. Muller
Paul Ruud

CESIFO WORKING PAPER NO. 5739
CATEGORY 10: ENERGY AND CLIMATE ECONOMICS
FEBRUARY 2016

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 5739*

What Forces Dictate the Design of Pollution Monitoring Networks?

Abstract

The U.S. Environmental Protection Agency (USEPA) maintains networks of pollution monitors for two basic purposes: to check and enforce the attainment of national ambient air quality standards (NAAQS) and to provide useful data for studying pollution and its effects. These purposes imply conflicting criteria for the locations of a limited number of monitors. To check the attainment of standards, monitors are placed where pollution levels are highest. Monitors are not required where standards have always been met and there are no new pollution sources. To provide useful data for studying pollution and its effects, monitors are placed to observe outcomes under a variety of pollution levels. This study asks the following questions. What factors affect when a monitor is retired from the network? What drives the decision to add a new site? What causes year-to-year changes in the number of monitors? We tackle these questions with a particular focus on the role of regulatory compliance and pollution levels in the context of monitors for tropospheric ozone (O3). Using a panel dataset of monitors in the contiguous US spanning the years 1993 to 2011, we find that peak O3 readings in the prior period are significantly associated with the regulator's decision of whether to add or to drop a monitor in the following period. While compliance with the NAAQS for O3 is not consistently associated with network composition, compliance with the PM2.5 NAAQS does appear to affect changes to the network.

JEL-codes: Q530, Q580, C230, C250.

Keywords: air pollution, regulatory compliance, pollution measurement.

Nicholas Z. Muller
Economics
Middlebury College
Warner Hall 305E
USA – Middlebury, VT 05753
nmuller@middlebury.edu

Paul Ruud
Department of Economics
Vassar College
124 Raymond Avenue #631
USA - 12604-0631 Poughkeepsie NY
paruud@vassar.edu

1. Introduction

The U.S. Environmental Protection Agency (USEPA) maintains networks of pollution monitors for two basic purposes: to check and enforce the attainment of national ambient air quality standards (NAAQS) and to provide useful data for studying pollution and its effects (40 C.F.R. § 58, Appendix D). These purposes imply conflicting criteria for the locations of a limited number of monitors. To check the attainment of standards, monitors are placed where pollution levels are highest. Monitors are not required where standards have always been met and there are no new pollution sources. To provide useful data for studying pollution and its effects, monitors are placed to observe outcomes under a variety of pollution levels.

Federal regulations that govern the monitoring networks reflect both purposes. In this paper, we make an initial systematic assessment of the relative importance of the regulations and the potential for biased sampling that may undermine the study of national public welfare.

Even for research into national public welfare conflicts about monitor placement arise. Often, different research questions suggest different placement designs. Exploring the transport of pollution requires a geographical distribution of monitor sites whereas measuring the health effects of pollution is best served by wide distributions of pollution levels and socioeconomic conditions across sites. Studying the welfare effects of pollution for American residents implies a monitoring network that provides a representative view of the experiences across that population.

The USEPA recently responded to the needs for representative sampling by organizing the National Core Network (NCore) multipollutant monitoring network, which began to operate formally in January 2011. But this is a modest start, requiring only one monitoring station for each state. The USEPA also runs a monitoring system assessment program that guides state, local, and tribal authorities responsible for monitor placement and maintenance. Our analysis complements these efforts by looking at the network as a whole.

In order to explore the factors that affect the spatial composition of the monitoring networks, this paper uses observations on tropospheric ozone (O_3) in the contiguous US between 1993 and 2011 gathered by the USEPA's Aeromatic Information Retrieval System (AIRS), (see http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm). Over this time period the number of monitoring stations in the US has increased from approximately 900 in 1993 to about 1,200 in 2011 (see figure 1).

O_3 is a criteria pollutant regulated by the USEPA. This pollutant is associated with a variety of respiratory illnesses and in more recent studies has been associated with premature mortality (Bell et al., 2004; Jerrett et al., 2009). O_3 also affects timber yields and crop production (Lesser et al., 1990). These myriad effects drive the determination of the NAAQs for O_3 which were first set in 1979. The standard was set at that time to be 120 ppb for the maximum hourly concentration over the course of a year. This approach held until 1997 when the standard was reduced to 80 ppb for the fourth highest daily maximum concentration averaged over three years. This standard was modified slightly to 75 ppb in 2008 (see http://www.epa.gov/ttnnaaqs/standards/o3/s_o3_history.html).

As stipulated by the Clean Air Act (CAA), the USEPA recently reviewed the NAAQS for ozone set in 2008. The standard was reduced again to 70 ppb. This new standard was motivated, in part, by additional evidence on the human health effects of ozone exposure. We focus on the ozone-monitoring network, providing a new view of the network that has produced data supporting the USEPA review.

That is, this paper explores the regulators' decision-making process in terms of locating air pollution monitors. What factors affect when a monitor is retired from the network? What drives the decision to add a new site? What causes year-to-year changes in the number of monitors? We tackle these questions with a particular focus on the relationship between regulatory compliance, pollution levels, and network design.

1.1 Policy Background

The USEPA's ambient air quality monitoring network has many components. The principle ozone-monitoring network is the State and Local Air Monitoring Stations (SLAMS). Within SLAMS, Photochemical Assessment Monitoring Stations (PAMS) measure ozone and its precursors specifically for areas of acute nonattainment. Since 2011, the USEPA also has used the Clean Air Status and Trends Network (CASTNET), which focuses on rural areas. Finally, Special Purpose Monitoring Stations (SPMS) provide ozone measurements from about 20 rural monitors that are part of the Portable

O_3 Monitoring System (POMS) network operated by the National Park Service (NPS). Figure A-1 shows the geographic distribution of these monitors.

The SLAMS network is a collection of monitors operated by state or local air monitoring agencies. The monitors are required to satisfy regulatory requirements and provide air quality information to public health agencies. As a result, the SLAMS network is focused on urban areas. As the regulatory agency, the USEPA must approve SLAMS sites and works with the state or local authorities on the design and maintenance of their network. While there are general monitoring requirements, the USEPA evaluates the state and local networks and may specify additional monitoring for particular circumstances.

Each state or local network must include a monitor site for one of six roles, all related to ozone. Three roles concern monitoring various pollution levels: (1) peak air pollution levels within the area covered by the network, (2) typical levels in densely populated areas, and (3) background concentration levels. A fourth role is to measure the impact of particular significant sources of pollution. Observing regional movement of air pollution between populated areas is the fifth role. Finally, monitors are dedicated to measuring pollution effects on visibility, vegetation damage, or other welfare-based impacts. SLAMS ozone monitors generally collect data exclusively on ozone. The regulations suggest no necessary interaction between siting ozone monitors and monitors for other pollutants.

In addition, SLAMS ozone monitoring networks must meet location requirements that depend on the size of the area covered and typical peak concentrations. Specific SLAMS ozone site minimum requirements are included in Table D-2 of 40 CFR Part 58, Appendix D, which is included as table A-1 in the appendix to this paper. The minima shown in A-1 are generally exceeded in order to cover all of the roles mentioned above. The regulations emphasize measuring maximum concentration and specifically state that every MSA must have at least one monitor for this purpose and specify siting requirements: "In many cases, these maximum concentration O_3 sites will be located 10 to 30 miles or more downwind from the urban area where maximum O_3 precursor emissions originate. The downwind direction and appropriate distance should be determined from historical meteorological data collected on days which show the potential for producing high O_3 levels. Monitoring agencies are to consult with their USEPA Regional Office when considering siting a maximum O_3 concentration site." (40 C.F.R. § 58, Appendix D, Section 4.1, Paragraph f)

Generally, monitor number and locations are proposed by state or local authorities and submitted to the USEPA, as part of the State Implementation Plan (SIP) required of all states. In practice, these groups work together to develop each monitoring network. The USEPA also provides technical guidance documents on ozone monitoring network design for evaluating the adequacy of each existing monitor, to relocate or retire an existing site, or to locate any new sites.

The regulations also address modification of existing monitoring networks (40 C.F.R. § 58.14). Generally, monitors required by an attainment or maintenance plan may not be altered. Discontinuation of a monitor is permitted when it "has shown attainment during the previous five years [and] … has a probability of less than 10 percent of exceeding 80 percent of the applicable NAAQS during the next three years based on the levels, trends, and variability observed in the past." Also, a monitor that "has consistently measured lower concentrations than another monitor for the same pollutant in the same [region] … during the previous five years" may be retired. Another monitor may replace a monitor that "is designed to measure concentrations upwind of an urban area for purposes of characterizing transport into the area" for the same purpose (40 C.F.R. § 58.14).

In 1993, the USEPA began requiring Photochemical Assessment Monitoring Stations (PAMS) sites in each ozone nonattainment area classified as serious, severe, or extreme. In 2006, new requirements dropped several siting specifications and required only two monitoring sites for each so-called PAMS area. The USEPA is currently considering additional changes to the requirements in order to give monitoring agencies more flexibility to pursue local objectives and to provide wider spatial distribution for research purposes. The PAMS network currently covers 25 areas with 75 sites. As shown in figure A-2, these sites are concentrated in California, Texas, and the northeastern seaboard.

The Clean Air Status and Trends Network (CASTNET) was established under the 1990 Clean Air Act Amendments. CASTNET was created to observe long-term trends in regional atmospheric sulfur, nitrogen, and ozone concentrations and deposition of sulfur and nitrogen pollutants in rural areas. The network is managed by the USEPA in cooperation with the National Park Service (NPS) and the Bureau of Land Management (BLM). In 2012, there were 91 CASTNET monitoring sites operating at 88 distinct locations. As figure A-3 shows, the USEPA operates sites in the east while the NPS operates sites in the west. The BLM operates four sites in Wyoming that joined the network in 2012.

In 2011, the USEPA put into operation a new National Core (NCore) network of monitors where multiple pollutants are measured. This network does not affect this study but we mention it for two reasons. First, the USEPA is considering requiring PAMS measurements at all NCore sites. This is one proposal of several addressing the desire, mentioned above, to give PAMS wider spatial distribution for research purposes. Second, the sites are chosen to be "representative" of a broad range of rural and urban sites, as opposed to focused in areas with high ambient pollution concentrations. Thus, the USEPA has acknowledged the unrepresentative nature of its SLAMS network.

1.2 Network Design

The two basic purposes of the USEPA monitoring network, to check and enforce the attainment of NAAQS and to provide useful data for studying pollution and its effects, represent a range of specific uses that are best served by different network designs. The purpose of this study is to estimate the relative importance of these two purposes and to serve as a first step in assessing inference problems that arise as a result. Two basic issues for any use are whether the monitor network data are representative and whether they are informative.

1.2.1 Representative Monitoring

The USEPA seeks to identify nonattainment of NAAQS wherever it occurs throughout the United States. To this end, monitors are maintained where nonattainment has occurred in the past and in locations where nonattainment is probable. At the same time, the USEPA wants to know the health effects of various pollution concentrations and to inform members of the general population about the pollution levels they face. Health effects are critical inputs to the choice of design values and their associated ambient standards. Providing accurate information to the public "sufficient to effectively participate in managing human health and environmental risks" (USEPA, 2016) is one of the basic missions of the USEPA. If the USEPA monitor network focuses on monitoring high pollution areas where nonattainment occurs or is probable then there is a risk that the USEPA cannot fulfill its mission for those people who live elsewhere.

Because monitors cannot be placed everywhere, the monitoring network must be used to predict (or forecast) criteria pollutant levels where there are no monitors. This is the fundamental motivation for representative monitoring. By monitoring a site that is typical of other, unmonitored locations, the USEPA could infer the conditions in those unmonitored areas, albeit with some uncertainty. Matters are, of course, more complicated because sites do not fall into a few types. Inference about unmonitored conditions requires extrapolation based on conditional probability models and estimation of model parameters. If the conditions of a conditional probability model characterize an informative distribution of pollution everywhere in the nation, then accurate information can be provided to public.

Usually, estimation rests on data collection, or sampling, that is also representative in the same conditional sense as the conditional probability model. Focusing monitors on high pollution areas results in two weaknesses related to conditioning. The first is that the variation in conditions may be too narrow to project accurately in other areas with markedly different conditions. We will discuss this weakness more in the next section on informative data. The second, more fundamental, weakness arises from the restriction that a probability model for pollution levels in all areas does not condition on pollution levels. Focusing on high pollution levels is an example of sampling conditional on pollution levels. From the perspective of statistical inference, the data from such sampling are unrepresentative or biased. Statistical inference that ignores this bias may be misleading.

For example, the USEPA's Office of Air Quality Planning and Standards manages a program called AirNow that is the USEPA's primary means of informing the public about air quality. AirNow is the national repository of real-time air-quality data and forecasts for the United States based on measurements from the ambient air monitoring networks. One of our long-term goals is to assess the potential biases that the current monitor network could have on AirNow forecasts. This paper takes a first step toward that goal by investigating sampling biases.

Another important use of the air quality monitoring data is research evaluating the effects of pollution on human health. Health is influenced not only by pollution but also by other environmental factors such as population density and socioeconomic characteristics. Sampling bias can confound these determinants of health. Researchers may then mistakenly attribute the effects of other factors to pollution exposure.

Note that sampling biases do not necessarily lead to misleading inference. Statistical inference may be able to account for sampling biases and avoid deception. In some cases, sampling bias may be present but inconsequential to inference.

1.2.2 Informative Monitoring

A second issue for monitor networks is how informative they are. Monitor placement affects the ability to measure the effects of such influences as pollution on health accurately. Even though it is unbiased, a measure may be ambiguous because a large amount of statistical uncertainty is present. If, for example, pollution concentration is similar over the monitor network then the influence of changes in concentration on health will be estimated imprecisely. Hence, the estimates will be uncertain.

In addition, monitor placement affects the degree of detail in the observed variation of pollution across regions. When neighboring monitors in a network produce highly correlated measurements, then several monitors are measuring essentially the same process. In such a case, one monitor can accurately forecast the conditions at the others. Alternatively, monitors might be placed so that they pick up the idiosyncrasies of each site so that their measurements are not highly correlated and forecasts based upon other monitors are inaccurate.

The USEPA and federal regulations have expressed attention to both aspects of creating an informative monitoring network. Required specifications for the roles of monitors ensure, in part, that monitors are not redundant. Remarks about representativeness also suggest the desire to obtain a wide variety of observed pollution levels for research on human health, ecological health, public welfare, and pollution transport. Specifically, data gathered at monitoring stations are used to assess the impact of exposure to air pollutants on a number of sensitive receptors: people, crops, timber, and ecosystems, to name a few. This phenomenon has broad importance in that Title I of the Clean Air Act (CAA) uses such scientific criteria—relationships between exposure and impact—to set the NAAQS. These standards are ostensibly set to protect human health and welfare with an adequate margin of safety. However, if collectively our sense of the threat that such pollutants pose to human health is derived from a biased set of observations then it is certainly possible, if not probable, that the NAAQs embody this bias.

Second, air quality models are tools used to estimate the link between emissions and concentrations. A critical role played by air quality models is the ability to estimate concentrations *in areas without monitors.* Importantly, these models are often calibrated to the monitor readings.[1] And if the set of observations is in some way biased, then the "correction" applied to the air quality models may not be appropriate. That is, the post-calibration estimates in locations without monitors will then embody the bias embedded in the monitor readings. Further, model projections in areas without monitors are (in part) used to inform new monitor placement decisions. Again, if the model is calibrated to a selected sample of monitors, then the estimates that are used to inform changes to network design are likely to be biased and future network design and arrangement is also likely to reflect this bias.

In its recent *Integrated Science Assessment for Ozone and Related Photochemical Oxidants*, (USEPA, 2013) the USEPA investigated the correlation among monitors in the networks of twenty cities (CITE). This investigation, admittedly, sought high correlations in support of health studies based upon aggregate pollution measures.

1.2.3 Empirical Approach.

In this context the current paper tests for endogenous monitor placement in the O_3 monitoring network using several approaches. First, the analysis explores the determinants of monitor density in each cross section from 1993 to 2012. That is, we regress two measures of network density (monitor counts by county and within a radius of each extant site) on lagged O_3 levels, and NAAQs compliance status, among

[1] This process, and guideline for it, is described by the USEPA at: http://www.epa.gov/ttn/scram/guidance/guide/final-03-pm-rh-guidance.pdf

other controls. We repeat this over all cross sections and then, using monitor fixed effects, assembling all years in a panel. The motivation for this initial specification is that if monitors in a given cross section are not placed endogenously with respect to O_3, the bias in the resulting surface is likely to be small. Detection of a link between lagged O_3 levels and monitor placement motivates more sophisticated modeling techniques.

Next, we use site-by-site observations and identify when a particular monitoring station is dropped (permanently) from the network. This approach encompasses locations that have been subject to measurement *at some time in the panel* but have been permanently dropped. By design it does not encompass locations that have never been monitored. A second approach identifies new sites in the network. An important consideration when modeling new monitors is to identify whether a new monitor is simply acting as a replacement for a retired monitor in basically the same location. We control for this by specifying an indicator for whether a monitor was dropped (one period prior) in the same county in which the new monitor is located.

These models make use of detailed information on each monitoring station including the number of (and reason for) missing observations and whether the station is managed by federal, state, or county governments or by private firms. The management controls are motivated by the prior discussion of different monitoring networks; these different network types (SLAMS, PAMS, CASTNET, e.g.) have somewhat different purposes and controlling for inclusion in the different networks is, therefore, likely to be important. Also included as controls are the age of the monitors, the number of monitors in a county, and the latitude and longitude coordinates of the monitors. Given the focus of the analysis we are particularly interested in testing whether lagged O_3 levels and lagged compliance status with the NAAQS affect the composition of the monitoring network. The "drop" probit models enable a test of whether lagged O_3 levels and regulatory compliance status influences the decision to remove a monitoring station from the network. The "new" probit models facilitate a test of whether lagged O_3 levels and regulatory compliance status affect the choice to add a monitor to the network.

The manner in which the NAAQS are implemented by the USEPA under the CAA also affects model specification. In particular, the USEPA requires that states submit State Implementation Plans (SIP) in order to demonstrate how each state will comply with the standards set under the CAA. In certain cases, the USEPA may promulgate a Federal Implementation Plan (FIP). This may occur if a state has not submitted a SIP (when it is required to do so), or if a submitted SIP is deemed to be incomplete or inadequate in some way (76 FR 48207, 2011). One aspect of a SIP that is especially relevant to the current analysis stipulates that plans for *monitoring* ambient pollution levels are typically delineated in or by a SIP (or its replacement FIP), (42 USC § 7410). As a result of this, regulators in a state facing a FIP may have different incentives in regards to monitor network composition. The econometric models control for SIP/FIP status.

The conceptual basis for potential associations between lagged O_3 readings, NAAQS compliance, and the decision to either add or drop a monitor to (or from) the network stem from network design decisions being embedded in the SIP approval process. Recall from the discussion above that SIPs cover, among other topics, rules for discontinuation and replacement of monitoring sites. Further, monitors in the PAMS network are placed in areas designated as out-of-attainment with the NAAQS.

1.2 Preview of Results

Beginning with the cross-sectional regressions, for each year, maximum O_3 readings are a significant and positive determinant of monitor counts when using either Poisson of negative binomial regressions. This finding holds for both the county and the distance-based (radius) monitor counts. Using the first-difference estimator, maximum O_3 readings are a significant and positive determinant of monitor counts in 15 out of 17 cross sections for the county specification. In the radius specification maximum O_3 readings are a significant and positive determinant of monitor counts in just 2 out of 17 cross sections.

In the panel, O_3 maximum readings are a significant and positive determinant of monitor counts for all county models. O_3 standard deviations are a significant and negative determinant of monitor counts for all county models. O_3 mean readings are also a significant and positive determinant of monitor counts for Poisson and negative binomial regressions for the county models. Using the distance based monitor counts O_3

maximums are a significant and positive determinant of monitor counts in the Poisson and negative binomial regressions.

O_3 readings are significantly associated with the decision to drop a monitor from the network throughout the 1993–2011 period. Specifically, a one part per billion (ppb) increase in the hourly maximum reading is associated with a -0.04% decrease in the probability that a monitor will be dropped in the subsequent time period($\alpha = 0.01$). We find limited evidence that O_3 NAAQs compliance status is associated with dropping a monitor; however, $PM_{2.5}$ compliance does predict the decision to drop a monitor from the network. The results also suggest that lagged ambient maximum readings are a significant factor in predicting whether to locate a new monitor in a given county. In particular, a one ppb increase to the hourly maximum reading is associated with approximately a 0.04% increase in the probability that a monitor will be added in the subsequent time period ($\alpha = 0.01$).

Because a new monitor may simply replace a retired monitor, one alternative estimation strategy is to restrict the sample to counties in which there *was not* a dropped monitor in the prior period in order to focus on net additions to the network. In this setting we find roughly equivalent marginal effects of ambient O_3 as documented above. However, we find evidence that O_3 NAAQs attainment status is a significant determinant of whether there is a new monitor in a given county in FIP states. In contrast, in SIP states, we find no evidence that attainment status affects the decision to add a monitor. Hence, whether a state manages its own compliance strategies or whether this role is assumed by the federal government is a significant factor in determining if lagged compliance status affects changes to the monitoring network.

So, summarizing the results of the current analysis, there are four essential findings.

1. Monitor counts are higher in areas with high maximum O_3 readings.

2. The regulator is less likely to drop an existing monitor for which prior maximum readings are high.

3. The regulator is more likely to add a new monitor to counties for which prior maximum readings are high.

    3.a. In counties that dropped a monitor during the prior period, the propensity of the regulator to add a new monitor is not affected by O_3 levels or attainment status.

    3.b. In counties that did not drop a monitor during the prior period, the propensity of the regulator to add a new monitor is positively associated with O_3 levels in states managed by a SIP and attainment status in states managed by a FIP.

4. Evidence of an association between adding a new monitor and prior O_3 attainment status is mixed.

    4.a. Regulators appear more likely to add monitors to counties with a legacy of non-attainment and this association is especially strong in FIP states.

    4.b. Regulators in states managed by a SIP are less likely to add a monitor to *previously unmonitored counties* that are chronically out of attainment.

The remainder of the manuscript is structured as follows. Section 2 touches on the areas in the literature to which this paper relates. Section 3 presents ours methods; this section describes the various model specifications as well as the data. Section 4 presents the empirical results and section 5 concludes.

2. Literature

To the authors' knowledge no papers in economics directly tackle the issue of pollution monitoring network design from the perspective of regulator choice. However, there are aspects of this paper that relate indirectly to areas in the economics literature. We begin with a discussion of papers that touch on the incentives faced by firms when the stringency of policy constraints varies across space. Hoel (1997) explores implications for plant or facility location in a context of multiple jurisdictions competing both in terms of growth or output and in terms of environmental quality. Hoel (1997), who draws on the work of Oates and Schwab (1996), notes that heterogeneity in stringency of environmental policy has implications for firms' location decisions. The interplay between exogenous and endogenous policy, explored by Markusen, Morey, and Olewiler (1993) is relevant in the current context as the federal government sets the O_3 NAAQs, but it leaves implementation up to the states (in most cases). Thus, the way in which states choose to implement federal standards may reflect industry composition within and across states, which in turn, may affect industry or market structure.

Many papers estimate the impacts associated with exposure to air pollutants (including tropospheric O_3). Notable examples include: Muller, Mendelsohn, and Nordhaus (2011) which conducted an environmental accounting exercise in the US economy in 2002, Adams et al., (1989) which quantified the value of crop yield losses due to O_3 exposure, and the USEPA's cost-benefit analyses of the Clean Air Act (USEPA, 1999; 2010) which, among other pollutants, assessed the impact on health, timber and agriculture of exposure to O_3.

A literature exists in the fields of atmospheric chemistry and physics specifically related to the design of air pollution monitoring networks. The World Health Organization outlined a list of nine objectives for the design of pollution monitoring networks (WHO, 1977). This broad list includes: the ability to evaluate health and environmental risk, evaluation of control strategies, and calibrating air quality models, among others. Only one objective appears to relate strictly to measurement without a direct connection to regulatory compliance: assessing spatial and time series trends in pollution levels. However, later papers zero-in on relatively few objectives in network design including: (1) maximization of the detection capability of peak pollution concentrations and (2) maximizing the detection capability of regulatory violations (Cheng and Tseng, 1997). Liu et al., (1986) state that early network composition was based on:

> "Subjective considerations; semiquantitative rules supported by experience; or sometimes, limited use of analytic tools like simple Gaussian Plume models."

Further, Liu et al., (1986) note that factors such as convenience in accessing monitoring stations (ostensibly for the purposes of installation or routine maintenance) also plays an important role in driving specifically where monitors are sited. And, as in Cheng and Tseng, 1997, placing monitors in such a way as to capture peak readings and regulatory violations appears to be the overarching objective. The problem, as Liu et al., (1986) note, is that optimization with respect to this goal requires some knowledge or prior information regarding the underlying spatio-temporal distribution of pollution. This can only come from two places: existing monitors or air quality models that are used to estimate levels. Since air quality models are calibrated to or validated against readings from monitors, network design is clearly endogenous with respect to existing information on pollution levels.

An additional objective in network design is minimizing the number of distinct monitoring stations (Venegas, Mazzeo, 2003). Liu et al., (1986) propose the Sphere Of Influence (SOI) approach which explores spatial correlations in predictions made by air quality models. This tack proposes siting a new monitor whenever the correlation between the readings at an existing station falls below a predetermined threshold value in a spatial correlation index.

Baldauf, Lane, and Marote (2001) identify a different objective that focuses on exposure analysis. Specifically, this approach optimizes design with respect to identification of impacts on human health. Potential monitoring locations are prioritized based on likely health risks to proximal populations. Baldauf, et al., (2001) describe a three-part optimization routine in which concentration, toxicity, and population density are used to identify differing levels of risk. The highest priority is given to potential locations with the highest risk. Although this approach embodies a slightly different objective in network design optimization, like the earlier papers, it is motivated by regulatory considerations,[2] not objective measurement of the pollution surface. Baldauf et al., (2002) explicitly state that:

> "Monitoring sites should be established at locations with the highest health risks to maximize the potential of obtaining a representative air quality measurement at the location(s) where adverse health effects occur."

However, selection of sites based on a metric that seeks out the highest health risks is not likely to yield a representative sample of air pollution measurements, by definition. That is, if the siting mechanism selects locations in the right-hand tail of the underlying distribution, it will produce measurements that tend to be larger than more typical observations from the distribution. This holds whether health risk or concentration levels comprise the objective function in network design.

Kainuma, Shiozawa, and Okamoto (1990) appear to move the discussion in a different direction by developing a multi-attribute strategy by which potential monitoring sites are ranked. These include factors such as effectiveness in providing readings on mean and episodic pollution levels, costs of building and maintaining the station, as well as

[2] Baldauf, Lane, and Marote (2003) note that the primary objective of the NAAQs is the protection of human health. Hence a methodology focusing on, effectively, a ranking of sites by health risks is still motivated by regulatory concerns, rather than purely measurement.

(but not limited to) population density. Kainuma et al., (1990) interview air quality regulators in an attempt to determine their preferences regarding the citing attributes. Air quality officers show a preference for networks with wide coverage and for coverage in highly polluted areas. Their ranking is declining in the number of stations (Kainuma et al., 1990).

3. Methods

The analysis of this paper focuses on a regulator's decision-making process regarding how to measure ambient O_3 pollution readings across space and through time. As mentioned above, if the primary goal or purpose of measurement were to obtain an unbiased estimate of the O_3 surface, the preferred approach would be to place the monitors exogenously (not necessarily randomly) with respect to observed levels and to distribute monitors more densely where the surface variation is greatest. This is described in some of the literature review above. Given a budget constraint, it is wasteful to put monitors in places where neighboring monitors are giving the exact same readings and better to put monitors where there is little correlation among neighboring monitors.. Historically, monitors are used to inform regulatory enforcement. As such, monitors are not randomly scattered across space. Instead, they are placed in areas where statutory exceedances are possible if not likely. Monitor placement is a function (at least partially) of the variable being measured. This contextual feature motivates an exploration of the factors that affect the spatial patterns in the network and the regulator's decision to make changes in the monitoring network[3].

3.1 Monitor Count

This analysis begins by examining factors that determine the number of monitors in various locations across the country. We begin cross-sectionally; for each monitor in each cross section (from 1994 to 2010) we tabulate the number of monitors in that county, $C^{c}{}_{i,t}$. The monitor count is then regressed on lagged ambient O_3 readings ($O_{i,t-1}$), the age of the monitoring station (Age_{it-1}), the fraction of missing observations at site *i*, $\boldsymbol{M}_{it-1}$, and the population in the county where monitor *i* is located, P_{it-1}. Note that (**M**) contains five covariates describing whether and why observations are missing. The reasons for missing data include maintenance, malfunction, damage, weather, and an anomaly. Model (1.c) also controls for location, L_i, (latitude and longitude) of the monitors. Model (1.c) is fitted using Poisson regression, negative binomial regression, and in first differences.

$$C^{c}{}_{i,t} = g(O_{i,t-1}, \mathbf{M}_{ct-1}, Age_{i,t-1}, P_{ct-1}) \qquad (1.c)$$

The second approach to modeling the monitor count computes the number of O_3 stations within a specified radius of each station in each cross section, $C^{r}{}_{i,t}$. Model (1.c) is fitted using this measure as the dependent variable. The default radius is 50 miles, and this measure is altered in a sensitivity analysis.

[3] We computed goodness-of-fit tests for these parametric models using the approach described in Andrews (1988). Despite fitting the data reasonably well, these tests reject the parametric models because the sample sizes are quite large. Given the qualitative agreement among our parametric models, we remain confident that the parameter estimates give an accurate representation of the patterns in the data.

In the count models, we employ three different measurements of ambient O_3 readings. The first is the O_3 season mean, a lagged average across monitoring stations in the county. The lag structure here is important because over such a long time period, there is likely to be considerable inertia in network design. That is, ambient concentrations many years earlier may affect the composition of the network which consists of capital installations that last many years. As such, the lag is a multi-year average back to the first year in the sample, 1993. So for observations in the year 1994, this is a one-year lag. For 1995, it is an average over 1994 and 1993, and so on up to 2011.

The second measurement is the O_3 standard deviation across monitoring stations in a given county, with the same lag structure as above. And the third is the maximum hourly O_3 level, across monitoring stations in a given county, with the same lag structure as above. We explore these different measures because the mean (what might be considered a default measure of O_3 levels) may not capture, in effect, what matters to the regulator in terms of whether or not to modify the monitoring network. That is, since 1979, the NAAQs for O_3 have been defined in terms of maximum hourly readings. (Note that the standards changed in 1997 and then again in 2008 and 2015, but their basic structure retains this hourly maximum-type definition.)

Model (2.c) adds controls for county-level attainment status, A_{it-1} with the O_3 NAAQs and the $PM_{2.5}$ NAAQs.

$$C^{c}{}_{i,t} = g(O_{i,t-1}, \mathbf{M}_{i,t-1}, Age_{i,t-1}, P_{i,t-1}, A_{i,t-1}) \qquad (2.c)$$

While this analysis focuses on O_3, we include $PM_{2.5}$ attainment status because some air pollution monitoring stations gather data on multiple pollutants. We use two different measures of attainment. The first measure of attainment is the sum of years, prior to *t*, in which a given county has been out of attainment with the current NAAQs. The measure changes through time for different counties in the sense that it is a running total. So, for example, if a given county was out of attainment in 1995 and in no other prior period, in 1996 this attainment measure would assume a value of unity. If, in 1996, this same county was out of attainment again, then in 1997 this measure would assume a value of two. The measure is computed separately for both O_3 and $PM_{2.5}$. The second measure is a categorical variable that assumes a value of unity if the monitor is in a county that was partially out of attainment in the prior period, and it assumes a value of two if the entire county was out of attainment in the prior period[4].

$$C^{c}{}_{i,t} = g(O_{i,t-1}, \mathbf{M}_{i,t-1}, Age_{i,t-1}, P_{i,t-1}, A_{i,t-1}, R_i) \qquad (3.c)$$

Model (3.c) controls for "ownership" of the monitors (R_i). These are a series of indicator variables coding whether each monitor is managed by federal, state, or county regulators, or if the stations are managed by a private firm.

The next step in modeling the distribution of monitoring stations across the country is to control for spatial fixed effects. Models (1.c), (2.c), and (3.c) are estimated by assembling each of the cross sections into a 17-year panel while controlling for monitor

[4] This is how attainment status is coded by the USEPA (see: http://www.epa.gov/oaqps001/greenbk/data_download.html)

fixed effects (γ_i) and the full set of year dummies (Y_t). The fixed-effects models are estimated using both the county-based and the radius-based monitor count specifications. Model (4.f) adds controls for changes to the NAAQs: the 1997 O_3 NAAQs revision, the 1997 and the 2006 $PM_{2.5}$ NAAQs revisions, indexed by N in (4.f). These are dummy variables that assume a value of unity *after* the regulatory change.

$$C^c{}_{i,t} = g(O_{i,t-1}, \mathbf{M}_{i,t-1}, Age_{i,t-1}, P_{i,t-1}, A_{i,t-1}, R_i, \gamma_i, Y_t, N_t) \qquad (4.f)$$

3.2 Decision to Drop Monitors

The next step in the empirical analysis uses a probit model to test whether O_3 levels and regulatory status affects the regulator's decision to remove a monitor from the network. Let (D_{it}) represent a dichotomous variable assuming a value of unity if monitor i is dropped from the network at time t and zero otherwise.

$$D_{it} = f(O_{it-1}, \mathbf{M}_{it-1}, Age_{it-1}, P_{it-1}, Y_t, L_i) \qquad (1.d)$$

Model (1.d) describes the decision to drop a monitor from the network at time t as a function of lagged ambient O_3 readings (O_{it-1}), the age of the monitoring station (Age_{it-1}), the fraction of missing observations at site i, $\boldsymbol{M}_{it-1}$, the population in the county where monitor i is located, P_{it-1}, as well as a year control, Y_t. Note that (**M**) contains five covariates describing whether and why observations are missing. Model (1.d) also controls for location, L_i, (latitude and longitude) of the monitors.

$$D_{it} = f(O_{it-1}, \mathbf{M}_{it-1}, Age_{it-1}, P_{it-1}, Y_t, L_i, A_{c,t-1}) \qquad (2.d)$$

Model (2.d) maintains the form from (1.d) while, as in (2.c), adding controls for county-level attainment status, A_{it-1} with the O_3 and $PM_{2.5}$ NAAQs.

$$D_{it} = f(O_{it-1}, \mathbf{M}_{it-1}, Age_{it-1}, P_{it-1}, Y_t, L_i, A_{c,t-1}, N_{t-1}) \qquad (3.d)$$

Model (3.d) includes indicator variables for the major regulatory changes: the 1997 O_3 NAAQs revision, the 1997 and the 2006 $PM_{2.5}$ NAAQs revisions, indexed by N in (3.d). These are dummy variables that assume a value of unity *after* the regulatory change.

$$D_{it} = f(O_{it-1}, \mathbf{M}_{it-1}, Age_{it-1}, P_{it-1}, Y_t, L_i, A_{c,t-1}, N_{t-1}, R_i, SIP_i) \qquad (4.d)$$

Model (4.d) controls for "ownership" or management agency of the monitors (R_i) and SIP status; states are coded (1) if they are governed by an approved state SIP and (0) if they are governed by a FIP.

3.2 Decision to Add Monitors.

We approach the specification of models that describe the decision to add a station to the monitor network in a similar manner to the models focusing on dropped monitors. There are, however, some important differences. In most cases, monitors are added to counties that *already have at least one* monitoring station. As such, the lag structure is one in which the county average becomes the relevant measure. That is, the analog to model (1.d), denoted (1.n) features county average O_3 measures, fractions of missing observations, monitor age, population, and location. Also note that we add a control for whether a monitor was dropped from the county in the prior period. The thrust is to distinguish between new monitors that replace a monitor that was eliminated versus a new monitor that in some sense adds to the network.

$$N_{it} = h(O_{ct-1}, \mathbf{M}_{ct-1}, Age_{ct-1}, P_{ct-1}, Y_t, L_c, Drop_{ct-1}) \qquad (1.n)$$

The next three models track closely to the approach used to model the dropped monitors discussed above.

$$N_{it} = h(O_{ct-1}, \mathbf{M}_{ct-1}, Age_{ct-1}, P_{ct-1}, Y_t, L_c, Drop_{ct-1}, A_{c,t-1}) \qquad (2.n)$$

Model (2.n) adds controls for county-level attainment status,A_{it-1}, as above.

$$N_{it} = h(O_{ct-1}, \mathbf{M}_{ct-1}, Age_{ct-1}, P_{ct-1}, Y_t, L_c, Drop_{ct-1}, A_{c,t-1}, N_{t-1}) \qquad (3.n)$$

Model (3.n) includes indicator variables for the major regulatory changes: the 1997 O_3 NAAQs revision, the 1997 and the 2006 $PM_{2.5}$ NAAQs revisions as in (3.d) above.

$$N_{it} = h(O_{ct-1}, \mathbf{M}_{ct-1}, Age_{ct-1}, P_{ct-1}, Y_t, L_c, Drop_{ct-1}, A_{c,t-1}, N_{t-1}, R_c, SIP_c) \qquad (4.n)$$

Model (4.n) controls for "ownership" of the monitors (R_i) and SIP status.

3.4 Data

The O_3 readings used in this analysis are publicly available from the USEPA's AIRS databases (USEPA, 2012). The data are comprised of hourly observations from individual monitoring stations over the period 1993 to 2011. The USEPA AIRS publishes extensive documentation on the monitors including reasons for missing observations, ownership or management of monitors, land use, and starting and ending date of service (see: http://www.epa.gov/ttn/airs/airsaqs/manuals/codedescs.htm).

Table A-2 in the appendix reports summary statistics for the O_3 monitoring data. The average O_3 reading across all sites in the panel is 33.5 ppb with a standard deviation of 7.64. The average maximum hourly reading is 103.1 ppb with a standard deviation of 22.6. Roughly 3 percent of the dataset are comprised of missing observations. Nearly all of these are due to monitor malfunctions or maintenance. Approximately two-thirds of the monitors are operated by state regulatory agencies, 11% are managed by county agencies, and 1% is operated by a private firm. Less than 1% is operated by the (federal) USEPA. Monitors for O_3 have been in use for an average of 15 years. The oldest monitor has been in use for 51 years.

Figure 1 indicates that in 1993 there were roughly 900 distinct stations recording O_3 in the contiguous US in 1993. The monitors were distributed among about 600 counties[5]. This network increased to approximately 1,200 stations in 2010. The number of stations has been roughly constant since 2003. Despite this, stations have been dropped and added (permanently) throughout the panel. As further evidence of continued change in the allocation of monitors across space, the number of counties with a monitor has increased to 750 in 2010.

Figure 2 shows changes to the monitoring network between 1993 and 2010. Dropped monitors have ranged between about 25 and 50 monitors dropped per year without a

[5] Note that the coterminous US is comprised of 3,100 counties in all.

discernible trend. New monitors were added at a rate of about 75 annually from 1993 to 2003 and then the rate of additions was more comparable to dropped sites. This pattern is of course reflected in the total number of monitors in the network with a distinct period of growth in the network up until about 2003.

The top-left panel of figure 3 shows that national average O_3 levels have been roughly constant over the 1993 to 2010 time period. The readings show a slight downward trend. There were two fairly major regulatory changes pertaining to O_3 that occurred in this time period. In 1997, the O_3 NAAQs were revised for the first time since 1979. The standard was reduced from a maximum hourly concentration of 120 ppb to 80 ppb. In 2008, this was further reduced to 75 ppb. While these changes have had a limited effect on the mean, the maximum O_3 readings have declined sharply since the early 1990's. The top-right panel of figure 3 shows this phenomenon. Along with a reduction in the maximum hourly readings, the O_3 standard deviation has declined since the 1990's (shown in the bottom-left panel of figure 3) as has the coefficient of variation (bottom-right panel).

In the appendix, figure A-4 and table A-3 show the differences in terms of O_3 levels by monitor ownership. Ownership is broken up into the following groups: federal monitors (which tend to be operated by the USEPA, but also include those run by the military and other agencies), state monitors which are operated by state environmental protection agencies, and non-federal monitors which are mostly comprised of sites run by city and county municipal governments. The key insight of table A-3 and figure A-4 is that federal monitors tend to be placed in areas of relatively higher O_3 than monitors run by other agencies.

Figures A-5 and A-6 in the appendix display the share of counties out of attainment with the $PM_{2.5}$ NAAQs and the O_3 NAAQs, respectively. Prior to 2005[6], NAAQs for PM was defined in terms of PM_{10}. Non-attainment rates were relatively low; figure A-5 indicates that less than 5% of counties with monitors were out of attainment. In 2006, the $PM_{2.5}$ NAAQs (proposed in 1997) was enacted. Non-attainment rates jumped to over 10% of monitored counties. In 2008, the second $PM_{2.5}$ NAAQs revision came into effect. Compliance rates with this standard are about 7.5%.

Figure A-6 in the appendix displays compliance rates with the O_3 standards. Prior to 2005, the acting standard for O_3 was a 1-hour maximum reading of 120 ppb. Compliance rates improved dramatically from the 1980's (more than 20% of monitored counties out of attainment) to 2005 (about 15% non-compliant). In 2006, an 8-hour standard was enacted. Non-compliance rates jumped back above 20%, but then rapidly declined back to 15%.

4. Results

Table 1 explores the differences in various O_3 measures according to whether a site is dropped or retained in the network or if a new site is added. For sites that are dropped, the (one year lagged) mean O_3 level is equivalent to sites that are retained. However, the

[6] Note that the date of passage and enactment of the NAAQs revisions may, and in fact do, differ. In specifying the NAAQs revision controls in the econometric models we employ the date of enactment.

hourly maximum reading is significantly lower at sites that are dropped than at retained sites. Specifically, the mean hourly maximum O_3 reading in the sample is 99 ppb at sites that are dropped the following year. The mean hourly maximum O_3 reading in the sample is 103 ppb at sites that are retained the following year (the difference is significant at $\alpha = 0.01$). Since the O_3 NAAQs are set according to the hourly maximum values, it makes sense that sites with higher maximum values are retained in the network; they are likely to be in counties that are (at least at risk of) being out of attainment the following period. Hence, such sites are important for compliance assessments. And recall from section 1.1 that, consistently low relative readings is grounds for removal of a monitor from the network (40 CFR 58.14).

The standard deviation measure is also significantly lower in sites dropped from the network than for retained sites. The standard deviation at dropped sites is 17.6 ppb, while at retained sites it is 18.2 ppb. This difference is significant at $\alpha = 0.01$. Monitors with greater variability are likely to have higher peaks readings than sites with lower standard deviations, all else equal. As such, NAAQs violations are probably more apt to occur and these more variable monitors are also likely to be in counties at risk of non-attainment. Hence, they tend to be retained.

The right-hand panel of table 1 displays the same cross-tabulations except for new and existing monitors. Recall that the measurements in this context are one-year lagged county averages. The mean O_3 level is marginally greater at new sites than at existing sites: the difference of 0.4 ppb is significant at $\alpha = 0.10$. The difference in the maximum hourly readings is about 5 ppb which is significant at $\alpha = 0.01$. The standard deviation shows a small, weakly significant difference. The right-hand panel of table 1 suggests that O_3 monitors are placed in locations with higher mean O_3 levels, higher peak readings, and greater standard deviations. In summary, table 1 provides initial evidence that changes to the monitor network are associated with differences in O_3 levels.

4.1 Monitor Count

Figure 4 plots the parameter estimates corresponding to maximum O_3 readings from the cross-sectional regressions. The results are from model (4.c). The top two panels correspond to Poisson regression. Beginning with the top left panel of figure 4, O_3 maximum readings are a significant, positive determinant of the monitor count for all 17 years when using the county monitor count. The effect on a one-ppb increase in O_3 maximum readings ranges between 0.02 and 0.04. The top right panel uses the distance count method. Notice that the magnitude of the parameter estimates is smaller and the standard errors are larger (the 95 percent confidence intervals are wider). Casual empiricism suggests a reason for the systematically lower precision when using the distance-radius monitor count models. Clean Air Act attainment status is defined or determined at the county level. Therefore, readings drawn from the monitoring network within a county are likely to drive whether monitors are added to or dropped from (the count) the network. The distance-radius approach may cross county (or even state) lines.

The middle panel shows the cross section results from the negative binomial models. The parameter estimates are quite similar to those produced by the Poisson models. That is, using the county count method (left panel) the effect of a one-ppb increase in maximum O_3 readings ranges between 0.02 and 0.04. The right-hand panel, employing the distance count, shows smaller parameter estimates and wider confidence intervals.

The bottom panel reports the results from the first-difference models. Using the county count, O_3 maximum readings are a significant, positive determinant of monitor counts in 15 out of 18 years. Using the distance count O_3 maximum readings are a significant, positive determinant of monitor counts in only 2 of 18 years. In fact, O_3 maximums are significantly, negatively associated with O_3 monitor counts in 6 out of 18 years.

Table 2 reports the partial results for the panel data monitor count models. The table is bifurcated such that the left-hand panel reports results from the county count models while the right-hand panel displays results from the distance count models. Beginning with the county count models, in specification 1.c, O_3 maximum readings are a significant (α = 0.01), positive determinant of the number of monitors in a given county for all three regressions (Poisson, negative binomial, and first differences). A one ppb increase in O_3 maximum readings yields between a 0.008 (Poisson and negative binomial) and a 0.07 (first difference) increase to the monitor count, per county.

O_3 mean readings are also significant (α = 0.01), positive determinant of the number of monitors for Poisson and negative binomial regressions. A one unit increase to the average O_3 reading yields a 0.03 increase to the monitor count. And O_3 standard deviations are a significant (α = 0.01), negative determinant of the number of monitors for Poisson and negative binomial regressions. A one unit increase in the O_3 standard deviation is associated with a 0.07 decrease in the monitor count. In the first difference regression, O_3 standard deviations also have a significant (α = 0.05), negative effect on the number of monitors in the first difference regression. These results are robust across the 1.c, 2.c, and 3.c models.

For the distance count models O_3 maximum readings are significantly (α = 0.01), positively associated with monitor counts in the Poisson and negative binomial regressions. A one ppb increase in O_3 maximum readings yields a 0.006 (Poisson and negative binomial) increase to the monitor count, per county. O_3 averages are significantly and negatively associated with monitor counts for Poisson (α = 0.10) and negative binomial (α = 0.01) regressions. O_3 standard deviations are positively associated with monitor counts in the negative binomial (α = 0.01) and first difference models (α = 0.05). These results are robust to the three different model specifications: 1.c, 2.c, 3.c. Comparing the county and distance monitor counts, only the association between the O_3 maximum readings and counts is robust to the different count approaches.

4.2 Dropped Monitors

Table 3 displays the results from the probit regressions in which the dependent variable measures whether a monitor is dropped (1) or retained (0) in the network (table 3 reports marginal effects). Table 3 reveals that dropped monitors are associated with lower peak O_3 levels in each of the four model specifications. For each model, a one unit increase in the maximum hourly O_3 reading is associated with roughly a 0.03% decrease in the probability of the monitor being dropped in the next period. The associations are significant at $\alpha = 0.01$. This result reinforces the cross-tabulations reported in table 1.

Using the cumulative measure of attainment status (sum of prior years out of attainment) with the O_3 NAAQs, we find evidence of an association between attainment status and whether a monitor is dropped from the network. In particular, in models 2.d and 3.d cumulative non-attainment is associated with a reduction in the probability of a monitor being dropped ($\alpha = 0.01$). This association is not detected in model 4.d which adds SIP status. In contrast, a one year increase in the number of periods out of attainment with the $PM_{2.5}$ NAAQs is associated with a 0.03% increase in the probability of a monitor being dropped. Note that the attainment measures employed in table 2 tabulates the sum of years out of attainment up to the year in which the regulator determines to drop or retain the monitor.

The passage and enactment of the 1997 O_3 NAAQs increases the probability that monitors are dropped from the network. In models (3.d) and (4.d) monitors are about 1.2% more likely to be dropped from the network after the 1997 NAAQs ($\alpha = 0.05$). Passage of the second $PM_{2.5}$ NAAQs revision is associated with a 0.8% increase in the probability of a monitor being dropped from the network. Counties in states that manage attainment with a SIP are about 7% more likely to have a monitor dropped than FIP states.

A number of other controls included in models (1.d) through (4.d) are significantly associated with dropping a monitor. As the average age of monitors increases, the likelihood of a monitor being dropped decreases up to approximately 20 years; beyond 20 years, increasing age of monitors increases the chance that a monitor will be dropped from the network. Of the four different reasons for missing observations, malfunction and maintenance are consistently significantly associated with dropping a monitor. A greater share of missing observations due to monitor malfunction is positively and significantly associated with the monitor being dropped ($\alpha = 0.01$) in all four models. Specifically, a one percentage point increase in the share of observations missing due to malfunction results in roughly a 10% increase in the probability of the monitor being dropped. Missing observations due to maintenance are also positively associated with a monitor being dropped in all four models ($\alpha = 0.01$); a one percentage point increase in the share of observations missing due to maintenance results in roughly a 13% increase in the probability of the monitor being dropped.

In terms of ownership of the monitoring stations, counties with a higher share of monitors operated by county regulators are 5.5% more likely to add an additional monitor than other counties with monitors run by other local regulators. Table 3 indicates evidence of a slight, downward yearly trend with the probability of adding a monitor dropping by 0.2% to 0.3% per year ($\alpha = 0.01$). The number of monitors in a county (not shown in table 3), the average age of monitors, and the population are also positively, and significantly associated with the probability of adding a monitor to the network.

4.3 New Monitors

Table 4, which also reports marginal effects, displays the results from the probit models focusing on new monitors in the network. Specifically, the response variable is coded 1 if a monitor was not active in any prior period and 0 otherwise. Because this tack focuses on new monitors, many of the covariates are one year lagged averages of observations in the county receiving the new monitoring station. Counties that had a monitor dropped from the network in the prior period are 12% more likely to have a monitor added ($\alpha = 0.01$) than counties without a dropped monitor in the prior period.

Table 4 indicates that maximum hourly O_3 readings are positively associated with the addition of a new monitor in models 1.n and 4.n. That is, the average maximum reading (across existing monitors in a given county) during the prior year is associated with an increase in the probability of a new monitor being located in that county of about 0.04%. This association is significant at ($\alpha = 0.05$). In models 2.n and 3.n, the average O_3 level is associated with a 0.07% increase in the likelihood of adding a new monitor ($\alpha = 0.05$). In addition, the standard deviation of O_3 readings is associated with a decrease in the probability of a new monitor being located in a county of between 0.16% and 0.27%. This association is significant at ($\alpha = 0.01$) in models (2.n), (3.n), and (4.n) and ($\alpha = 0.05$) in model (1.n).

A one-unit increase in O_3 NAAQs attainment status is associated with an increase of 0.04% in the probability of a new monitor being added to the network ($\alpha = 0.01$). Attainment status with the $PM_{2.5}$ NAAQs is not associated with the probability of adding a new monitor. There is no evidence that the passage of the NAAQs revisions impact the decision to add a new monitor to the network. If a state employs a SIP to manage NAAQs compliance, the likelihood of adding a new station to the network increases by about 2% ($\alpha = 0.01$) relative to a state subject to direct federal management under an FIP. Having a monitor dropped from the network in the same county (one year earlier) increases the probability of adding a monitor by 12% ($\alpha = 0.01$).

4.3.1. New Monitors and SIP, Drop Status.

The importance of a county having dropped a monitor in determining whether that same county subsequently adds a monitor speaks to a larger point in the analysis. That is, the question of whether to replace a monitor that has mechanically failed or has been permanently removed from the network for some other reason is a distinct question from whether to add a new monitor to the network in order to obtain observations in areas not previously subject to measurement. In order to examine these two questions separately, table 5 reports results from model (4.n) applied to two restricted samples.

First, table 5 reports results from model (4.n) fitted to data from counties that *did not* have a monitor dropped in the prior period. Second, table 5 reports results from model (4.n) fitted to data from counties that *did* have a monitor dropped in the prior period. Table 5 also differentiates between states that implement approved SIPs from states that are subject to FIPs. The impetus for this latter distinction being that SIP states may face different incentives (than FIP states) that are related to regulatory enforcement and may have a direct bearing on network composition.

The top panel of table 5 shows the results when the sample is restricted to counties in which a monitor was not dropped in the prior period. These cases constitute net additions to the network. The first (left hand) column indicates that the marginal effects of O_3 readings (maximums, means, and standard deviation) are roughly equivalent to those reported for model (4.n) in table 4 with the exception of weak evidence that average O_3 levels affect the propensity to add a monitor ($\alpha = 0.10$). Specifically, the average maximum reading (across existing monitors in a given county) during the prior year is associated with an increase in the probability of a new monitor being located in that county of 0.05% ($\alpha = 0.10$). The O_3 standard deviation is negatively associated with the installation of a new monitor ($\alpha = 0.01$). These marginal effects are relatively robust to including just SIP states. There is no evidence that maximum or average O_3 readings affect the decision to add a monitor in states subject to an FIP. We detect weak evidence that O_3 standard deviations are associated with a decrease in the likelihood of adding a monitor ($\alpha = 0.10$) in FIP states.

Also of interest in table 5 is the association between attainment status and the probability of adding a new monitor. As in table 4, O_3 NAAQs attainment status shows a significant association with the probability of adding a monitor across all states. When the sample is restricted to just SIP states there is no evidence of an association. In contrast, for states subject to a FIP, there is strong evidence of a positive, significant association: being out of attainment in a state subject to direct federal management (as implied by the FIP) increases the probability of adding a new monitor by 0.11% (α = 0.01). For counties without a dropped monitor, O_3 NAAQs attainment status appears to be a factor in determining whether a new monitor is added to the existing network *in states that are subject to a FIP*. Attainment status has no impact on new monitors in states governed by their own SIP.

The findings in the top panel of table 5 yield some insight as to regulatory strategies in FIP and SIP states. Since SIPs and FIPs specify, in part, strategies for monitoring O_3 the fact that O_3 readings are only significant in the SIP sample suggests state level regulators incorporate information on pollution levels into network composition while federal regulators (those that develop the FIPSs) do not. In contrast, federal regulators make additions to the network based on cumulative evidence of O_3 NAAQs attainment. This is more of a long-term, stable measure which is based on existing federal regulation.

In terms of $PM_{2.5}$ attainment, table 5 reports weak evidence of an association with the decision to add a monitor in all states and in FIP states. In all states, the effect on a one-year increase in years out of attainment is a 1.9% increase in the likelihood of adding a monitor (α = 0.10). In FIP states, the effect on a one-year increase in years out of attainment is a 4.3% increase in the likelihood of adding a monitor (α = 0.10). In counties that did not drop a monitor in the prior period, attainment status has no impact on new monitors in states governed by their own SIP. Further, we find no evidence that O_3 NAAQs revisions affects new monitor placement.

The bottom panel of table 5 focuses on counties in which a monitor was dropped in the prior period. In all states, O_3 peak readings reduce the likelihood of adding a monitor (replacing a dropped monitor). Specifically, a one unit increase in O_3 maximum readings reduces the probability of replacing a dropped monitor by 1.9% ($\alpha = 0.01$). This effect is driven by SIP states. That is, the association is also detected in only SIP states and there is no evidence of an association for FIP states. In addition, there is no evidence that O_3 NAAQs attainment status is associated with the likelihood of adding a monitor to the network in counties where a monitor was dropped in the prior period.

Table 6 explores the significance of two different attainment measures in determining: the monitor count (column 1), the probability a monitor is dropped from the network (column 2), and added to the network (column 3). In each case, the reported coefficients are estimated in model specification (4) which includes controls for O_3 readings, missing observations, and monitor ownership. The first measure of attainment is the sum of years in which a given county has been out of attainment with the current NAAQs. (This is the default measure used throughout the paper.) Recall that the measure changes through time for different counties in the sense that it is a running total. The measure is computed separately for both O_3 and $PM_{2.5}$. Table 6 indicates that there is no evidence of an association between O_3 attainment status (in the cumulative measure) and the probability of dropping a monitor, or the monitor count. However, this measure is positively associated with the decision to add a monitor. Attainment status for $PM_{2.5}$ is significantly associated with the probability of dropping a monitor. First, one additional year of cumulative, non-attainment is associated with a 0.03% increase in the likelihood of dropping a monitor ($\alpha = 0.01$). Table 5 suggests that this effect manifests in counties without a dropped monitor in the prior period.

The second measure of attainment status is an indicator variable which is coded (2) if the entire county was out of attainment in the prior period, (1) if a part of the county was out of attainment in the prior period, and (0) if the county was deemed to be compliant. For the dropped monitor models, O_3 attainment status is not associated with changes to the monitoring network. For the add monitor model, being out of attainment is weakly associated ($\alpha = 0.10$) with an increase in the probability of adding a monitor; a one unit increase in attainment status is associated with an increase of 0.004 to the likelihood of adding a monitor. For $PM_{2.5}$, a one unit increase in non-attainment status is associated with a 0.7% increase in the likelihood of dropping a monitor ($\alpha = 0.01$).

Table 7 extends the analysis by incorporating monitors that are added to the network *in counties that do not have a monitor in the prior period.* The drawback of this tack is that none of the lagged measures of O_3 levels can be included. However, because USEPA *estimates* attainment status for all counties in the U.S. we can fit a model with reasonable explanatory power to the more inclusive sample of new monitors. Column (1) of table 7 simply reports the results of fitting model (4.n) to the original sample (denoted: "No New Counties") for the purposes of comparison with the spatially extended sample. In the original sample, only O_3 attainment status and the SIP indicator are associated with the probability of adding a new monitor. However, including previously unmonitored counties into the sample produces some significant, and different, associations.

First, when the new monitors in previously unmonitored counties are added, O_3 attainment status is negatively associated with the likelihood of adding a new monitor; a one unit increase in the years of non-attainment status is associated with a 0.1% decrease in the likelihood of adding a monitor ($\alpha = 0.01$). This is greater than a two-fold increase in the marginal effects relative to the default specification. Further, SIP states are more likely to add a monitor. Although evidence of this association was detected in the sample without previously unmonitored counties, the magnitude of the effect of SIP status is nearly one order of magnitude larger in the expanded sample. The second $PM_{2.5}$ NAAQs revision is also associated with an increased probability (3.0% increase, $\alpha = 0.01$) of adding a monitor.

The third column restricts the sample to SIP states (again inclusive of previously unmonitored counties). The results for SIP states mirror those of the full sample in that the O_3 attainment measure is negatively associated with the probability of adding a monitor ($\alpha = 0.01$), and the magnitude of the effect is nearly equivalent to the full sample. Further, the effect of the second $PM_{2.5}$ NAAQs revision is also significant and similar in magnitude to the full sample. The results for FIP states suggest that the associated between O_3 attainment and the decision to add a monitor is weaker. First, the magnitude of the effect is one-half that estimated in SIP states. Second the level of significance drops to α = 0.05. The association between the second $PM_{2.5}$ NAAQs revision and adding a monitor persists in FIP states.

The fact that prior O_3 non-attainment is inversely related to the decision to add a monitor *when the sample includes previously unmonitored counties* suggests that regulators are averse to spatial expansions to the network in counties in which attainment status is in question. Coupled with the fact that SIP states are nearly 20% more likely to add monitors, the negative association with prior O_3 attainment status suggests that state regulators are relatively eager to add new monitors, but *not* in counties with a history of non-attainment.

In contrast, when the sample is restricted to counties previously monitored, regulators appear to target additional monitors at problematic (chronically out-of-attainment) counties. A few candidate explanations emerge. First, regulators may choose to add monitors to non-attainment counties to try to *lower* the average reading in that county (in an attempt to improve future attainment status). This might be achieved by strategically placing the monitor in cleaner zones within the county. Second, recall from table 5 that the positive association between O_3 attainment status and propensity to add a monitor is strongest in FIP states. Since these states are directly managed by federal regulators, this positive association may be intended to increase precision of O_3 measurements by adding more observations (new monitors). These findings provide evidence of a complex relationship between NAAQs attainment status and network composition that depends on SIP/FIP status, and whether the sample under consideration includes previously unmonitored counties.

5. Conclusions

In the US, air pollution monitors are used to assess environmental conditions and for purposes of regulatory enforcement. As we noted previously, if the primary objective of measurement were to derive an unbiased estimate of the O_3 surface, the approach would be to place the monitors exogenously with respect to observed levels and to allocate monitors more densely where the surface variation is greatest. However, because the monitors are used for enforcement, the monitors are *not randomly placed.* The stations tend to be in areas where compliance with existing standards and rules is in question. The goal of this paper is to test whether monitor location is endogenous with respect to ambient pollution levels and regulatory compliance status. A secondary goal is to describe the factors that systematically affect the monitor network.

Specifically, the paper asks: What factors affect when a monitor is retired from the network? What drives the decision to add a new site? What factors are associated with the spatial distribution of monitors? We tackle these questions using a panel dataset of monitors from the O_3 network over the period 1993 to 2011. Three specifications are employed: a count model with monitor density proximal to each extant station as the response variable; a probit model with the decision to drop existing monitors from the network as the response variable; and a probit model with the decision to add new monitors to the network as the response variable.

The empirical results of the paper can be boiled down to four essential points.

1. Monitor counts are higher in areas with high maximum O_3 readings.

2. The regulator is less likely to drop an existing monitor for which prior maximum readings are high.

3. The regulator is more likely to add a new monitor to counties for which prior maximum readings are high.

   3.a. In counties that dropped a monitor during the prior period, the propensity of the regulator to add a new monitor is not affected by O_3 levels or attainment status.

   3.b. In counties that did not drop a monitor during the prior period, the propensity of the regulator to add a new monitor is positively associated with O_3 levels in states managed by a SIP and attainment status in states managed by a FIP.

4. Evidence of an association between prior O_3 attainment status is mixed.

   4.a. Regulators appear more likely to add monitors to counties with a legacy of non-attainment and this association is especially strong in FIP states.

   4.b. Regulators in states managed by a SIP are less likely to add a monitor to *previously unmonitored counties* that are chronically out of attainment.

More specifically, the empirical results indicate that O_3 readings are significantly associated with the decision to drop a monitor from the network throughout the 1993–2010 period. Specifically, a one part per billion (ppb) increase to the hourly maximum reading is associated with a -0.04% decrease in the probability that a monitor will be dropped in the subsequent time period($\alpha = 0.01$). We find some evidence that O_3 NAAQs compliance status is negatively associated with dropping a monitor.

The results also suggest that lagged ambient maximum readings are a significant factor in determining to add a new monitor to the network; a one ppb increase to the hourly maximum reading is associated with approximately a 0.03% increase in the probability that a monitor will be added in the subsequent time period ($\alpha = 0.01$). O_3 compliance status is significantly and positively associated with the regulator's decision to add a monitor to the network ($\alpha = 0.01$).

Because a new monitor may simply replace an old, retired monitor, one alternative estimation strategy is to restrict the sample to counties in which there *was not* a dropped monitor in the prior period in order to focus on net additions to the network. In this setting we find roughly equivalent marginal effects of ambient O_3 as documented above in SIP states. And, we find evidence that O_3 NAAQs attainment status in the prior period is a significant determinant of whether there is a new monitor in a given county in FIP states. In contrast, in SIP states, we find no evidence that attainment status affects the decision to add a monitor. Hence, whether a state manages its own compliance strategies or whether this role is assumed by the federal government is a significant factor in determining if lagged compliance status affects changes to the monitoring network. When the analysis is expanded to encompass new monitors in previously unmonitored counties, we detect evidence that regulators in SIP states are less likely to add a new monitor to counties chronically out of attainment.

This analysis suggests future research on a number of fronts. First, further work could apply this apparatus to other pollutants. Do these results found herein hold for $PM_{2.5}$ monitoring stations or for sulfur dioxide? Second, future research could test for selection bias in the O_3 surface resulting from the endogenous placement of monitors. This might explore whether selection, and the mechanisms driving selection, changed over time and to what extent regulatory constraints have affected estimates of such bias. Third, authors could explore the removal of selection bias from the O_3 surface and test whether this impacts our understanding of how O_3 affects health and other valuable receptors such as crops and trees.

Tables

Table 1: O_3 Measures for Dropped and Retained Monitors

Seasonal O_3 Measure (ppbv)	**Drop Status**		**Add Status**	
	Drop = 1	Drop = 0	New = 1	New = 0
Mean	33.25	33.48	33.95*	33.53
	(0.289)	(0.052)	(0.301)	(0.066)
Maximum	99.18***	103.24	106.01***	101.65
	(0.889)	(0.154)	(0.870)	(0.200)
Standard Deviation	17.61***	18.16	18.40*	18.16
	(0.158)	(0.028)	(0.154)	(0.037)
Obs.	731	21,293	620	9,518

For t-tests: *** p<0.01, ** p<0.05, * p<0.1

Table 2: Monitor Count and O_3 Levels: 1993–2012.

	Lag O_3 Readings in County			Lag O_3 Readings within 50 Mile Radius		
Model 1.c	**O_3 max**	**O_3 mean**	**O_3 sd**	**O_3 max**	**O_3 mean**	**O_3 sd**
Poisson	0.008***	0.029***	-0.071***	0.006***	-0.011*	0.017
	(0.001)	(0.008)	(0.013)	(0.002)	(0.006)	(0.011)
Negative	0.008***	0.029***	-0.071***	0.006***	-0.011***	0.017***
Binomial	(0.001)	(0.008)	(0.013)	(0.001)	(0.001)	(0.003)
First	0.069***	-0.015	-0.352**	-0.012	-0.015	0.135**
Differences	(0.019)	(0.112)	(0.160)	(0.011)	(0.035)	(0.064)
Model 2.c	**O_3 max**	**O_3 mean**	**O_3 sd**	**O_3 max**	**O_3 mean**	**O_3 sd**
Poisson	0.008***	0.030***	-0.071***	0.006***	-0.011*	0.015
	(0.001)	(0.008)	(0.013)	(0.002)	(0.006)	(0.011)
Negative	0.008***	0.030***	-0.071***	0.006***	-0.011***	0.015***
Binomial	(0.001)	(0.008)	(0.013)	(0.002)	(0.006)	(0.003)
First	0.069***	-0.015	-0.352**	-0.012	-0.015	0.135**
Differences	(0.019)	(0.112)	(0.160)	(0.011)	(0.035)	(0.064)
Model 3.c	**O_3 max**	**O_3 mean**	**O_3 sd**	**O_3 max**	**O_3 mean**	**O_3 sd**
Poisson	0.008***	0.029***	-0.070***	0.006***	-0.011*	0.015
	(0.001)	(0.008)	(0.013)	(0.002)	(0.006)	(0.011)
Negative [A]	0.008***	0.029***	-0.070***	0.006***	-0.011***	0.015***
Binomial	(0.001)	(0.008)	(0.013)	(0.002)	(0.001)	(0.003)
First	0.069***	-0.015	-0.352**	-0.012	-0.015	0.135**
Differences	(0.019)	(0.112)	(0.160)	(0.011)	(0.035)	(0.064)

Robust Standard Errors in Parenthesis (except for negative binomial models): *** p<0.01, ** p<0.05, * p<0.1. All models contain full list of controls as in (1.c), (2.c), and (3.c). Poisson and negative binomial employ site fixed effects, first-differences employ county fixed effects. n = 24,460.

Table 3: The decision to drop a monitor: 1993–2009.

	(1.d)	(2.d)	(3.d)	(4.d)
Year	-0.000831***	-0.000901***	-0.00178***	-0.00174***
	(0.000264)	(0.000282)	(0.000612)	(0.000606)
O_3 max, t-1 (ppb)	-0.000345***	-0.000304***	-0.000301***	-0.000331***
	(9.76e-05)	(0.000102)	(0.000101)	(0.000103)
O_3 sd, t-1 (ppb)	-0.000843*	-0.00141***	-0.00128***	-0.00101**
	(0.000473)	(0.000480)	(0.000482)	(0.000483)
Age	-0.00260***	-0.00248***	-0.00247***	-0.00228***
	(0.000345)	(0.000343)	(0.000342)	(0.000342)
Age^2	5.96e-05***	5.60e-05***	5.57e-05***	5.10e-05***
	(9.06e-06)	(9.01e-06)	(8.99e-06)	(8.96e-06)
Share Missing Obs. from Damage	0.0592	0.0555	0.0573	0.0614
	(0.128)	(0.131)	(0.133)	(0.130)
Share Missing Obs. from Malfunction	0.105***	0.104***	0.105***	0.100***
	(0.0170)	(0.0171)	(0.0171)	(0.0172)
Share Missing Obs. from Maintenance	0.149***	0.137***	0.136***	0.128***
	(0.0418)	(0.0419)	(0.0417)	(0.0425)
Share Missing Obs. from Weather	-0.0841	-0.0884	-0.109	-0.0985
	(0.154)	(0.153)	(0.157)	(0.151)
Sum of O_3 Non-Attainment Years		-0.000207***	-0.000209***	-0.000114
		(7.61e-05)	(7.60e-05)	(7.66e-05)
Sum of $PM_{2.5}$ Non-Attainment Years		0.000334***	0.000324***	0.000326***
		(5.42e-05)	(5.38e-05)	(5.36e-05)
O_3 NAAQs Revision			0.0116**	0.0114**
			(0.00567)	(0.00564)
PM NAAQs Revision (2006)			-0.00690	-0.00724
			(0.00531)	(0.00528)
PM NAAQs Revision (2009)			0.00881**	0.00812**
			(0.00414)	(0.00412)
SIP State				0.00753**
				(0.00373)
Wald Chi^2	257.49	291.38	295.64	387.06
Pseudo R^2	0.035	0.040	0.040	0.049

Robust Standard Errors in Parenthesis: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. All models contain controls for monitor latitude, longitude (linear and quadratic forms). n = 24,996. Marginal effects reported. Dependent Variable: 1 if Dropped O_3 Monitor, 0 if Active Monitor in Subsequent Period.

Table 4: The decision to add a new monitor: 1994–2010.

	(1.n)	(2.n)	(3.n)	(4.n)
Dropped Monitor, t-1	0.123***	0.123***	0.123***	0.121***
	(0.00551)	(0.00551)	(0.00552)	(0.00553)
Year	-0.00219***	-0.00304***	-0.00414***	-0.00420***
	(0.000395)	(0.000456)	(0.000766)	(0.000772)
O_3 max, t-1 (ppb)	0.000361**	0.000215	0.000227	0.000365**
	(0.000141)	(0.000149)	(0.000149)	(0.000145)
O_3 mean, t-1 (ppb)	0.000435	0.000655**	0.000654**	0.000444
	(0.000311)	(0.000313)	(0.000313)	(0.000316)
O_3 sd, t-1 (ppb)	-0.00166**	-0.00244***	-0.00235***	-0.00269***
	(0.000788)	(0.000822)	(0.000834)	(0.000838)
Share Missing Obs.	0.118	0.124	0.130	0.118
from Damage	(0.137)	(0.137)	(0.136)	(0.133)
Share Missing Obs.	0.0402	0.0471	0.0460	0.0424
from Malfunction	(0.0389)	(0.0387)	(0.0388)	(0.0398)
Share Missing Obs.	-0.174**	-0.206**	-0.205**	-0.129
from Maintenance	(0.0873)	(0.0889)	(0.0884)	(0.0861)
Share Missing Obs.	-0.625*	-0.596*	-0.563	-0.523
from Weather	(0.363)	(0.348)	(0.352)	(0.349)
Age	3.38e-05	7.46e-06	3.79e-06	4.18e-06
	(2.78e-05)	(3.09e-05)	(3.14e-05)	(3.18e-05)
Sum of O_3		0.000418***	0.000416***	0.000443***
Non-Attainment Years		(0.000128)	(0.000127)	(0.000126)
Sum of $PM_{2.5}$		0.000158	0.000148	0.000166
Non-Attainment Years		(0.000114)	(0.000114)	(0.000116)
O_3 NAAQs			-0.000446	0.000616
Revision			(0.0101)	(0.0102)
PM NAAQs			0.0117	0.0121
Revision (2006)			(0.00967)	(0.00970)
PM NAAQs			0.00820	0.00769
Revision (2009)			(0.00746)	(0.00741)
SIP State				0.0211***
				(0.00551)
Wald Chi^2	1,012.83	1,041.78	1,054.34	1,136.03
Pseudo R^2	0.203	0.206	0.207	0.213

Robust Standard Errors in Parenthesis: *** p<0.01, ** p<0.05, * p<0.1. All models contain controls for monitor latitude, longitude (linear forms). n = 12,824. Marginal effects reported. Dependent Variable: 1 if New O_3 Monitor, 0 if Active Monitor in Previous period.

Table 5: The decision to add a new monitor: Sample restricted to areas with and without concurrent dropped monitor in SIP/FIP states.

		Drop = 0	
VARIABLES	All States	SIP	FIP
O_3 max, t-1 (ppb)	0.000349***	0.000356**	-1.70e-05
	(0.000130)	(0.000139)	(0.000316)
O_3 mean, t-1 (ppb)	0.000494*	0.000573*	0.000855
	(0.000279)	(0.000321)	(0.000646)
O_3 sd, t-1 (ppb)	-0.00317***	-0.00324***	-0.00315*
	(0.000757)	(0.000849)	(0.00165)
Sum of O_3	0.000378***	0.000184	0.00111***
Non-Attainment Years	(0.000115)	(0.000126)	(0.000305)
Sum of $PM_{2.5}$	0.000236**	0.000137	0.000326
Non-Attainment Years	(0.000100)	(0.000120)	(0.000200)
O_3 NAAQs	-0.0109	-0.0147	0.00149
Revision	(0.0105)	(0.0115)	(0.0248)
PM NAAQs	0.0190*	0.0116	0.0433*
Revision (2006)	(0.0101)	(0.0110)	(0.0231)
PM NAAQs	0.00877	0.00456	0.0137
Revision (2009)	(0.00640)	(0.00694)	(0.0150)
Observations	11,984	8,836	3,148
		Drop = 1	
	All States	SIP	FIP
O_3 max, t-1 (ppb)	-0.0190***	-0.0184**	-0.0173
	(0.00613)	(0.00735)	(0.0115)
O_3 mean, t-1 (ppb)	-0.000704	-0.000169	-0.00331
	(0.00125)	(0.00155)	(0.00282)
O_3 sd, t-1 (ppb)	0.000563	-0.00268	0.00516
	(0.00264)	(0.00391)	(0.00414)
Sum of O_3	0.000805	0.000171	0.00103
Non-Attainment Years	(0.000974)	(0.00120)	(0.00201)
Sum of $PM_{2.5}$	-0.00184**	-0.000517	-0.00311***
Non-Attainment Years	(0.000902)	(0.00138)	(0.00118)
O_3 NAAQs	0.0940	0.0623	0.101
Revision	(0.0688)	(0.0821)	(0.123)
PM NAAQs	-0.0582	-0.0662	0.00385
Revision (2006)	(0.0683)	(0.0803)	(0.119)
PM NAAQs	0.0341	-0.0256	0.0945
Revision (2009)	(0.0612)	(0.0778)	(0.100)
Observations	840	533	304

Robust Standard Errors in Parenthesis: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. All models contain controls for monitor latitude, longitude (linear forms). Marginal effects reported. Dependent Variable: 1 if New O_3 Monitor, 0 if Active Monitor in previous period.

Table 6: Alternative Measures of NAAQs Attainment.

Attainment Measure	Poisson Count	Drop Monitors	Add Monitors
Sum of O_3	0.0009	-0.000114	0.000443***
Non-Attainment Years	(0.0045)	(7.66e-05)	(0.000126)
Sum of $PM_{2.5}$	-0.0013	0.000326***	0.000166
Non-Attainment Years	(0.0011)	(5.36e-05)	(0.000116)
One year lag O_3	-0.0037	0.00224	0.0040*
Attainment Status	(0.0070)	(0.0014)	(0.0024)
One year lag $PM_{2.5}$	0.0022	0.0074***	0.0032
Attainment Status	(0.0139)	(0.0013)	(0.0031)

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: The decision to add monitors in counties without existing stations.

	No New Counties	With New Counties	With New Counties SIP States	With New Counties FIP States
Sum of O_3	0.0004***	-0.00104***	-0.00131***	-0.000701**
Non-Attainment Years	(0.0001)	(0.000151)	(0.000179)	(0.000330)
Sum of $PM_{2.5}$	0.0002	0.000261*	6.45e-05	0.000387
Non-Attainment Years	(0.0001)	(0.000153)	(0.000209)	(0.000246)
O_3 NAAQs	0.0006	-0.00385	-0.00659	0.00661
Revision	(0.0102)	(0.0131)	(0.0149)	(0.0275)
PM NAAQs	0.0121	-0.000684	-0.00883	0.0184
Revision (2006)	(0.0097)	(0.0129)	(0.0148)	(0.0265)
PM NAAQs	0.0077	0.0301***	0.0273***	0.0369**
Revision (2009)	(0.0074)	(0.00922)	(0.0106)	(0.0188)
SIP	0.0211***	0.1985***		
	(0.0055)	(0.0060)		
Observations	12,284	13,318	9,787	3,531
Pseudo R^2	0.104	0.059		0.112

Robust Standard Errors in Parenthesis: *** p<0.01, ** p<0.05, * p<0.1. All models contain controls for monitor latitude, longitude (linear forms), population, and Year. Marginal effects reported. Dependent Variable: 1 if New O_3 Monitor, 0 if Active Monitor in previous period.

Figures

Figure 1: National O_3 Monitor Count.

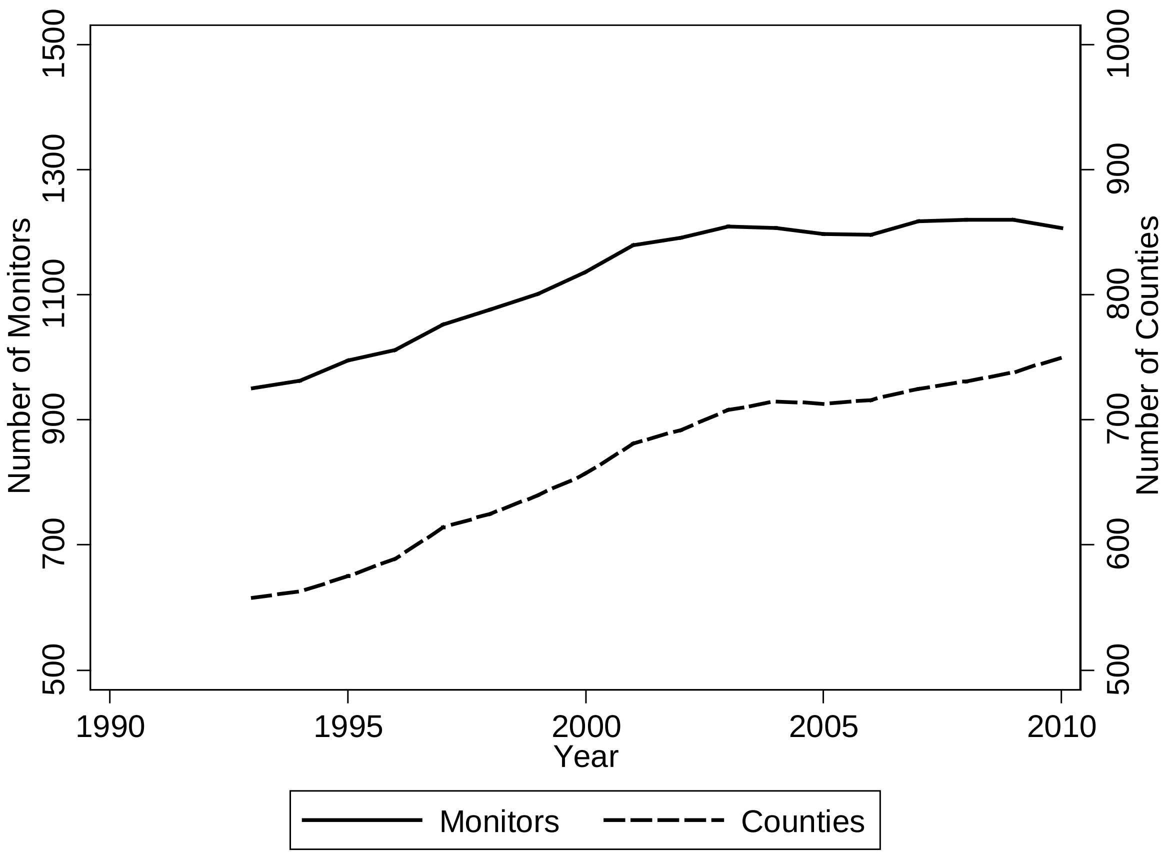
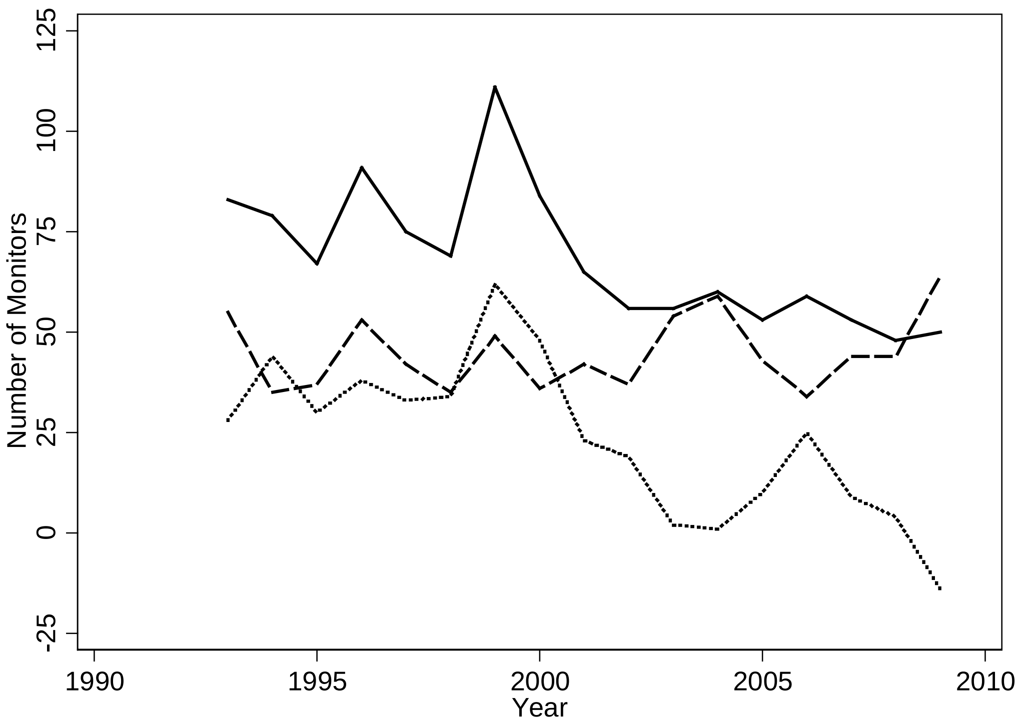
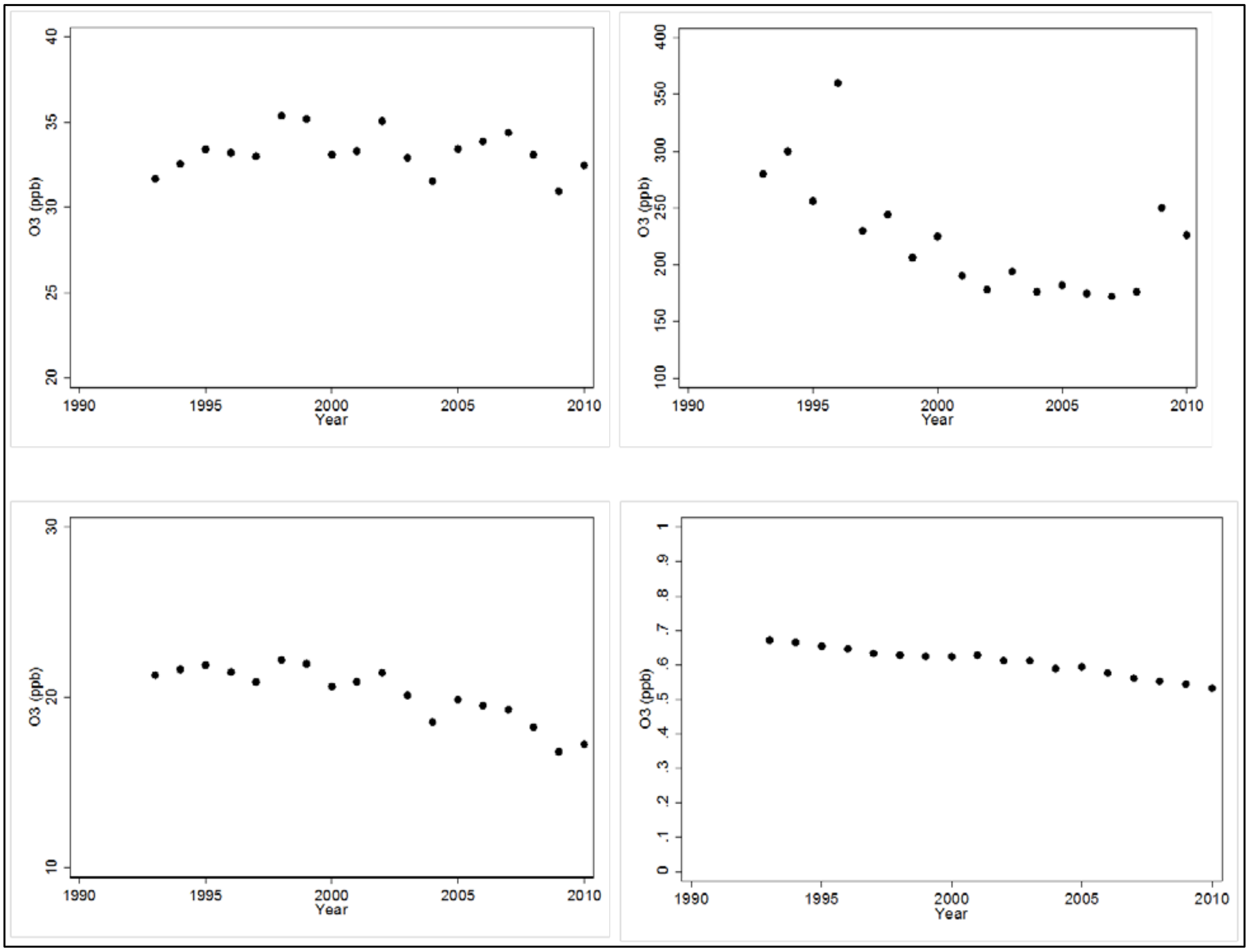


Figure 2: Changes to the Monitor Network (1993-2010).



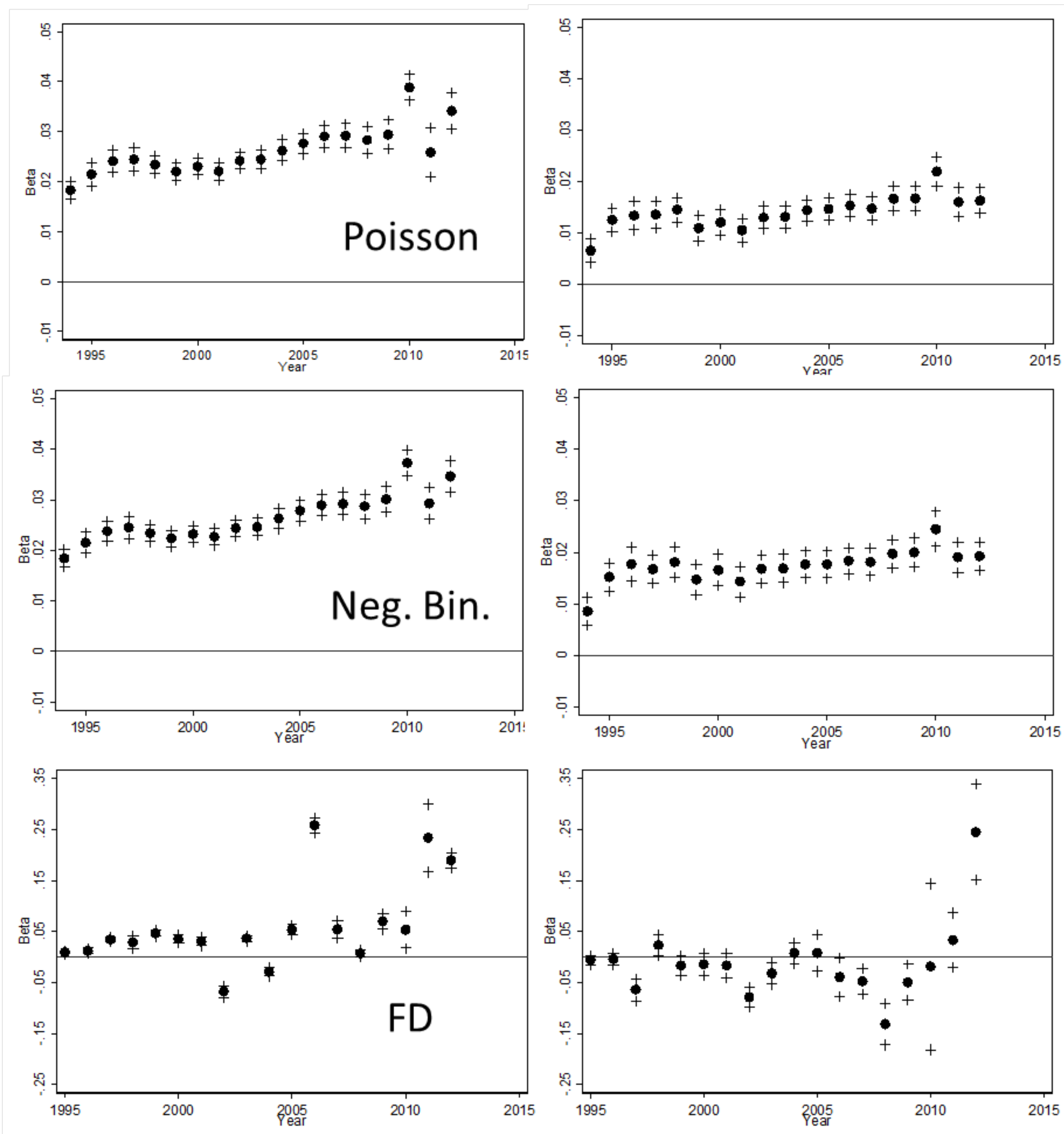
Dot: (New – Dropped Sites), Solid: New Sites, Dash: Dropped Sites.

Figure 3: National Average O_3 Readings.



Top left: Arithmetic mean. Top right: Hourly maximum. Bottom left: Standard deviation. Bottom right: Coefficient of Variation.

Figure 4: Cross sectional regression coefficients for maximum O_3 readings in monitor count models.



Left hand panel corresponds to county monitor count models. Right hand panel corresponds to distance monitor counts. Dots are the mean parameter estimate for O_3 maximum readings. Pluses indicate 95% confidence intervals. Dependent variable in all models: monitor count. Results presented from fitted model 4.c.

References.

Adams, R.M., Glyer, J.D., Johnson, S.L., McCarl, B.A., 1989. A reassessment of the economic effects of ozone on United States agriculture. *Journal of the Air Pollution Control Association* 39: 960-968.

Andrews, D.W.K. "Chi-Square Diagnostic Tests for Econometric Models: Theory," Econometrica, Vol. 56, No. 6 (Nov., 1988), pp. 1419-1453

Baldauf, R.W., D.D. Lane, and G.A. Marote. 2001. "Ambient air quality monitoring network design for assessing human health impacts from exposures to airborne contaminants." *Environmental Monitoring and Assessment.* 66: 63-76.

Baldauf, R.W., D.D. Lane, G.A. Marotz, H.W. Barkman, T. Pierce. 2002. "Application of a risk assessment based approach to designing ambient air quality monitoring networks for evaluating non-cancer health impacts." *Environmental Monitoring and Assessment.* 78: 213–227.

Bell, Michelle L., Adrian McDermott, Scott L. Zeger, Jonathan M. Samet, and Francesca Domenici. 2004. "Ozone and Short-Term Mortality in 95 US Urban Communities, 1987-2000." *J. of the Amer. Medical Association,* 17: 2372-2378.

Chang, N.B., C.C. Tseng, 1997. Optimal design of a multipollutant air quality monitoring network in a metropolitan region using Kaoshung, Taiwan as an example. *Environmental Modeling and Assessment*. 57: 121-148.

Hoel, M. 1997. "Environmental Policy with Endogenous Plant Locations." *Scandanavian Journal of Economics.* 99(2): 241-259.

Jerrett, M., R.T. Burnett, C.A. Pope, K. Ito, G. Thurston, D. Krewski, Y. Shi, E. Calle, and M. Thun. 2009. "Long-term Ozone Exposure and Mortality." *The New England Journal of Medicine.* 360: 1085- 1095.

Kainuma, Y., K. Shiozawa, and S. Okamoto. 1990. "Study of the optimal allocation of ambient air monitoring stations." *Atmospheric Environment.* 24B(3): 395–406.

Lesser, V.M., J.O. Rawlings, S.E. Spruill, M.C. Somerville. 1990 . Ozone Effects on Agricultural Crops: Statistical Methodologies and Estimated Dose-Response Relationships.Crop Science. 30, pp. 148-155.

Liu, M.K., J. Avrin, R.I. Pollack, J.V. Behar, and J.L. McElroy. 1986. “Methodology for designing air quality monitoring networks.” *Environmental Monitoring and Assessment.* 6: 1–11.

Markusen, J.R., E.R. Morey, N. Olewiler. 1995. “Competition in regional environmental policies when plant locations are endogenous.” *Journal of Public Economics.* 5: 55–77.

Modak, P.M., B.N. Lohani, 1985. “Optimization of Ambient Air Quality Monitoring Networks.” *Environmental Monitoring and Assessment.* 5: 1–19.

Muller, Nicholas Z., Robert Mendelsohn, and William D. Nordhaus. 2011. “Environmental Accounting for Pollution in the United States Economy.“ *American Economic Review.* 101: 1649-1675.

Oates, W., R. Schwab. 1996. "The Theory of Regulatory Federalism: The Case of Environmental Management," in W. Oates, The Economics of Environmental Regulation (Aldershot, U.K.: Edward Elgar. 1996), pp. 319-331.

United States. Federal Register. 76 FR 48207.

United States. United States Code, 42 USC § 7410.

United States. Code of Federal Regulations, 40 C.F.R. § 58.

United States Environmental Protection Agency. 1999. *The Benefits and Costs of the Clean Air Act: 1990--2010. USEPA* Report to Congress. USEPA 410-R-99-001, Office of Air and Radiation, Office of Policy, Washington, D.C.

United States Environmental Protection Agency. 2010a. *The Benefits and Costs of the Clean Air Act: 1990--2020. USEPA* Report to Congress. Office of Air and Radiation, Office of Policy, Washington, D.C.
http://www.epa.gov/air/sect812/aug10/fullreport.pdf

United States Environmental Protection Agency. 2013. Final Report: Integrated Science Assessment of Ozone and Related Photochemical Oxidants. http://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=247492

United States Environmental Protection Agency. 2016.
http://www.epa.gov/aboutepa/our-mission-and-what-we-do.

Venegas, L.E., N.A. Mazzeo. 2003. "Design methodology for background air pollution monitoring site selection in an urban area." *Intl. Journal of Environment and Pollution*. 20(1-2): 185–195.

World Health Organization (WHO). 1977. "Air Monitoring Programme Design for Urban and Industrial Areas." Global Environmental Monitoring System. WHO Offset Publication No., 38.

Appendix.

Figure A-1: Map of U.S. ambient O3 monitoring sites in operation during 2006-2010 in *Welfare Risk and Exposure Assessment for Ozone: Final*

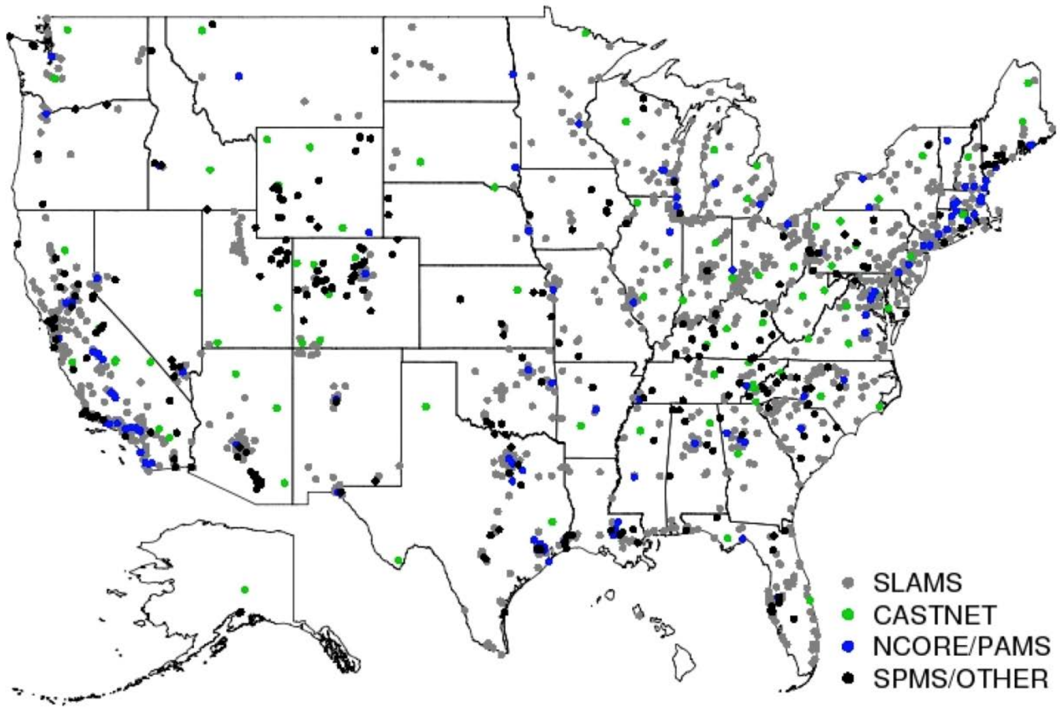


Figure A-2. Summary of Existing PAMS Site Locations

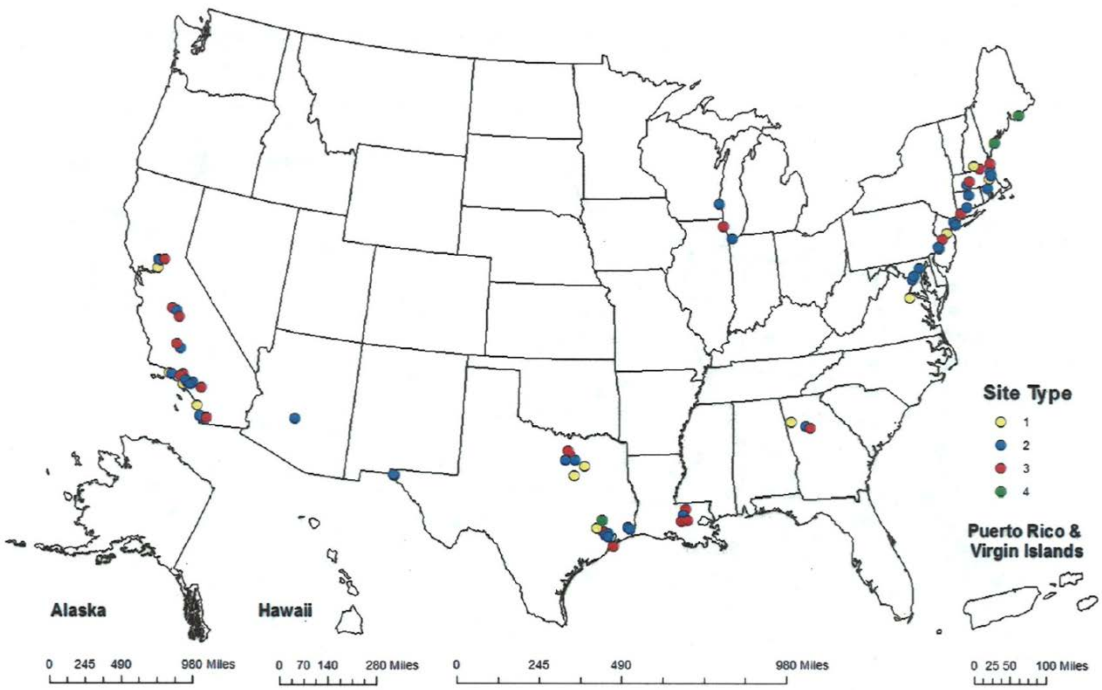


Figure A-3: CASTNET Sites Operational During 2012

Figure A-4: O_3 Levels by Monitor Ownership

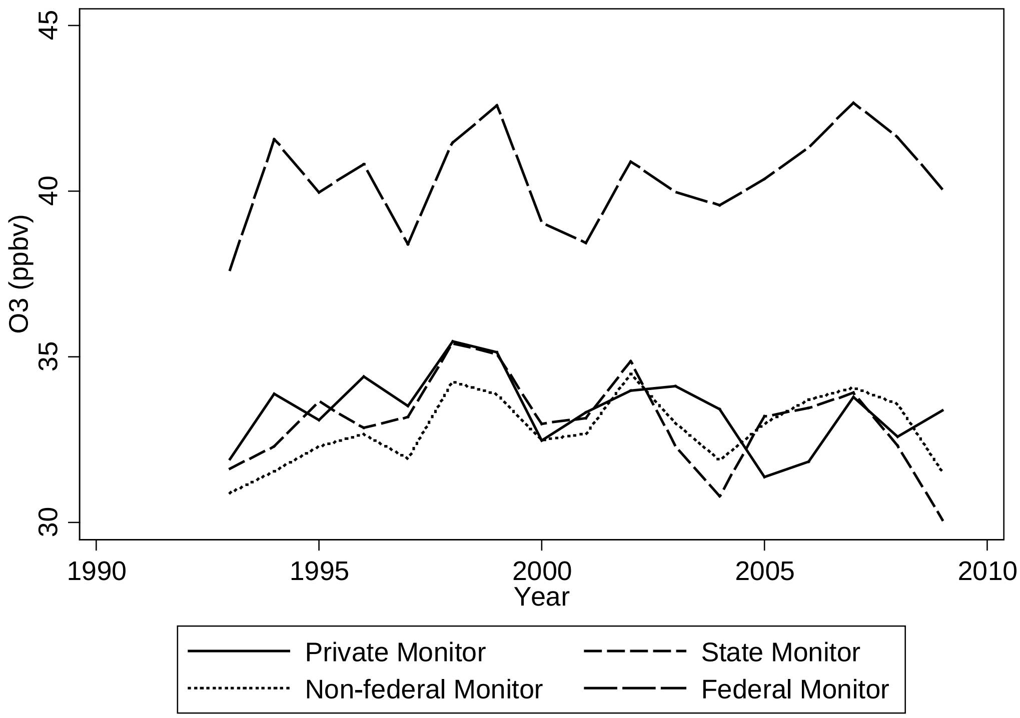
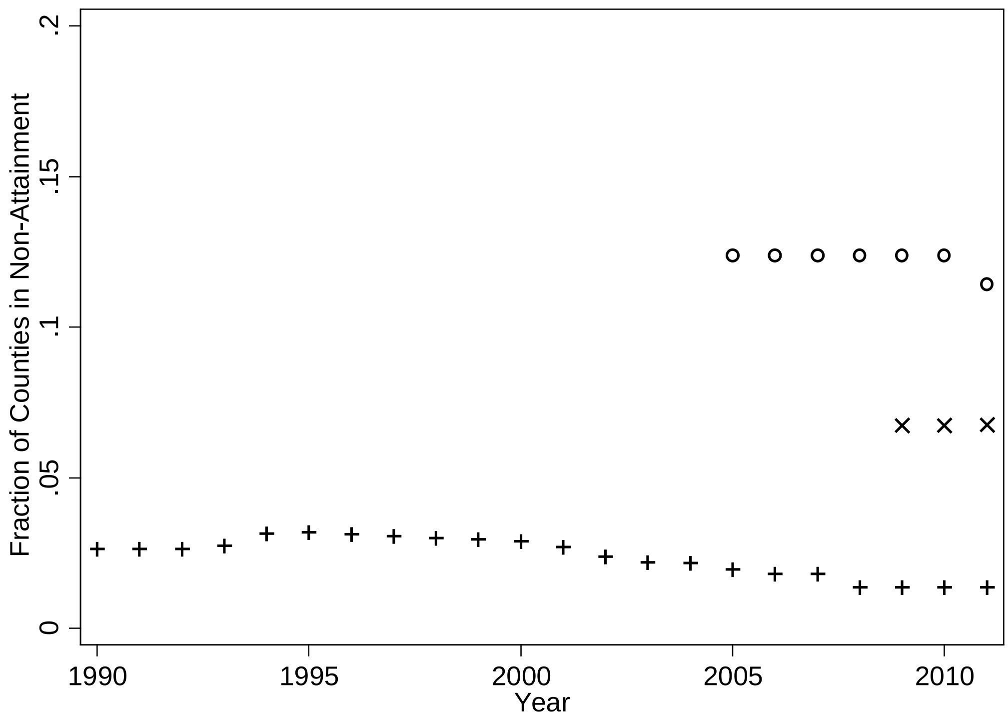


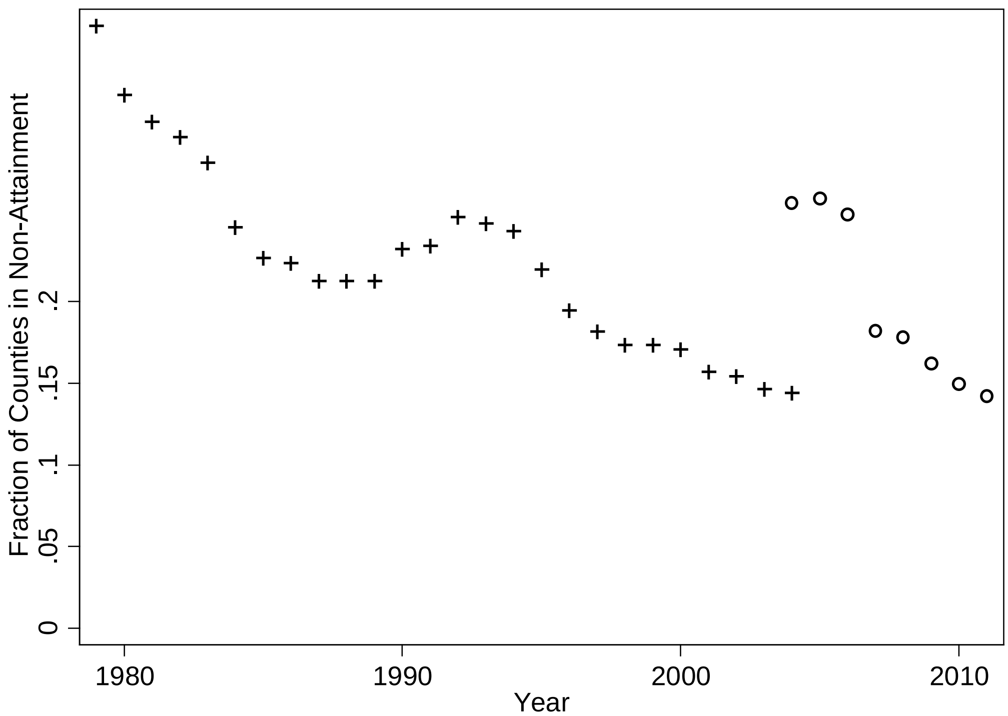
Figure A-5: PM NAAQs Attainment Status.



+ = PM_{10} NAAQs, o = $PM_{2.5}$ 1997 NAAQs, x = $PM_{2.5}$ 2006 NAAQs.

Source: http://www.epa.gov/oaqps001/greenbk/data_download.html

Figure A-6: O_3 NAAQs Attainment Status



+ = 1 Hour O_3 NAAQs, o = 8 Hour O_3 NAAQs 1997.

Source: http://www.epa.gov/oaqps001/greenbk/data_download.html

Tables.

Table A-1.

MSA population	Most recent 3-year design value concentrations ≥85% of any Ozone NAAQS	Most recent 3-year design value concentrations <85% of any Ozone NAAQS
>10 million	4	2
4-10 million	3	1
350,000-<4 million	2	1
50,000-<350,000	1	0

Source: Table D-2 of 40 CFR Part 58, Appendix D**:** Network Design Criteria for Ambient Air Quality Monitoring, Section 4 - Pollutant-Specific Design Criteria for SLAMS Sites.

Table A-2. Summary Statistics.

Variable	Mean	Std. Dev.	Min	Max
Drop	0.033	0.179	0	1
O_3 mean (ppb)	33.475	7.639	0.86	68.68
O_3 max (ppb)	103.10	22.562	21.4	256
O_3 sd (ppb)	18.147	4.111	3.078	42.36
Year	2002.8	4.543	1995	2010
Missing	0.032	0.058	0	1
Damage	0.000	0.006	0	0.68
Malfunction	0.015	0.048	0	1
Maintenance	0.014	0.028	0	1
Weather	0.001	0.008	0	0.39
Anamoly	0.003	0.010	0	0.64
Private	0.013	0.112	0	1
State	0.673	0.469	0	1
County	0.107	0.309	0	1
EPA	0.004	0.066	0	1
Other	0.170	0.376	0	1
Attainment Status (t-1)	0.474	0.538	0	2
Attainment Status (t-2)	0.484	0.539	0	2
Latitude	37.183	4.771	25.39	48.95
Longitude	-94.43	16.430	-124.63	-67.06
Age	14.59	10.03	0	51
Pop	694810	1267741	741	9830420
NAAQs 1997	0.892	0.310	0	1
NAAQs 2008	0.200	0.400	0	1

Table A-3. O_3 Levels by Monitor "Ownership" Type.

	Mean O_3 (ppb)			
Years	Federal Monitors	State Monitors	City, County, or District Monitors	Private Monitor
1993 /1994	37.69 (0.051)	31.71 (0.014)	30.85 (0.022)	31.99 (0.046)
1994 /1995	41.90 (0.054)	32.29 (0.010)	31.46 (0.022)	34.25 (0.056)
1995 /1996	40.31 (0.059)	33.70 (0.014)	32.33 (0.022)	33.13 (0.055)
1996 /1997	41.08 (0.053)	33.08 (0.013)	32.52 (0.022)	34.52 (0.057)
1997 /1998	38.61 (0.050)	33.36 (0.013)	31.65 (0.020)	33.70 (0.057)
1998 /1999	42.30 (0.052)	35.71 (0.013)	34.08 (0.021)	35.21 (0.070)
1999 /2000	42.99 (0.053)	35.32 (0.013)	33.87 (0.020)	34.41 (0.068)
2000 /2001	39.01 (0.051)	33.15 (0.012)	32.29 (0.019)	32.86 (0.068)
2001 /2002	38.88 (0.049)	33.32 (0.012)	32.64 (0.019)	33.79 (0.065)
2002 /2003	41.64 (0.053)	34.92 (0.012)	34.51 (0.019)	35.00 (0.069)
2003 /2004	39.80 (0.044)	30.80 (0.010)	31.96 (0.017)	32.98 (0.065)
2004 /2005	39.49 (0.042)	30.80 (0.010)	31.92 (0.017)	32.70 (0.064)
2005 /2006	40.37 (0.044)	33.21 (0.011)	32.98 (0.018)	31.37 (0.065)
2006 /2007	41.32 (0.043)	33.46 (0.010)	33.71 (0.018)	31.85 (0.064)
2007 /2008	42.67 (0.039)	33.92 (0.010)	34.08 (0.017)	33.80 (0.066)
2008 /2009	41.62 (0.039)	32.32 (0.010)	33.58 (0.017)	32.58 (0.058)
2009 /2010	40.06 (0.036)	30.09 (0.009)	31.53 (0.016)	33.39 (0.059)