ABSTRACT

Title of Dissertation: ASSESSING THE EFFICACY OF NPDES REGULATION: PERMIT WRITERS' DECISIONS, PLANTS' RESPONSES, AND IMPACT OF POLLUTANTS ON WATER QUALITY

Lopamudra Chakraborti, Doctor of Philosophy, 2008

Directed By: Professor Kenneth E. McConnell
Department of Agriculture and Resource Economics

This study provides indirect evidence that the Clean Water Act (CWA), implemented through the National Pollution Discharge Elimination System (NPDES) regulation, has been effective in terms of inducing certain 'best practice' responses from the different 'agents' involved in water pollution and its regulation. Given that cost benefit analyses of the CWA have yielded favorable conclusions, the chapters in this dissertation collect empirical evidence on whether NPDES permit writers pay attention to downstream water quality, if plants are sensitive to ambient pollution, and finally if pollutant discharges have an impact on downstream quality. Previous empirical studies incorporating ambient water quality in effluent limit or abatement choice, or pollutant inputs as a determinant of downstream water quality could not be found. These intermediate relationships are studied with Biochemical Oxygen Demand (BOD) as the primary pollutant and hence Dissolved Oxygen (DO) as the main indicator of water quality. Monthly panel data comprising a sample of 100 plants from Maryland, Virginia and Pennsylvania and 79 pairs of (downstream and upstream) water quality monitoring stations over a period of about 14 years, from 1990 to February 2004, was used. Positive evidence on efficacy of the NPDES regulation is found in all the three aspects investigated. On how regulation is implemented: if average water quality prevailing during past permit cycle is increased by one percent, then limits on BOD concentration (quantity) in the 'new' cycle would be made less stringent by 0.617 (0.322) percent. On how polluters respond to downstream water quality: if average DO prevailing during past three years is reduced by one percent, then concentration (quantity) discharges relative to effluent limits is reduced by 1.301 (1.558) percent. Finally, on how pollutant discharges from point sources have an impact on ambient water quality: if sum of BOD concentration is increased by one mg/L, then downstream net of upstream DO is reduced by 0.005 mg/L.

ASSESSING THE EFFICACY OF NPDES REGULATION: PERMIT WRITERS' DECISIONS, PLANTS' RESPONSES, AND IMPACT OF POLLUTANTS ON WATER QUALITY

By

Lopamudra Chakraborti

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2008

Advisory Committee:
Professor Kenneth E. McConnell, Chair
Professor John K. Horowitz
Professor Erik Lichtenberg
Professor Charles Towe
Professor Maureen Cropper

Dedication

To my parents

Acknowledgements

I am greatly indebted to Professor Kenneth McConnell, my advisor. He has provided me with excellent advice and consistent guidance, to say the least, from the time of prospectus development through the completion of this dissertation. I cannot thank him enough for his incredible patience and valuable time that he spent through the years. He has always made me aware of the fundamentals on how to make sound arguments to the range of complexities in econometric issues. I am also indebted to Professor John Horowitz who has immensely helped me during the initial and middle stages of development of my thesis. Between the two of them, they have often shared and exchanged the roles of incomparable educators as they introduced me to graduate level research. I am also grateful to Dr. Lichtenberg, Dr. Towe and Dr. Cropper who gladly accepted to serve on my committee.

I am thankful for the generous advise and feedback provided by several officials with the MDE, MDDNR, VADEQ, and the EPA. To name a few, Curtis Dalton, Roger Stewart, Bill Romano, Kyle Winter, Sherman Garrison, Ed Partington, Emery Harriston, and Shane Knipschild.

I have also received generous help and advise from professional staff in my department. Among them are: Jeff Cunningham, Barbara Burdick, Chuck McCormick, Katherine Faulkner, Jane Doyle, Curtis Henry, Linda Wilkinson, and Chris Aggour.

Last, but not the least, I thank my family for providing me with immense motivation and support since childhood. I am also indebted to Martin, my husband, without his unconditional help, patience and understanding, during the arduous stages of my thesis writing, I would not have been able to achieve my goal.

Table of Contents

List of Tables

Appendix Tables

List of Figures

Chapter 1:Assessing the Efficacy of NPDES Regulation: Permit Writers' Decisions, Plants' Responses, and Impact of Pollutants on Water Quality

Introduction

The purpose of this study is to investigate the efficacy of water pollution regulation in the US with the foreknowledge that cost benefit analyses of the Clean Water Act (CWA) have generally been 'positive'.[1] In order to control point source discharges, National Pollution Discharge Elimination System (NPDES) permits are issued by state or federal governments, which specifies limits either on quantity, and/or concentration of Biochemical Oxygen Demand (BOD)[2] in the effluents discharged by industrial facilities as well as wastewater treatment plants (also known as Publicly Owned Treatment Works (POTWs)).

Studies such as the one conducted for the EPA in 2000 address questions on whether the substantial investments on water pollution abatement have yielded results in terms of improved water quality.[3] The findings of this report were: between 1968 and 1996, effluent discharge of BOD5[4] from POTWs reduced by about 45%. On the outcomes side, comparing the "worst-case" dissolved oxygen (DO) level *before* (1961-

---

[1] For example, two reports by the EPA estimated (separately) costs and benefits of the CWA to be around (approximately) $14.1 billion, in 1997 (USEPA, 2000c), while benefits amounted to about $11 billion per year in the mid-1990s (USEPA, 2000a).

[2] Biochemical or biological oxygen demand (BOD) refers to the same chemical procedure of biological organisms using up dissolved oxygen.

[3] "From 1970 to 1995, EPA provided $61.1 billion in Federal Construction Grants Program funds to help fund new or upgrade existing POTWs" (USEPA, 2000b).

[4] The BOD-5 day test measures the amount of DO consumed by the decomposition of carbonaceous and nitrogenous matter in a sample of the wastewater (under laboratory conditions e.g. $20^{\circ}C$) over a five-day period. It has a detection limit of 1 mg/L. However, according to more recent studies such as Gray (1999) the BOD- 5 day test measures only carbonaceous oxidation.

1970) and *after* (1986-1995) the CWA revealed that for 214 out of the 311 reaches[5], below POTW outfalls, DO levels had improved, while the number of reaches with less than 5 mg/L[6] decreased from 167 to 97.

This study examines if the NPDES regulation induces certain 'best practice' responses from the different 'agents' concerned with water pollution and its abatement. In particular, whether permit writers implementing regulatory policies, and behavioral responses of the polluters (POTW and industries), the latter bearing its physical impact (of pollutant discharges) on ambient water quality, have not lost sight of the goal of the CWA. In other words, do they appear to pay attention to downstream[7] water quality[8] when making their (daily) monthly average abatement and permitting decisions (usually) spanning for 5 years, and if water quality responds to lower discharges by the point sources? Hence, effectiveness of regulation is assessed by collecting evidence on these intermediate relationships with the hindsight that findings from various reports of the EPA have shown that water quality has improved since the inception of the CWA.

The three causal linkages, that are studied, are spelled out in this paragraph. The first linkage examines the regulator's decision i.e. if water quality prevailing in the past permit cycle affects effluent limit chosen for the current cycle (Chapter 2). The second relation investigated is that between abatement decisions of a plant and lagged ambient water quality (Chapter 3). Specifically, does water quality prevailing at a past time

---

[5] Average length of these reaches was 10 miles.

[6] To maintain aquatic life a minimum DO standard of 5 mg/L is required at any time for all surface waters (in the US).

[7] Downstream ambient water quality refers to DO measured at a location that is downstream to the point(s) of effluents discharged by one (or more) major point source polluter(s).

[8] The term (ambient) water quality is used synonymously with in-stream dissolved oxygen (DO) throughout this study even though DO is only one dimension of assessment of ambient water quality. However, observed DO still serves as a good indicator of overall water quality since it is directly related to many stream uses such as fishable and swimable.

period and at a location downstream to the point of outfall of a plant determine how much pollutant it discharges in its ‘receiving’ water? This effect would be induced by regulation. And finally, the technical effect of whether contemporaneous pollutant discharges (at all) affects ambient water quality (Chapter 4).

Conventional pollutants have been the focus of most regulatory efforts (Helland 1998; Magat and Viscusi 1990) specially since toxics such as heavy metals are attributed almost solely to industrial activities. By contrast, manufacturing and sewage treatment plants (as well as agricultural farms and urban runoffs) discharge these pollutants. Among the five conventional pollutants[9], BOD is chosen since data records of monthly average pollutant discharges, tracked by the EPA and states, are most abundant. The pollutant that has been studied most, concurrently with BOD, is Total Suspended Solids (TSS). Majority, if not all, of these previous studies have found empirical evidence of similarities in pollutant discharge behavior (etc) of these two pollutants. However, monthly records of TSS from centralized databases are not as frequent as BOD (Bandyopadhyay, 2002). Predominant source of other common pollutants namely nitrogen and phosphorus as non-point source run-offs (apart from POTWs) is a well-established fact. Moreover, during the time period of this study, nitrogen discharges from point sources were not regulated. Data on these two pollutants are just beginning to be formally collected and recorded from the relevant point sources (and non-point sources that are relatively easy to monitor such as feed lots). BOD has also been found to be representative of other pollutants to a reasonable extent; in particular, technology for control of BOD is linked with reductions in nitrogen pollution.

---

[9] According to 40 CFR (Code of Federal Regulations) §401.16, section 304(a)(4) of the CWA: these are BOD, pH, fecal coliform, TSS, oil and grease.

The CWA requires the EPA and authorized states to: 1) develop water quality standards for all surface waters; 2) monitor these waters; and 3) identify and list those waters not meeting water quality standards. A water quality standard is the combination of its designated use and the water quality criteria designed to protect that use.[10] Once a designated use has been assigned to a water body the ‘anti degradation’ policy provides three tiers of protection of water quality. Protect existing uses, protect fishable and swim-able uses for higher water quality water bodies and prevent increase in discharges that affect the water quality of water bodies with exceptional natural resource and or wildlife significance.

BOD from the effluent discharges of point sources (e.g. pulp and paper mills and sewage treatment plants) is the pollutant directly linked to reduction of ambient DO, apart from nutrient pollution from sewage treatment plants and farms etc. The Streeter-Phelps (S-P) equation discussed in Chapter 4 outlines the process of evolution of downstream DO due to BOD discharges in the effluents of point sources. Plants that are regulated by monthly average limits on BOD5 that can be discharged in their effluents are required to (self) report their monthly average discharge of BOD in the Discharge Monitoring Reports (DMR) filed with the authorized regulator.

---

[10] Interaction between the socio-political environment and the economic ability of a certain region to sustain or improve its water quality, along with various mechanisms of discussion such as public hearings, legislation and evaluation are considered in the ‘use attainability’ analysis.

Data

The current sample of 100 (26 from MD, 22 from PA, and 52 from VA) 'major' NPDES plants is tracked for a time period of about 14 years from January 1990 to February 2004, on a monthly basis. The EPA's Permit Compliance System (PCS) database provides effluent discharge data from the monthly DMRs filed by major sewage treatment and manufacturing plants that face a BOD5 limit in terms of concentration (in mg/L) and/or in quantity (in lbs/day). Major municipal dischargers include all facilities with design flows greater than one million gallons per day, or facilities serving populations greater than 10,000, or facilities with EPA/State approved industrial pretreatment programs i.e. they receive industrial process wastewater. Major industrial facilities are determined based on specific ratings criteria developed by the EPA or the authorized State (USEPA, 1996).[11] Monthly data on pollutant discharge, effluent limits, limits start date, inspections, design flow and type of plant as captured by sic code (54 sewage treatment, 4 public sector and 42 manufacturing plants) was obtained by a FOIA submitted to the EPA.[12]

There are 86 (94) plants with BOD5 concentration (quantity) monthly discharge data. These plants do not comprise the entire universe of major polluters classified under the NPDES program across the three states of MD, PA and VA. It is a subset of the major plants, facing BOD5 limits, discharging into a free flowing stream or river rather than a lake or the Chesapeake Bay. However, some of the plants included in the sample

---

[11] Classification of an industrial plant as a major generally involves consideration of factors relating to the significance of the discharger's impact on the environment, such as: nature and quantity of pollutants, character and assimilative capacity of the receiving waters, presence of toxics, and compliance history (NPDES Overview, Indiana Department of Environmental Management (IDEM) Online).

[12] Information of different components of data available from the PCS database can be accessed from the documentation on PCS User Support (USEPA, 2001a).

either discharge into the Delaware River (which is classified as an estuary in certain locations) or near the mouth of a river. The nearest downstream (to the plant's discharge point) water quality monitoring station with dissolved oxygen data, for these and all other cases, are identified using EPA's (GIS) BASINS (Better Assessment Science Integrating point and Nonpoint Sources) software (USEPA, 2001b), which integrates the locational information of the major point source polluters tracked by the EPA's PCS database and the ambient water quality monitoring stations (tracked by STORET).

EPA's central database STORET is the primary source for ambient water quality data. This database incorporates data from an intricate network of monitoring stations on almost every stream and river segment mapped in the reach files network (RF1) (the latter is also digitized by the USGS). Monthly dissolved oxygen data for the period 1990 to 1999 can be obtained from the Legacy Data Center (LDC), which contains historical water quality data under the STORET system. Subsequently, monthly DO data for the monitoring stations in Maryland (MD) are obtained from the Chesapeake Bay Monitoring Program (CBP). For Virginia (VA), monthly data for the entire time period was obtained from a VADEQ official (Roger, Stewart; Water Quality Assessment and Planning Scientist). This data is also available on VADEQ's website. For Pennsylvania (PA), STORET was the only source for monthly records of DO data for the time period considered.

In the current sample, there are 79 pairs of downstream and upstream stations with data on monthly dissolved oxygen over the approximately 14-year period. A unique pair of upstream and downstream stations could be identified for 59 of the 97[13] "major" manufacturing and sewage treatment facilities sampled. For the remaining 38, 26 of

---

[13] Three plants did not information on design effluent flow.

them had one other major facility discharging into the same stream segment and hence they had the same pair of upstream and downstream stations. The last 12 plants had two other plants polluting into the same stream segment i.e. there were three plants discharging “in between” the same pair of upstream and downstream monitoring stations.

The same sample of plants and monitors are considered for the empirical analysis in all the three models (outlined below). Specifically, the two months data on effluent limits from the last year in the sample (2004) could not be dropped from the estimation sample considering that even after calculating seasonal averages[14] for each plant specific permit cycle, did not leave one with a large sample size since permits were “generally” issued with a 5 year time span. For the abatement and water quality models, monthly data is used, and hence to control for annual trends (if any) the two months of the last year are not included.

The Permits Decision of the Regulator

The causal relationship examined in this model is that of the level of ambient water quality prevailing during the preceding, plant-specific permit-cycle in the effluent limits choice of the regulator, for the subsequent cycle. In the 1990s, water pollution regulation was believed to be mostly inflexible with the uniform requirement of plants meeting their corresponding secondary treatment standard; thereby, leaving no room for incorporating ambient water quality in permitting decisions.

---

[14] Some plants in the sample faced seasonal permits which is a Water Quality Based Effluent Limit (WQBEL), since it meant that they could discharge a lower amount of pollutant during the critical summer conditions when the “assimilative capacity” of streams was at its lowest.

For sewage treatment plants, technology based standards for BOD discharge means that they have to meet the secondary treatment requirement of polluting no more than 30 mg/L in their effluent wastewater on a monthly average basis. For industrial plants, Best Professional Judgment (BPJ) is exercised to determine the relevant technology-based limits. Broadly speaking, the effluent limitation guidelines (ELG) for manufacturing facilities specify that the performance i.e. magnitude of abatement of pollution has to be compared with the best existing performance within that industrial (sub) category, and the cost of a plant to reduce pollution by a certain amount has to be compared with that of a sewage treatment plant (with similar design flow e.g.).

The objective of modeling the permitting decision is to test whether regulators have responded to downstream ambient water quality while choosing limits for a plant. For monitoring locations that meet ambient standards, a positive influence of past cycle average water quality is expected since regulators want to make sure that ambient standards will not be violated in the (near) future, at the same time as possessing the knowledge that abatement is costly to plants (POTWs and industrial firms alike).

The Abatement Choice of the Polluter

The purpose of modeling this behavior is to find evidence on the second causal relationship hypothesized; namely, during the time period of this current study, did the polluters pay attention to downstream ambient water quality when deciding on their (daily) monthly average abatement decisions? Induced by regulation to ‘observe’ the downstream water quality in order to avoid harsher effluent limits (and consequent penalties of non-compliance) or driven by (environmental) camaraderie, a positive finding can be interpreted as the permittees acting efficaciously since they adjust to ambient water quality even before it can has been formally incorporated in their permitting system. It will also provide an additional explanation for the plant operators' and managers' decision to overcomply when it has been shown that abatement is costly (in particular, marginal cost of abatement is non-zero or positive).

For almost the entire time period of 1990 to 2003, in many parts of the US, water pollution due to elevated levels of nitrogen had become an increasing environmental concern. Nutrient pollution from non-point sources such as agriculture and livestock, not to mention point sources such as sewage treatment plants, led to the pervasive problem of alarmingly high levels of dissolved nitrogen and phosphorus in the ambient waters. This led to excessive algal growth, which thwarted aquatic life by consuming high amounts of DO. At that time there was no federal statutory requirement to control nitrogen discharges from even the point sources, similar to the BOD discharges, which also used up ambient DO.[15] As a result, one would have witnessed a downward pressure on water quality, despite the 'good' performance of the point sources in terms of controlling BOD in their effluents much below the permitted limits. Polluting facilities in turn would have to either 'accommodate' this stress on water quality by increasing their efforts to reduce their pollution of BOD, or more explicitly 'compensate' for it in the situation that its effluent limits of BOD were made more stringent. Since these WQBELs were based on non-Total Maximum Daily Load (TMDL) mechanism for allocation of pollution burden non-point sources of pollution were effectively 'left out' of the regulatory realm.

---

[15] Most treatment plants did not face either a formal limit built in a TMDL for its' receiving sub-watershed, or were given financial incentives to adopt nitrogen control technology.

TMDLs emerged as the future of water pollution regulation in the new millennium, since the regulators were obligated to amend local water quality problems.[16] Section 303(d) of the CWA directs States to identify and list waters known as water quality limited segments (WQLSs), in which the current (required) controls of a specified pollutant are inadequate to achieve water quality standards. For each WQLS, the State is required to establish a TMDL of the pollutant that the water body can receive without violating water quality standards. Consequently, the manufacturing and sewage treatment plants as well as the previously 'unregulated' nonpoint sources were legally enforced to be responsive to ambient water quality. In the absence of this explicit mechanism, the treatment plant operators and managers of industrial firms had to be more prudent in terms of being receptive to the status of ambient quality in the 'receiving' waters. In other words, plants might be responding to the same 'trigger' that could result in more stringent effluent limits as part of their strategic decision-making; thereby, exercising their foresight in incorporating factors in their abatement decisions that are central to the regulator's concern. Hence, lagged ambient water quality measured by concentration of DO is used as an explanatory factor in the abatement decision of a polluting plant. A positive direction of impact is anticipated since plants want to avoid expected costs of non-compliance and related intervention activities once a more stringent WQBEL is imposed in response to a poor status of downstream water quality.

---

[16] Local water quality concerns have resulted in state-level initiatives to control nitrogen discharges by requiring reductions via TMDLs (which would include the 'contribution' of non-point source pollution (NPS)) and providing incentives to encourage voluntary reductions. "Maryland, for example, encourages nitrogen reductions by providing financial incentives for plants seeking to upgrade their facilities via the provision of capital construction grants and low interest loans" (Akobundu 2004). More recently, incentives are giving way to mandatory requirements. 66 major treatment facilities have to achieve Enhanced Nutrient Removal (ENR) technology to meet concentrations of 3.0 mg/l (parts per million) or less total nitrogen, and 0.3 mg/l or less total phosphorus (by 2010, per their scheduled upgrades) under the Point Source Strategy (the limits for these plants were based on achieving the WQ standards in the Maryland Regulations for the Chesapeake Bay and its tributaries).

The Water Quality Observed Downstream

The third causal relationship examined is the ambient water quality observed as a result of the pollutants discharged by point sources. In particular, the analysis focuses on whether, and by how much, pollutant “inputs” from ‘major’ point sources have had a significant impact on downstream DO controlling for water quality prevailing at an upstream location i.e. before the point(s) of effluent outfall of the plants. Positive empirical evidence would be another yardstick to assess the effectiveness of water pollution policy since major point sources faced the brunt of the regulation during a time period when TMDLs were being formulated on paper. In other words, if BOD discharges from point sources *did* have a significant impact on net downstream DO, then policy implemented through the NPDES permits (to which plants respond while making their discharge decisions) *did* have an impact on its target i.e. ambient water quality. A negative impact of BOD discharges from one or more[17] plants is expected since the physical impact of BOD is to consume dissolved oxygen, by definition (the S-P model).

Given that the S-P equation cannot be generalized i.e. its coefficients cannot be estimated across distinct monitoring locations, a simple linear model[18] is specified where sum of BOD concentration, from all the relevant polluters, is the primary explanatory variable. S-P tells us that the impact of BOD discharge is maximum when the pollutant has traveled a certain distance downstream and not at the point of effluent outfall; beyond this trough the impact of BOD reduces and DO increases. In the absence of knowledge

---

[17] Availability of monthly water quality data downstream (and upstream) to each plant-monitoring station pair was a constraint.

[18] Besides, past empirical studies have resorted to linear, multivariate analysis; given that DO is distributed (close to) normal.

on the exact distances, defining the ‘impact zone’ for BOD (on DO), which are specific to each pair of monitors and plants, some candidates for upper and lower ranges is tried within which a significant impact of BOD can be felt. Subsequently, the coefficient for the ‘impact range’ is compared with the non-linear (NL) estimates of a modified S-P equation (depending on data availability of the different components of the original S-P).

Concentration rather than BOD loads are considered relevant for purposes of assessing the impact of discharges on downstream water quality in the S-P framework. A case in point: numerous TMDLs are assigned, in the US, where the water quality models predict the concentration of BOD that will be necessary to meet the ambient standard. Specifically, given that background i.e. upstream pollution is zero under “7Q10”[19] conditions, BOD in the river/wastewater mixture is solely represented by effluent concentration. The corresponding wasteload allocations were made based on the design effluent flow of plants.

Has Regulation Been Effective?

In retrospect, water quality was, in general, good for the sampled locations. 3-5 year average DO was around 9.5 mg/L, much higher than the minimum 4-6 mg/L standard for aquatic life. This provides an appropriate background for the underlying framework of the three aspects investigated. Positive evidence on efficacy of the water pollution regulation is found in all the three aspects of how regulation is implemented,

---

[19] 7Q10 is defined as the critical stream flow; it is the minimum 7 consecutive day average stream flow that has a recurrence interval of 10 years.

how polluters respond (to downstream water quality controlling for effluent limits), and finally how point sources' decision to abate have an impact on ambient water quality.

For limits on BOD concentration, if mean (median) water quality prevailing over the entire span of past cycle is increased by one mg/L, then permit levels in the 'new' cycle would be made less stringent by 1.328 (1.176) mg/L. For quantity permits, the coefficient is 46.356 and 38.593 lbs/day depending on whether the mean or median value for preceding cycle water quality is considered. In terms of percentages, the estimated influence of water quality for concentration and quantity permits is 0.617 (0.556) and 0.322 (0.268) percent, respectively. Hence, some evidence is found on the regulators utilizing limits on both effluent concentration and loadings as policy instruments, in particular, to insure that the ambient water quality goals are met. Additionally, inference can be drawn on past cycle water quality assuming greater importance when deciding on the concentration effluent limits.

On an average, past period water quality seems to have a significant influence on current period (monthly) abatement decision. The results (for the entire sample) show that if past three years mean (median) DO is reduced by one percent, then relative concentration discharges is reduced by 1.301 (1.257) percent, while relative quantity discharges is reduced by 1.558 (1.657) percent. Upon restricting the sample to at least 50 percent of the total monthly observations, the estimates for relative concentration discharges is larger at 2.001 (1.903) percent, while coefficients for relative quantity discharges is marginally smaller at 1.395 (1.593) percent.

Sum of BOD5 discharges by major point sources seem to exert a significantly negative impact on the change in downstream DO from that of an upstream location. The coefficient obtained from the best candidate for capturing the impact of BOD is the range of greater than 2 and less than 26 miles. If sum of BOD5 concentration is increased by one mg/L, then downstream net of upstream DO is reduced by 0.005 mg/L (significant at 5%). Despite the ad-hoc methodology utilized, the estimated impact of sum of BOD5 concentration is ‘comparable’ (albeit less in magnitude), when compared with the NL estimates of the “partial” S-P equation (with coefficients around –0.008 to –0.011 mg/L). Hence, policy is found to have an impact on its target (ambient water quality), indirectly, through the significant impact of abatement behavior of plants on downstream DO.

Chapter 2:A Model of Permit Setting

Introduction

The objective of this chapter is to answer the question: to what extent do regulators take into account ambient water quality when deciding on the maximum amount of a pollutant that plants should be allowed to discharge in their effluents. Specifically, water quality measured at locations downstream to a plant's point of effluent discharge (outfall) during past permit cycles is anticipated to play a role in the permits chosen for the current cycle. This observation is based on the notion that preventing deterioration of the nation's water quality is the quintessential objective of water pollution regulation i.e. the CWA.

The States are authorized to designate uses for all waterbodies within their boundaries. The corresponding ambient quality standards that are required to meet these uses are identified from the scientific literature. For example, if maintaining aquatic life is identified as the use for a certain reach then the underlying ambient water quality standard of 4-5 mg/L at any point in time is implied irrespective of the state in which the waterbody is located in.

Code Of Maryland Regulations (COMAR) classifies most of its waterbodies according to eight designated uses.[1] Hence, protection and propagation of aquatic life is an implicit component of all the designated uses listed above. However, the requirement

---

[1] Use I and I-P: Water Contact Recreation, and Protection of Nontidal Warmwater Aquatic Life and Public Water Supply,
Use II and II-P: Support of Estuarine and Marine Aquatic Life and Shellfish Harvesting and Tidal Fresh Water Estuary,
Use III and III-P: Nontidal Cold Water Aquatic Life and Public Water Supply, and
Use IV and IV-P: Recreational Trout Waters and Public Water Supply.

that “[t]he dissolved oxygen concentration may not be less than 5 milligrams/liter at any time” (http://www.dsd.state.md.us/comar/26/26.08.02.03%2D3.htm) for Use I waters e.g. originates from the scientific records of sustenance of life in water if DO is higher than this standard, and is not at the discretion of the state regulators.

Operating under the overarching rule of anti degradation, once the uses have been chosen, the state/local regulators have discretion in choosing effluent limits for point sources such that the specific categories of uses are met by each river or stream segment. The EPA centralizes all self-reported discharge data (submitted by the authorized states) and monitoring activities (both federal and state level) in the Permit Compliance System database. The federal regulator has discretion in conducting inspections and subsequent enforcement actions based on the data submitted by the state regulators. However, the state regulators themselves undertake almost 90% of the monitoring activities (as seen in the inspections data).

Downstream[2] water quality prevailing during past permit cycle is considered as the relevant determinant for the permit choice decision. When the regulator reaches a decision on the effluent limit for a pollutant he/she, most likely, considers the ambient water quality in the past months/years of the preceding permit cycle. On the other hand, water quality prevailing in the ‘same’ (current) cycle is endogenous to the permitting and pollutant discharge behavior of the regulator and the plant, respectively. Current ambient water quality depends on upstream water quality, and the plant’s decision to abate, which in turn is influenced by the effluent limit assigned to the plant. However, water quality observed in the past permit cycle is exogenous to the regulator’s choice of limits in the subsequent cycle, and hence is considered as the appropriate explanatory variable.

---

[2] See Chapter 4 for details on how downstream monitoring stations are identified.

For purposes of investigating the efficacy of regulation, in ensuring that the goals of the CWA continue to be met, a crucial aspect is whether the National Pollution Discharge Elimination System (NPDES) permits, issued or renewed (during the time period of this analysis) by the state regulators, were technology or water quality based. As mentioned in Chapter 1, technology based limits for sewage treatment plants meant that they could discharge no more than the secondary treatment standard of 30 mg/L of BOD on a monthly average basis. On the other hand (as the name suggests), water quality based limits are determined by local water quality considerations. In particular, their objective is to ensure that the ambient standard required for maintaining the designated use of a stream segment will not be violated. Hence, evidence on the influence of ambient water quality in the effluent limit(s) chosen by a regulator, during a phase when technology based permits was apparently the common practice, is an interesting linkage to examine.

During the 1990s, permitting decisions were largely perceived to be inflexible to factors determining local costs and benefits of water pollution regulation. Consequently, it was common knowledge that permits chosen by the regulator did not incorporate ambient water quality, preference for environmental quality, or abatement costs. As Keplinger (2003) pointed out, the NPDES permits "focused on the technical feasibility of achieving effluent standards as the primary factor in establishing pollution control in contrast to the former clean water legislation, with its emphasis on state-administered ambient standards" (p.1058). On a similar note, an editorial in the Yale Law Journal (1982) emphasized that reliance on technology based standards meant that a cost-effective allocation of the pollution burden was not being achieved "because the impact of a given discharge depends on meteorological factors, topography, and other physical parameters specific to the discharge points" (p.798).[3]

The above discussion provides the main motivation for modeling the permitting process: to test whether regulation has been efficacious in the context of using tools such as effluent limits to ensure that ambient water quality standards are maintained. Simply expressed, to examine whether permit writers were paying attention to ambient water quality when making their permit issue or renewal decisions. Presently, there are very few empirical studies on the regulator's choice of permit levels, and the couple that are there (discussed in the next section) utilize cross-section data; thereby, leaving the question of whether permit writers incorporate water quality prevailing in previous permitting cycles in their limits choice, for the next cycle, unanswered.

As will be seen later, evidence is found on the practice of permit writers assigning BOD5 limits that were water quality based, but were derived from non-TMDL[4] (Total Maximum Daily Load) based waste load allocations (WLA). Given the technology-based effluent limits of a plant, a dissolved oxygen sag analysis was conducted, i.e., the lowest concentration of ambient DO under critical low flow conditions was simulated. If the DO level generated did not meet the ambient standard required to meet the designated use of the stream, water quality based limits were invoked. Presently, the regulators have greater access to actual water quality data for review due to the increased frequency and consistent monitoring of water quality at different locations. As a result, even for BOD,

---

[3] Secondly, in the long run, technology-based standards were identified as providing the weakest incentives for innovations in abatement technology. Plants have a strong disincentive to adopt expensive, but more efficient technologies because this provides a justification to the regulators to impose, in all likelihood, more stringent pollution control standards i.e. permits.

[4] Termed so by the Virginia Department of Environmental Quality, VADEQ.

TMDL based water quality limits are starting to be implemented.[5] This chapter presents evidence that both concentration and quantity permits are utilized as instruments by the regulators to insure that ambient standards required for existing uses continue to be met.

Review of the Literature

McConnell and Schwarz (1992) is one of the very few studies that model the permitting decision of the regulator for sewage treatment plants. They estimate the actual and design effluent concentration (i.e., the permit level) choices made by the regulator as a system of equations. The design effluent is instrumented in the actual effluent choice model. They use a non-linear, instrumental variable, three-stage least-squares estimation method. The design effluent concentration is determined by the factors laid out in the following reduced form equation.

$$\begin{aligned}\tilde{E}_d &= \delta_0 + \delta_1 \tilde{F}_d + \delta_2 \tilde{F}_a + \delta_3 \tilde{I}_d + \delta_4 \tilde{I}_a + \delta_T' \tilde{Z}_T + \delta_u' \tilde{Z}_u + \delta_6 Y \\ &\quad + \delta_7 \left(\overline{C}_k + \ln\left(1 + \varphi e^{(\tilde{C}_o - \tilde{C}_k)}\right)\right) + e \end{aligned} \tag{2.1}$$

where a ~ denotes natural log, and

$$\overline{C}_k = \theta_0 + \theta_2 \tilde{I}_d + \theta_3 \tilde{F}_d + \theta_k' \tilde{Z}_k$$

The actual effluent concentration is determined by (the following technological constraint):

---

[5] In the absence of recorded water quality data the permit writer has to rely on the conventional evaluation, i.e., stringent permits are chosen only if the technology based limits are not enough to meet the 4-5.0 mg/L for DO under critical low flow (drought) conditions. Availability of measured water quality data means that she/he can verify whether ambient standards are actually violated by the water body at any point in time, especially during low flow seasons.

$$\tilde{E}_a = \lambda_0 + \lambda_1 \tilde{E}_d + \lambda_2 \tilde{F}_d + \lambda_3 \tilde{F}_a + \lambda_4 \tilde{I}_d + \lambda_5 \tilde{I}_a + \lambda_T' \tilde{Z}_T + \varepsilon_T \tag{2.2}$$

In equation (2.1), permits are a function of the design flow rate F_d, the actual flow rate F_a, the design BOD influent concentration $\tilde{I}_d$, the actual BOD influent concentration I_a, additional technological factors that eventually determine the amount of pollution discharged by a plant Z_T,[6] "exogenous" factors such as ambient water quality and the regions' concern for environmental matters Z_u, state income Y, annualized plant design and building costs C_k, annual O&M costs at the plant C_o, and the exogenous variables affecting plant construction costs Z_k. Actual effluent concentration depends on permit levels, design and actual flow, and influent concentration, and other technological factors affecting the actual amount of pollutant discharged by a plant.

Their (cross-section) sample of sewage treatment plants, across the U.S., have about 50% of the plants that face permitted concentrations of BOD5 lower than the secondary treatment standard (no higher than 30 mg/L, or 85% removal of BOD, or whichever is more stringent). McConnell and Schwarz note that data on plant costs were the most difficult to obtain. Ambient water quality was not included in their regressions even though it was considered as a factor that affects the regulator's taste for pollution control. "The NPDES-permitting system requires states to consider water quality when granting permits to wastewater treatment plants. However, in our model, water quality in the plant's region might appropriately be considered endogenous to the regulator's decision about plant effluent" (p.61). Additionally, due to the absence of consistent ambient water quality within even 20-50 miles of the plant locations factors such as mean

---

[6] For example, since presence of advanced nutrient removal technology may affect the actual effluent concentration of a plant a dummy variable was included in the set of technological constraints.

flow and velocity of the receiving water body were considered as proxies. Data on mean flow i.e. average rate of flow of the plant's receiving water body was obtained from Dewald et al (1985).

McConnell and Schwarz mention that greater mean flow or velocity allows the river to absorb pollution with more ease; thereby, implying that the damages from the same amount of a pollutant in a plant's effluents are reduced. Their results show that mean flow has a negative and significant effect (at 5%) on permits chosen. The estimated coefficient can be interpreted as: if mean flow of the receiving water body is increased by one percent, then effluent concentration permits would be reduced by 0.027 percent. Rise in mean flow by one unit implies that ambient water quality is higher, which in turn implies that the damages of pollution from the same amount of a pollutant in a plant's effluents are reduced. As a consequence, demand for pollution abatement is reduced implying that effluent limits can be made less stringent i.e. increased by a certain amount. Hence, one would expect a positive coefficient on the mean flow coefficient instead of the negative effect found by McConnell and Schwarz.[7]

Some of the other results are: less stringent permit levels are chosen if the costs of pollution control of a sewage treatment plant are higher. They find that if costs of abatement are increased by one percent design effluent levels are increased by 0.51 – 0.68 percent. More specifically, they find that construction costs play a significantly greater role as opposed to O&M costs in their design effluent choice model.

---

[7] Similarly, county level population density and population growth rates are expected to exert a downward pressure on ambient water quality. Ceteris paribus, if density is increased by one unit, then water quality is reduced by a certain amount, implying that the damages from the same amount of a pollutant in a plant's effluents is increased, which in turn means that demand for abatement is increased, leading to more stringent permit levels being chosen. Again, we expect the sign on state income and state environmental index to be negative implying that regulators in more "environmentally concerned" regions are obligated to set lower design effluent levels.

McConnell and Schwarz conclude that local regulators might be considering costs and benefits of pollution control in making their decisions about effluent limits, despite the prevalent perception of rigidity in the regulation. Their findings refute the common notion that design BOD effluent concentrations were determined exogenously, i.e., “federal regulations set minimum “secondary” treatment levels for BOD that give local officials little flexibility in choosing effluent quality” (p.55).

DeShazo and Lerner (2004) estimate an empirical model where each plant’s “effluent limit” is regressed on a set of plant, firm, industry, and local community characteristics. Plant specific discharge standards i.e. “effluent limits” refer to BOD and TSS (total suspended solids) loads rather than concentration of the pollutants in the effluent discharge.

$$LIMIT_i = f\begin{pmatrix} PLANT_i CHARS., FIRM_i CHARS., INDUSTRY_i CHARS., \\ COMMUNITY_i CHARS, POLITICAL_i CHARS., REGION_i CHARS. \end{pmatrix} \qquad (2.3)$$

The dependent variable is log (LIMIT). The controls for differences in technology-specific effluent limit guidelines and best available technologies across plants were: product type (pulp and paper), plant technologies (kraft, sulphite, deinked etc), and plant capacity. The following measures of firm size were considered separately: an indicator of whether the firm owned another plant (regulated by NPDES) in the same state, an indicator of whether the firm owned another plant in the U.S. regulated under the NPDES, and two continuous measures for firm size. The non-discrete measures were: interaction of the indicator of presence within the state with the log of the additional firm capacity in the state (omitting the capacity of that plant), and interaction of the indicator of national presence with the log of the additional capacity at the national level (omitting the capacity of that plant). Industry size was captured by the log of the total production capacity of the pulp and paper industry in a given state (excluding the capacity of the firm that owns that plant on the left hand side). The Herfindahl Index for the pulp and paper industry in a given state captured industry concentration (where the firm that owns the plant on the left-hand side was excluded from the calculation). Local conditions included were: the log of the population in the plant's county, the LCV (League of Conservation Voters) score served as a proxy for demand for environmental amenities in a given state, the log of surface water withdrawn from the county (in millions of gallons per day),[8] and the log of the total water area in the plant's state.[9] State primacy i.e. whether the State had the authority to administer NPDES permits, federal oversight, and their interaction, and regional dummies were the other characteristics included.

A cross section sample (around January 1995) of 219 major, pulp and paper plants (across the US) from the EPA's PCS database was considered. Equation (2.3) was estimated by OLS correcting for plant level heteroskedasticity. They did not use a panel-data approach stating unavailability of time series data on permit levels, which are revised every five years, and that plant and industry specific variables change very little over a ten year period. Since pulp and paper mills are significant dischargers of conventional pollutants, they report the results for the average monthly limit on BOD (and separately for TSS).

The size of water bodies in a state, indicating the state's assimilative capacity to absorb effluent discharges, was positive and significant in all specifications (at 5% and 10% level). Water withdrawn (for all uses) was negative and significant at the 1% level

---

[8] The higher the water withdrawal (for all the uses), the more the water is used for various purposes, and the larger is the damage caused by the effluents of a plant.

[9] A given amount of pollution will be more damaging the smaller the total water area. Higher log of water area implies less damage from the effluents and thus higher expected effluent limits.

in all models; thereby, indicating that higher water use implies lower allowable effluent limits. It's estimated coefficients ranged from 0.05 to 0.06% and -0.22 to -0.40% respectively. DeShazo and Lerner find that larger firms are able to secure significantly less stringent limits for their plants (14 to 43 percent less stringent than smaller firms). Industry concentration was negative and significant at the 1% level, while its square was positive and significant. The coefficient on the size of the rest of the industry in the state was negative and significant at the 1% confidence level, which indicates that the larger the size of rest of the industry, the more stringent the plant's limits will be. The coefficient on the size of the population in a given county was negative and significant (at the 1% level), which was interpreted as state regulators assigning more stringent limits in more populated areas.

Decker (2003) estimates a model on the permit approval process rather than examining the effluent limits chosen by the regulator. In particular, his focus was on the length of time it takes environmental agencies to issue permits for new industrial facilities under the NPDES program. Between 1990 and 1998, 68 major and 82 minor industrial facilities applied for a new permit from the six Midwestern states comprising EPA's Region V: Michigan, Illinois, Indiana, Ohio, Minnesota, and Wisconsin. The model is:

$$DUR = f\begin{pmatrix} MAJOR, ENFLAST3, TRI1, CHTRI, PART3350, RELUNEMP, \\ GREENRAT, TOURINC, FIRMREV, POPDEN, CONSERV, \\ ExQUAL, MAJOR \times GREENRAT \end{pmatrix} \tag{2.4}$$

The dependent variable measures the time lapse, in days, between permit application and approval. $MAJOR$ is a dummy variable which is equal to one if the new construction project is major. $ENFLAST3$ is the number of enforcement actions taken against all the other existing plants, located in the same state and owned by the same firm, for statutory violations under the Clean Air Act (CAA), CWA, and Resource Conservation and Recovery Act (RCRA) during the 3-year period prior to the permit issue date. $TRI1$ measures the thousands of tons of chemicals, listed under the Toxic Release Inventory (TRI), released by all the other plants, in the same state and owned by the same company, 1 year prior to permit application date. $CHTRI$ is the change (e.g. reductions) in the TRI levels. $PART3350$ is a dummy variable that equals one if the company in concern was a 33/50 Program participant. $RELUNEMP$ is the amount by which the percentage unemployed in the county exceeds the corresponding number for the state, during the year of application. $GREENRAT$ measures the membership rates (per 10,000 state residents) in the Sierra Club and the Natural Resource Defense Council. $TOURINC$ is the revenue generated from parks and related recreation tourist activities relative to the personal income of the state. $FIRMREV$ is the total revenue of the firm seeking the discharge permit for a plant during the year the permit was issued. $POPDEN$ is the population of the county where the new facility will be located relative to the land area. $CONSERV$ captures the percentage of Republican state voters in the most recent presidential election preceding a firm's permit application date. $ExQUAL$ is a control for state level differences, by capturing the state's expenditures on water quality.

Given the nature of the dependent variable, Decker employs duration analysis, in particular, he estimates equation (2.4) using Weibull MLE method (as well as OLS with ln(DUR) as the dependent variable). His findings (similar in both the estimation methods reported above) are discussed next. If the proposed facility is categorized as major, then it takes three times longer to get approval. As for environmental compliance of major dischargers, an additional violation at any one of the firm's existing plants will increase subsequent permitting time by about 81 days (for the average, median, permit). Decker interprets this result as one of the underlying incentives for firms to maintain a good compliance record (at all other facilities) such that the chances of minimizing regulatory red tape associated with obtaining new permits are greater. As expected, RELUNEMP, CONSERV and PART3350 have significant, negative impacts on days till permit is approved while TR1 lengthens permit times. Voluntary pollutant release information is interpreted as facilitating permit approval process. In the absence of previous compliance history for new plants, regulators turn to related environmental information that are likely to reveal the potential compliance behavior of the new facility, under the same ownership.

It is worth mentioning that ambient standards and hence ambient water quality are expected to play an even greater role in the permits choice decision of a regulator. In particular, water pollution regulation is entering a mature phase where all the sources of pollution will be incorporated within its rulemaking purview (Boyd, 2000). The number of TMDLs that are being formulated and implemented in the new millennium is on the rise; even though, pollution from point sources has been successfully controlled. On the other hand, the "contribution" of non-point sources has been climbing steadily, since it has gone effectively unregulated from the time of inception of the FWPCA. Boyd adds the caveat that even though this is the right direction for water pollution regulation to move; controlling pollution from non-point sources by implementing best management practices, and the equally tall order of successfully enforcing these rules might present significant challenges to the regulatory authorities.[10]

The next section describes the recommended practices of the permit writer. It is a reliable and practical starting point for investigating what factors aught to influence the regulator's choice of permit levels for pollutants discharged in a plant's effluents.

Review of the Permit Application and Approval Process

The EPA's 1996 NPDES Permit Writers' Manual, documents the description and guidance provided to regulators who determine the permit levels to be assigned to an individual plant. It outlines the factors that a regulator has to take into consideration while deciding on the appropriate permit level. The federal regulations contained in 40 CFR (Code of Federal Regulations) §122.21 require that applications for new discharges be made no later than 180 days before discharges actually begin. Alternatively, applications for permit renewals (i.e., for existing dischargers) must be made at least 180 days before the expiration of the existing NPDES permit.[11] For new dischargers, the description of the permit application process and most of the key pieces of information that go into the review and ultimate decision on permit levels translate into expected or anticipated terms. For instance, expected date of start of operation and volumes of

---

[10] Not to mention that TMDLs are developed after extensive water quality monitoring and modeling.

[11] According to USEPA (1996), "an expired NPDES permit remains in effect until the new permit is issued as long as the application for permit renewal was submitted on time and complete (per 40 CFR §122.21). However, if State law does not allow expired permits to remain in effect until a permit is reissued or if the permit application is not on time and complete, the facility is considered to be discharging without a permit from the time the permit expired until the effective date of the new permit" (p.37).

effluents to be discharged, etc., whereas other aspects such as treatment technologies and plant design would typically be already in place.

NPDES permits for BOD 5-day are issued with a 5-year cycle. Termination of the permits prior to their scheduled expiration date occurs in situations such as noncompliance with the conditions of the permit, misrepresentation or omission of relevant facts by the permittee, or endangering human health or the environment. Structural changes in the plant design or capacity, as and when, reported by the plant might also trigger “untimely” decisions of assigning new permit levels to the polluter (Bandyopadhyay and Horowitz, 2006). The permits for direct dischargers to surface waters (municipal and industrial) can be categorized into two basic types: technology-based and water quality based limits.

Generally, the NPDES discharge permits that are issued to municipal facilities limit the following water quality characteristics: BOD5, TSS, E. coli, total residual chlorine, dissolved oxygen, and pH (Curtis, H. Dalton, P.E., Maryland Department of the Environment (MDE), Water Management Administration, Wastewater Permits Program, Municipal Surface Discharge Permits Division (MSDPD); email communication).[12] Additional parameters such as ammonia nitrogen, total kjeldahl nitrogen (TKN), total nitrogen, and total phosphorus may also be limited depending on receiving stream conditions, or applicable regulation. For sewage treatment plants, technology based permits require meeting the secondary treatment standard of at least 85% removal of BOD, which ensures the minimum level of effluent quality that can be attained, given the existing technology. In terms of effluent limits on the concentration of pollutants in a

---

[12] Discussions in this sub-section draw heavily from the feedbacks provided by a Maryland official in the Municipal Surface Discharge Permits Division of the Wastewater Permits Program, and a Virginia official in the Division of Water Quality Programs, both of them permit writers themselves.

plant's wastewater discharges, this standard translates to an average for a 30-day period of *either* BOD of or CBOD of 25 mg/l, *and* TSS of 30 mg/l, *and* either no dissolved oxygen limit *or* a D.O. limit of 5.0 mg/l,[13] *and* no ammonia or TKN limits.[14] Consequently, if a municipal plant faces a BOD5 limit that is less than 30 mg/l it means that the receiving waters were determined to be water quality limited as opposed to "effluent limited" (Kyle I. Winter, P.E., Manager, Office of Water Permit Programs, Division of Water Quality Programs, VADEQ; email communication). WQLSs do not have sufficient waste load assimilative capacity to allow the discharge of secondary level treated wastewater.

Table 2.1: Pollutant Requirements for Technology-Based Standards

Pollutant	Concentration, mg/L (30-day average)
BOD or CBOD	30 or 25 and
TSS	30 and
DO	none or 5 and
Ammonia or TKN	none

Receiving-stream water quality evaluations are usually undertaken when a new permit cycle is about to begin, to determine the required level of treatment for BOD5, TKN, and ammonia. Computer models are used to estimate the impact of the wastewater discharge on the receiving stream's dissolved oxygen. Initially, secondary level treatment of BOD5 = 30 mg/l and TKN = 25 mg/l is assumed in the stream model input. If the model predicts that the receiving stream dissolved oxygen will remain above the stream standard then secondary limits are selected as permit limits. If the model predicts that the

---

[13] Permit limits for total residual chlorine, pH, and dissolved oxygen are set equal to the receiving stream water quality standard.

[14] The ammonia and nitrogen specifications for pollutant control were incorporated during more recent times (post mid nineties) when water quality problems were not being resolved despite good performance by the point source polluters.

receiving stream dissolved oxygen standard will be violated with secondary level treatment then stricter BOD5 and TKN limits are chosen. The permit writer has access to the engineering reports on the impact on ambient water quality when making her/his decision.

Specifically, BOD5 limits (as opposed to annual nutrient loading limits established through a TMDL evaluation) are developed to protect the receiving stream dissolved oxygen during design stream flow conditions (Curtis, H. Dalton). COMAR (Code of Maryland Regulations) 26.08.01.01.B (18) defines "design stream flow" as the minimum 7 consecutive day, average stream flow that has a recurrence interval of 10 years (7Q10). This is representative of a drought condition. Kyle I. Winters also mentioned that critical low flow conditions is used to simulate the worst water quality that could arise on account of a plant discharging its secondary standard of 30 mg/L.[15] If new information on critical low flow conditions becomes available to the regulator even during an ongoing permit cycle for a plant, the permit writer is obligated to undertake a DO sag or a TMDL analysis (whichever is relevant for the specific facility), to ensure that minimum standards are not violated at any point in time, as well as aquatic life can be maintained with a margin of safety in the near future. This is one of the reasons why some plants face a new permit level before its usual 5-year span expires, apart from more obvious reasons such as noncompliant behavior of plants.

---

[15] Critical conditions are important because they describe the factors that combine to cause a violation of water quality standards and will help in identifying the actions that may have to be undertaken to meet water quality standards. Essentially, they are a combination of environmental factors (e.g., flow, temperature, etc.) with an acceptably low frequency of occurrence, and in this sense it represents a reasonable “worst-case” scenario condition. Hence, the 7 day, 10-year low-flow (7Q10) design condition is often used as the critical condition because the ability of a water body to assimilate pollutants without exhibiting adverse impacts is at a minimum.

Seasonal limits (summer: low flow and high temperature, and winter: high flow and low temperature) are examples of permits that are more stringent than technology based limits. These are incorporated in the permits if water quality evaluation determines that secondary limits are not adequate during the "summer" period. The permittee is therefore not required to meet the more restrictive summertime based BOD5 (less than 30 mg/l) limit during wintertime, since a less stringent wintertime BOD5 limit is included in the permit. Due to lower stream temperature and higher flow conditions, the assimilative capacity of the receiving stream can accommodate the less stringent wintertime limits, normally set at a maximum allowable BOD5 of 30 mg/l (Nov. through April). For example, Akobundu (2004) reports that in a personal interview an official of the Municipal Permits Division at the MDE explained that seasonal limits in Maryland reflect the fact that water quality conditions generally deteriorate during the summer due to elevated levels of BOD.[16] In the dataset used by Akobundu, many MD plants faced seasonal BOD limits where the summer BOD limit is lower than the winter BOD limit. In particular, 42 plants in MD[17] had seasonal limits, where the winter limit was at 30mg/l and the summer limit ranged from 5mg/l to 26mg/l with most plants at summer limits between 10mg/l and 20mg/l.

Municipal discharge permits limit both the concentration and quantity (loading) of BOD5. According to the regulation 40 CFR §122.45(f), permit writers must apply the

---

[16] Excess production of algal blooms in the water due to nitrogen over-enrichment was implicated; along with conditions such as increased availability of sunlight, and warmer temperatures. When these blooms die, they deplete the dissolved oxygen in the water, increasing the oxygen demand.

[17] Chesapeake Bay Program (CBP) office's Point Source database used by Akobundu includes plants discharging at 1,000 gallons per day. This criterion is distinct from the restriction used in this current study, i.e. picking up only the "major" polluters (WWTPs and manufacturing plants). Major municipal dischargers are those facilities, which have design flows of greater than one million gallons per day, or facilities serving populations greater than 10,000, or facilities with EPA/State approved industrial pretreatment programs i.e. they receive industrial process wastewater (USEPA, 1996).

secondary treatment standards (of 30 mg/L etc) as mass-based limits using the design flow of the plant (USEPA, 1996). In other words, the design wastewater flow is used to establish the appropriate BOD5 loading that will protect the stream during design (critical) stream conditions. The relation between quantity and concentration based limits is:

Quantity based limit = Design Flow X Concentration based limit X Conversion Factor[18] (2.5)

For industrial dischargers, technology based permits are derived using national effluent limitations guidelines or best professional judgment (BPJ) on a case-by-case basis. This involves comparing the cost of an industry to reduce its pollutant discharge with the cost of a publicly owned sewage treatment plant for similar levels of reduction of a pollutant loading. Best existing performance, on an average, by plants within the same industrial category or subcategory are also considered while deriving the appropriate limit for a manufacturing plant.

For industrial dischargers too the regulations of 40 CFR §122.45(f)(1) "require that all permit limits, standards, or prohibitions be expressed in terms of mass units (e.g., pounds, kilograms, grams)" (USEPA, 1996). Exceptions include situations when limitations based on mass are infeasible because the mass or pollutant cannot be related to a measure of production. However, the appropriate limit must ensure that dilution will not be used as a substitute for treatment. A provision under 40 CFR §122.45(f)(2) allows a permit writer to express limits in additional units (e.g., concentration units).[19] Expressing limits in terms of concentration as well as mass encourages the proper

---

[18] The conversion factor when the concentration permit is expressed in mg/L is usually 8.34 with the units of (lbs/day)/(mgd)(mg/L).

[19] When limits are expressed in more than one unit the permittee must comply with both.

operation of a treatment facility at all times by discouraging the reduction in the level of treatment during low flow periods. NPDES permits focused on the technical feasibility of achieving effluent standards (Keplinger, 2003). The long-term average flow is utilized for the calculation of concentration limits based on the already assigned mass limits, since it reflects the range of concentrations that could be expected in a well-operated plant.

The regulator has to ensure that ambient standards will not be violated in the near future, even for the industrial plants, by conducting a DO sag analysis. A permit writer must first confirm through the water quality evaluation that ambient water quality conditions have sufficient waste load assimilative capacity before a technology based BOD5 permit limit can be (re-) established. Finally, unlike the treatment plants, an effluent limit lower than 30 mg/L is not indicative of a water quality based permit, since one of the most crucial factors that determine the permitted level chosen by the regulator is the type of production facility.[20]

All facilities applying for an individual NPDES permit submit general facility information including the Standard Industrial Classification (SIC) code, which captures the nature of the production process of each facility. Municipal Application

---

[20] To elaborate, the effluent limitation guidelines accessed from the Electronic Code of Federal Regulations at http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?c=ecfr;rgn=div5;view=text;node=40%3A28.0.1.1.14;idno=40;sid=2568c8aaa953f0203903ea71d58bc3da;cc=ecfr for subcategories under the Organic Chemicals, Plastics, and Synthetic Fibers, industrial category are discussed as examples. For instance, the technology based BPT effluent limitation for Thermosetting Resins, Specialty Organic Chemicals, and Bulk Organic Chemicals plants are above 30 mg/L, at 61, 45, and 34 mg/L. In this situation a BOD5 concentration limit of 30 mg/L, for a plant, might actually denote a more stringent water quality based effluent limit rather than the technology-based limit. Alternatively, plants within the subcategories of Rayon Fibers, Thermoplastic Resins, and Other Fibers already have technology based BPT limitations under 30 mg/L: 24, 24, and 18 mg/L respectively. Hence, in the absence of more plant specific information, if a plant in this category faced a BOD5 concentration limit under 30 mg/L, it would not necessarily imply existence of a water quality based limit. Finally, only one subcategory: Commodity Organic Chemicals has a BPT effluent limitation of 30 mg/L, identical to the uniform, secondary treatment standard of sewage treatment plants.

Requirements include submission of information on technology such as collection system types, description of treatment practices and plant design and schedule of improvements. Information on areas served, total population served, description of influent, including major industrial facilities discharging to the system, number of discharge points, total volume discharged, and receiving water body are also required to be submitted by the permittee. Non-municipal dischargers applying for an individual NPDES permit are required to submit additional detailed facility information, which can adequately characterize the nature and quantity of pollutants in their effluents and their impact on the receiving water.

The permit writer might also collect and review any additional background information on the facility.[21] In-house file information typically includes the current permit level, the fact sheet or statement of basis for the current permit, the Discharge Monitoring Reports (DMRs), the compliance inspection reports, and correspondence or information on changes in plant conditions or problems, and compliance issues.[22] Lastly, public participation activities such as providing public notices, collecting and responding to public comments, and holding public hearings (as necessary) are conducted.[23] Hence, one expects the demand for environmental quality of the region to play a role in the final permit level that is assigned by the regulatory agency.

---

[21] For example, a permit writer might wish to discuss compliance issues, changes, or history of complaints with compliance personnel who conducted previous inspections of the facility.

[22] In addition the regulator might take into consideration if the polluting facility is a chemical industry belonging to the group of TRI, or even related environmental permits that could provide site-specific background information about the types of pollutants and waste-streams at a facility, for example, hazardous wastes regulated by the RCRA permits and CAA permits.

[23] The public notice is the vehicle for informing all interested parties and members of the general public of the contents of a draft NPDES permit or of other significant actions with respect to a NPDES permit or permit application. Since the draft permit is usually submitted for public notice after it has undergone internal review by the regulatory agency that is issuing the permit, in situations where an evidentiary hearing is called, the permit writer's primary responsibility is to be thoroughly familiar with the technical basis for the permit conditions.

Finally, once a permit level has been issued to a plant, section 402(o) of the CWA and NPDES regulations under 40 CFR 122.44(I)(2), prohibit the relaxation of effluent limitations in reissued permits i.e. a modification of limits that are less stringent than the previously established levels (USEPA, 1996). Anti backsliding addresses two specific situations: (i) BPJ technology-based limits after less stringent effluent guidelines are promulgated later, and (ii) when a permittee seeks to relax limits based on State treatment, or water quality standards. Exceptions for BPJ based limits can be granted if substantial alterations or additions to a permitted facility are made, or new information that was not available at the time of permit issuance becomes available. In addition to applicable effluent limitations guidelines (ELGs), potential violations of water quality standards are taken into consideration before exceptions are deliberated. For limits based on State or water quality standards exceptions can be granted if the existing limit is based on TMDL or WLA, provided that water quality standards, including anti-degradation will not be violated.

Using the information on the various aspects of the permit issuance process, a permits model is estimated in the next section. It enables one to estimate (and assess) the impact of past cycle water quality on permits chosen in the subsequent cycle, controlling for other aspects that also determine effluent limits. These factors are: polluting-facility specific characteristics such as design flow (size of abatement technology) and type of plant (describing the production process), and location specific features such as socio-demographic characteristics of the surrounding population, and state level jurisdiction.

The Permits Model

In equation (2.6), the dependent variable is the permit level in cycle c and season s for plant i, P_{ics}. Each time period of this model is defined by permit cycle and season. A permit cycle is defined by the phasing out of an old permit, which is replaced by a new permit level that is assigned to a polluter. Under the NPDES program, plants discharging BOD5 in their effluents face limits on either concentration, or quantity, or both. The permit level for concentration and quantity of BOD5 are modeled separately.

$$P_{ics} = \alpha + \gamma DO_{j(c-1)s} + \tau designflow_{i(c-1)s} + \pi sic_i + \sum_{k=1}^{6} \lambda_{ik} sociodemographics_{ik} + \theta_r + \varepsilon_{ics} \tag{2.6}$$

The primary explanatory variable is the measure of downstream ambient water quality prevailing during the previous permit cycle $(c-1)$ and season s $DO_{j(c-1)s}$ at monitoring location/site j. The other determinant is plant size, which is the average design flow in the previous cycle $(c-1)$ and season s $designflow_{i(c-1)s}$. It captures the expected scale of the production or treatment process of each plant and hence is an upper limit to the volume of effluents that the plant is designed to process or generate. Additional factors/controls are: structural characteristics associated with the production process as captured by the sic code sic_i, time invariant socio-economic aspects based on the zip code of plant i,[24] and θ_r which captures the effect of differences in state regulations, where r denotes the state in which plant i is located.

---

[24] Socio-demographic and economic characteristics are mostly highly correlated with each other; hence their individual coefficients have to be interpreted with caution. Ceteris paribus, communities with higher level of environmental awareness are anticipated to have more stringent permit levels assigned to the plants, since the regulator faces a stricter audience in terms of exhibiting an active interest in activities such as assessing the appropriateness of the permit levels to be assigned to a polluter, not to mention the water quality itself. Median income and poverty (unemployment) rates etc. influence the level of environmental

The error term in cycle c and season s for plant i, ε_{ics}, is expected to exhibit temporal dependence, in particular, within the same plant and from one cycle to the next (by season). Serial correlation arises since other unobserved physical aspects, not included in the permits decision model, associated with the production and abatement or wastewater treatment processes are essentially captured by the error term.

We know that the permit level assigned to a plant reflects the current abatement/treatment technology e.g. the vintage of the plant and costs of abatement. Other factors related to the operation of a plant are also embedded in the permit levels. Essentially, information on plant specific factors used by the regulator at the time of deciding on the permit levels are not readily accessible to outside researchers e.g. in the form of a dataset compiled by the regulatory authorities. For instance, permit levels are applied to sewage treatment plants after considering technological/treatment practices such as collection system types, location specific factors such as areas served, total population served, number of industrial plants discharging into its systems and volume and type of influents. In addition, if plants submit information on scheduled structural and operational changes, then the new permit levels are expected to reflect these changes that will be incorporated in the plant's operation/treatment in the new cycle. In particular, for plants that do not report any changes in operational procedures or structural aspects, new permit levels are anticipated to be close to the levels prevailing in the previous cycle; consequently, giving rise to serial correlation of the error terms.

---

consciousness of the ruling political party and hence stringency of the state government in passing more environmentally favorable legislations. On the other hand, holding all other determinants constant, local populations with a higher proportion of workers employed by the manufacturing sector, i.e., locations with higher 'leverage' polluters means less pressure on the regulator to implement stricter permitted levels.

Typically, plant level dummy variables are included to control for unobserved heterogeneity arising mostly due to time invariant factors. Plant level fixed effects are however not included in this analysis, to control for plant specific characteristics such as the ones mentioned above. Design flow, i.e., the scale of abatement technology is perfectly correlated with plant level fixed effects implying that it is a plant specific attribute that is invariant across cycles and seasons for each plant in the current sample.

There are three channels of impact of past cycle water quality on effluent limits. One is when the permit writer does a water quality evaluation (also called an oxygen sag analysis) for a specific polluter before deciding on the revised permit level to be assigned in the new permit cycle.[25] This exercise is usually done for all new, or expanding (scheduled structural or operational changes in production and/or abatement), or existing dischargers applying for a permit renewal. The second channel is (if and) when new information regarding receiving stream quality becomes available to the permit writer (not necessarily coinciding with a plant's permit cycle), e.g., water quality collected since the last evaluation has implications for reductions in critical low flows, then permit level for existing dischargers are modified to incorporate the results of this updated evaluation.[26] [27] The above protocol remains effective even when the polluter has a clean record of environmental performance with respect to its pollutant discharges being well

---

[25] That is, whether it should continue to be the technology based limits, or if it should be made a more stringent to water quality based limits, or if the water quality based limit itself should be made more stringent.

[26] A recent trend (since the new millennium) has been to completely modify, even the methods of evaluation; namely, a TMDL analysis for the sub-basin is conducted instead of the DO sag analysis for a specific polluter.

[27] Ex-post this would mean that a new permit cycle for a plant becomes effective even though (per the 5-year norm) a revision of its effluent limits is not on the regulator's schedule.

within its effluent limit.[28] Thirdly, modifications to reissued permits (i.e. exceptions to anti-backsliding), in particular, less stringent ones can be granted for BPJ technology-based limits and State or water quality standards based limits under the overarching condition that ambient standards will not be violated.

A positive direction of impact of past period water quality on effluent limits is expected. All else constant, if past cycle ambient water quality is increased by one unit, in a stream segment that meets its ambient standard, then less stringent effluent limits could be chosen by the regulator in the current permit cycle (since abatement is costly to plants). Alternatively, if past cycle water quality is reduced by a certain amount, in a stream segment that meets its designated standard, the regulator might choose a more stringent i.e. lower effluent limits in the current permit cycle (e.g., a water quality evaluation reveals that under critical (stream) conditions ambient standard will not be met in the future).[29]

---

[28] Uses have to be attained, unless the regulator finds evidence of significant socio-economic impact on the affected community on account of the excessive financial burden on the polluters resulting from them having to meet the current water pollution standards. The Chesapeake Bay Use Attainability Analysis (UAA) (approved by the EPA in 2005) is a notable example of refining, in particular, infeasibility of meeting current designated uses based on natural, anthropogenic, and economics conditions.

[29] According to the provisions laid down in the regulations it could be argued that the direction of impact of past water quality on new permits might be negative. In particular, authorized agencies can conduct an “interim economic analysis” to determine appropriateness of the current designated use (for higher than existing use waterbodies). In the case that the regulator does not find a substantial and widespread economic and social impact (on the permittees and the affected community) arising out of the costs of pollution control needed to meet the more stringent effluent limits, the permitted levels of BOD might actually be lowered so that the higher water quality standards can be maintained. However, during the time period of this analysis non-TMDL WLAs (for BOD) was the reality with the overwhelming trend of stream segments projected to not meet their designated use (also their existing use of fishable and swimable) under critical stream flow conditions. Consequently, incidence of upgrading designated uses or refining existing uses for free-flowing, higher water quality bodies was practically nonexistent during the 1990s.

Data

This section describes the nature of the data that is used to construct the dependent and explanatory variables, and controls included in the permits model. For the current sample of 100 plants, the time period of about 14 years from January 1990 to February 2004 is categorized by plant specific permit cycles. In the initial years of this dataset, most plants were facing permit levels that were assigned to them in the late eighties (with some of them going back to mid eighties). Revised permits were assigned around the mid-nineties period.[30] For most of the nineties, the plants seemed to be facing about a 5-year permit cycle, per the statutes of the existing regulation. However, starting in the late nineties a number of the plants faced permit cycles that were shorter than the usual five year time span: ranging from one year to three years. In fact, there are two plants that experienced six permit cycles over the January 1990 to February 2004 period. Five plants sampled have 5 permit cycles, 22 have 4 cycles, 51 have 3 permit cycles and 15 have two cycles. Five other plants have only one cycle in this time period most likely on account of phasing out of permits.

Within a cycle, the monthly average permit level assigned to each polluter is constant, except for seasonal variation allowed for some plants. For these plants, higher discharges are allowed during "winter"[31] season (usually October through March) than during the rest of the year. As such, these permits reflect the need to discharge lower BOD in the plant's effluents in the "summer" than in the "winter", in order to protect water quality. The structure of their monthly permits was distinct from the other plants, which had a constant permit level irrespective of the season. Permit cycle for all the

---

[30] EPA's PCS field code called limits start (and end) date is used to identify the plant specific permit cycles.

[31] "Winter" denotes low temperature and high flow conditions, as opposed to high temperature and low flow conditions prevailing during "summer".

plants (including plants that did not face seasonal limits) is divided up as “summer” and “winter” averages, in order to accommodate the plants that exhibited seasonal variability. In the current sample there are 15 (17) such plants with seasonal variation in their permits for BOD5 concentration (quantity). Hence, the model that is estimated is categorized by plant-specific cycles (and seasons), which are determined by the regulator’s issue of a new permit to each plant.

In the current sample, there are 86 (24 in MD, 17 in PA and 45 in VA) plants facing concentration permits. There are 94 plants facing quantity limits on BOD with 23 plants in MD, 20 in PA and 51 in VA. As seen in Figure 2.1 and Figure 2.2 below concentration permits, by season, are declining (over 1990-2003) in MD and VA while in PA permits went down and then were made lax in the more recent time periods (a U shaped curve). Figure 2.3 and Figure 2.4 show that for limits on load, seasonal average limits were more or less constant for plants in MD and VA while for PA limits seem to have an inverse U shape, were relaxed earlier but have been made more stringent recently.

By type of plants, 54 POTWs and 4 public sector facilities faced limits based on both concentration and BOD5 load. Amongst manufacturing plants, 28 faced concentration limits while 36 had permits for BOD5 load assigned to them.

Figure 2.1: High Temp. Average Concentration Permits by State

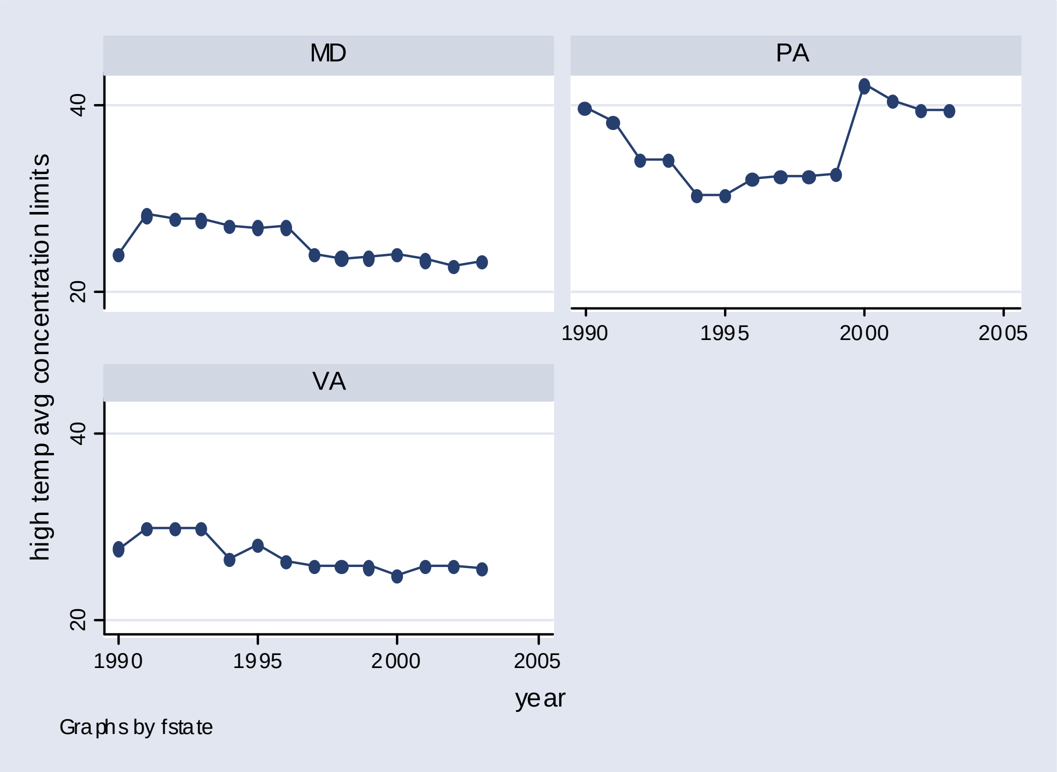
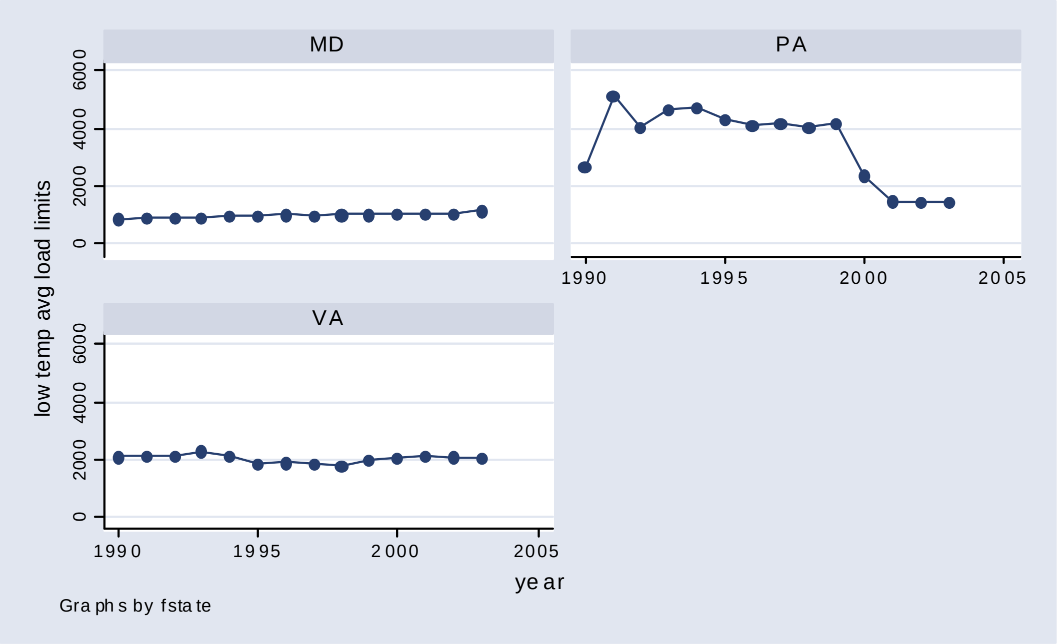


Figure 2.2: Low Temp. Average Concentration Permits by State

Figure 2.3: High Temp. Average Load Limits by State

Figure 2.4: Low Temp. Average Load Limits by State



The min and max of average concentration permits (by season) ranged from 5 mg/L to 111 mg/L (Table 2.2). Seasonal average quantity permits on the other hand exhibit much more variability across plants. The min and max of average quantity permits (by season) ranged from 9.7 lbs/day to 42400 lbs/day for BOD5 loadings (Table 2.2).

As mentioned before, a crucial aspect of classification of effluent limits (for purposes of the analysis in this chapter) is the distinction between technology based and water quality based limits. Contrary to the distribution of concentration permits, quantity limits do not assume a common value, both within and across plants, that would be even remotely indicative of the type of limits (i.e. technology or water quality based) faced by either municipal or industrial plants. This is primarily because of design flow differences across plants.

Table 2.2: Summary Statistics of Permits, Past Cycle DO, Design Flow and Socio-demographic Data

Variable	Mean	Min	Max
Permit cycle average BOD5 concentration limit by season, mg/L	27.78	5	111
Permit cycle average BOD5 quantity limit by season, lbs/day	1868.64	9.68	42400
Entire permit cycle mean (median) downstream DO, mg/L	9.64 (9.57)	4.72 (3.8)	13.62 (13.3)
Last three years of past cycle mean (median) downstream DO, mg/L	9.59 (9.52)	4.75 (3.41)	13.84 (13.7)
Last two years of past cycle mean (median) downstream DO, mg/L	9.57 (9.50)	4.65 (2.88)	13.66 (13.6)
Last one year of past cycle mean (median) downstream DO, mg/L	9.55 (9.46)	4.10 (2.88)	14.8 (14.8)
Excludes last two years of past cycle mean (median) downstream DO, mg/L	9.67 (9.62)	4.76 (2.65)	13.97 (13.7)
Excludes last one year of past cycle mean (median) downstream DO, mg/L	9.67 (9.61)	4.77 (2.65)	14.67 (14.4)
Design flow, million gallons/day	54.89	0.03	2000
Percent non-white	16.81	0	79.09
Median household income, thousands of $	29.54	13.13	60.59
Percent car-pooling	16.59	6.75	30.13
Percent employed in the manufacturing sector	23.01	3.77	53.76
Total population in the zip-code, thousands	14.66	0.49	68.44

Note: Table A 1, Table A 2, Table A 3, Table A 4, Table A 5 and Table A 6 in the appendix present detailed summary statistics of the data.

The number of permits with technology-based limits is not readily available for the state of Maryland, as has been confirmed by the MSDPD official, Curtis, H. Dalton. He added: “I believe a majority of the municipal discharge permits do not have technology based limits for BOD5, TKN and ammonia” (email communication). However, he was unable to sift the appropriate figure for major sewage treatment plants in Maryland with BOD5 limits (relevant for this current study since it restricts attention to ‘major’ plants).[32] In Virginia, a preliminary search in the state’s water control board information revealed that at least 12 plants included in the current sample have water quality based limits, termed as non-TMDL waste load allocations.[33] The concentration-based limits for the 10 treatment plants were either lower than 30 mg/L or they faced seasonal limits. Using the above two criteria, almost 60% of the 54 sewage treatment plants in the current sample face seasonal average concentration permits that are more stringent than the secondary treatment standard of 30 mg/L.[34] [35]

There are gaps in the monthly average permit levels. The reason for these gaps is hard to discern. It may be a consequence of reporting requirements or non-reported monthly data. However, the most likely source of missing data seems to be updates in the state-level PCS databases which could not be reconciled with the EPA’s PCS database. A noteworthy example of this type of missing data is when monthly

---

[32] He could verify that the number of permits with BOD5 limits below 30 mg/l is close to 100 (including major and minor), or approximately 40% of the total number of municipal permits.

[33] It could be verified that the following treatment plants had water quality based limits (quantity limits in kg/day are indicated for some of them in parenthesis): VA0021199 (273 and 454), VA0022390 (52.8 and 125), VA0025020 (1173), VA0025291 (182), VA0025305 (681), VA0060593 (1907), VA0069345 (257), VA0085952 (133), VA0060844 (818), and VA0025054 (130 and 260) (see regulation 9 VAC 25-720: http://www.townhall.state.va.us/L/GetFile.cfm?File=E:/townhall/docroot/103/2208/3866/Text_DEQ_3866_v1.pdf). Two industrial plants with water quality based permits were also identified in this cursory search: VA0002178 (1570), VA0002160 (272).

[34] This upper limit corresponds to at least 85% removal of BOD from the treatment plant’s effluents, given its technology.

[35] Conversely, almost 10% of these plants have concentration permits that are less stringent than the secondary treatment requirement.

observations at the beginning of a permit cycle are not present for multiple plants in Virginia. The other two states also have missing data but usually only for a couple of months.[36]

The primary explanatory variable is the measure of plant specific, past permit cycle seasonal average water quality. Permit writers most likely incorporate entire past cycle water quality in their permitting decisions since the minimum ambient water quality standard of 4-5 mg/L DO is necessary at any point in time and hence every month and cycle considered. Even though, others such as Earnhart (2007) suggested that the regulator most likely chooses a plant's permit level for the new cycle, with a one to three year "lead time" before the expiration of the preceding limit.[37] However, consultations with actual permit writers, confirmed that occasionally effluent limits are revised in case new data on critical flow become available before the new cycle becomes effective, even after the draft has been approved in the public notice period (Kyle I.Winters, P.E.).

Monthly water quality data is averaged over the time period of the preceding permit cycle and by season because effluent limits are assigned responding to the seasonal variation in the ambient water quality observed in streams and rivers.[38] The max and the min of the preceding cycle seasonal average dissolved oxygen ranged from

---

[36] A DEQ official in Virginia claimed that "DEQ has either changed, or made changes to, the databases used to process discharge monitoring data several times since 1993; it is possible that some of the changes resulted in lost data during the upload into EPA's PCS database" (Kyle I. Winters; email communication).

[37] Hence, an alternative indicator for past water quality was tried: average downstream water quality prevailing in the previous cycle, up until one to two years before the new permit level becomes effective.

[38] The well-documented seasonal trend of dissolved oxygen is that during the high temperature (impact on solubility of oxygen) and low flow (impact on flushing and re-aeration), summer and fall months (as observed in MD, PA and VA), ambient concentration of DO is lower. Due to the same physical processes, dissolved oxygen concentrations are anticipated to be higher during the low temperature and high flow, winter and spring months. Hence, BOD limits are more stringent during the high temp and dry seasons, and higher during the low temp and wet seasons.

4.72 to 13.6 mg/L (Table 2.2).[39] On an average, water quality in the previous cycle was recorded at 9.6 (with 50% of the observations also around 9.6). This mean value was well above what is required to maintain "aquatic life support" use. It is the use directly related with and hence defined by ambient concentrations of dissolved oxygen. It is also part of the "baseline" use (fishable/swimmable) that every stream or river system across the US needs to meet.

In the absence of knowledge on the exact time span over which permit writers examine past water quality, six alternative measures of preceding cycle DO are considered. These are past permit cycle seasonal average DO (mean and median), mean and median DO of last three, two and one year(s) of preceding cycle, by season, and mean and median DO excluding last two and one year(s) before past cycle ends, by season. The last two measures of water quality (in Table 2.2) attempt to capture past cycle water quality under the presumption that the permit writer might decide on the effluent limits with a 'lead' time of one to two years.

The different measures of seasonal, past-cycle average dissolved oxygen do not reveal much variation in terms of the values assumed by them. This in turn implies that ambient water quality might not have exhibited a significant trend, i.e., undergone substantial ups or downs over the span of a permit cycle (2-5 years) of the sampled plants. Such a pattern would be witnessed when the assimilative capacity of waterbodies successfully 'absorbs' the excesses in pollution discharges from various sources of pollution: most likely with the 'aid' of certain types of easily controllable pollution sources such as plants (as opposed to farms). In this context, the point sources would

---

[39] Seasonal variability in ambient dissolved oxygen implies that the lower end of the distribution most likely, prevailed during the high temp and low flow seasons.

have had to assimilate the impact of ‘permanent’ shifts in the pollution discharge from nonpoint sources by controlling their own effluents.

The other explanatory variable is design flow, which captures a plant’s size in terms of abatement technology. It is expected that all else equal, plants with higher design flow (i.e. higher volume) will pollute more on an average, in terms of pounds of pollutants discharged per day in its effluent, and hence will be assigned a higher level of quantity permit. This measure is actually time-invariant both within and between cycles and seasons, for all dischargers in the sample. It ranged from 0.03 to 2000 million gallons per day (Table 2.2). Three plants do not have information on design flow. Age of the plant or more precisely, the vintage of the abatement /treatment technology of the plant, aught to be a factor in determining the choice of permits. However, there is no proximate data that might capture this plant specific feature.

Zip code level data on the local socio-demographic characteristics are obtained from the 1990 US Census Summary Tape File 3B. These effects would capture variations due to indirect regulatory pressure; thereby, resulting in different intensities of the same policy being implemented across locations. On an average, the percent of population that is non-white is 17, the median income of a household is almost 30,000$, the car-pooling population is 17%, percent employed in industries is 23, the total population is almost 15,000, and the percentage of residents living in urban areas is 30%. Nature of business of the plant as indicated by the SIC code is also an important factor in determining the permit levels assigned to the polluter. There are 54 sewage treatment plants, 6 power generation plants, 4 food and kindred products, 4 textile mill products, 7 paper and allied products, 10 chemicals and allied products, 6 petroleum refining and related industries, 1 rubber and miscellaneous plastics products, 2 leather and leather products, 1 fabricated metal products, except machinery and transportation equipment, 1 transportation equipment, 1 justice, public order, and safety, and 3 national security and international affairs. State level dummy variables are included to control for disparities in the stringency of state level regulation. Significant coefficients on these dummies would indicate differences in regulation i.e. variations in policy such as different classifications of the designated uses of water body segments adopted in a state.

Estimation of the Permits Model

The estimation is conducted on two sub-samples out of the total of 100 plants. First, the subset of sewage treatment plants that either faced effluent limits on BOD concentration lower than 30 mg/L or seasonal limits is considered. Three reasons drive this focus on POTWs: for manufacturing (and other public sector) plants, an effluent limit lower than 30 mg/L does not necessarily indicate that the plant is facing water quality based effluent limits (WQBELs); second, for POTWs that continue to face 30 mg/L as their permitted concentration level, ex-post, one might not find any empirical evidence of the influence of past water quality on limits chosen by the regulator. Lastly, irrespective of whether these POTWs actually underwent changes in the permit levels assigned to them, one can expect that an empirical relationship between past-cycle DO and effluent limits can be evinced. This speculation is valid as long as one has systematically identified these effluent limits as most probable candidates for WQBELs. This claim, in turn, holds true primarily due to the ‘inertia’ of ambient water quality pattern in the current sample of monitoring locations (as seen in the descriptions on the summary statistics of the various measures of past cycle DO).[40] There were 32 (out of a total of 54) treatment plants that faced concentration permits more stringent than the 30 mg/L standard, or seasonal limits were applicable for them.

The second subset of plants (manufacturing as well as POTWs) underwent at least one change in their permitted levels (concentration or quantity), during the time period of this analysis. This criterion is applied, first, so that permitting decisions of industrial dischargers can be examined, albeit, with the shortcoming of not differentiating between technology or water quality based limits. Specifically, one might be able to find evidence on the impact of past cycle water quality on the new limits chosen by the regulator (e.g. a transition from technology to water quality based limit). ‘Inertia’ in the various measures of past cycle water quality is again relied upon, since majority of them probably witnessed only one shift in their permit levels over the 14 year time span. Consequently, during the cycles when no further revision in permits is observed, a relationship might still be found since no significant time trends have been detected in the seasonal average ambient water quality by permit cycle. There were 21 and 66[41] plants that faced a change in their concentration and quantity permits at least once over the period 1990 to Feb. 2004.

Table 2.3 below presents the results of estimating equation (2.6) (repeated below) after pooling the above two sub-samples. The merged sample is comprised of 38 and 74 manufacturing plants and public sector facilities that underwent at least one change in their permitted levels, and/or sewage treatment plants with permits more stringent than 30

---

[40] Not to mention the cross-sectional variability in water quality across different monitoring locations.

[41] Two plants dropped out of the regression sample from the total of 66 since they did not have data on design flow.

mg/L.

$$P_{ics} = \alpha + \gamma DO_{j(c-1)s} + \tau designflow_{i(c-1)s} + \pi sic_i + \sum_{k=1}^{6} \lambda_{ik} sociodemographics_{ik} + \theta_r + \varepsilon_{ics}$$

The model regresses seasonal average concentration/quantity limits on past cycle seasonal average water quality, design flow, type of production facility, sociodemographic aspects, and state level controls. Correlation in error terms within plants, and from one permit cycle to the next, is implemented primarily in order to address the unobserved heterogeneity problem that arises once the time series element of multiple plant specific permit cycles is introduced. The FGLS method is used to estimate the parameters of the linear regression model in which the errors are serially correlated; in particular, plant specific error terms (by season) are correlated with that of the preceding cycle error within the same plant (and season). The procedure (StataCorp, 2007) involves first obtaining an estimate of the serial correlation coefficient by regressing the residuals obtained from running OLS on equation (2.6) on residuals for the previous cycle. Next, the original model is transformed using the Prais-Winsten transformation (Prais and Winsten, 1954). Finally, GLS estimates of the original parameters are obtained, conditional on the estimated serial correlation coefficient.

The coefficients below show that the seasonal average (mean and median) dissolved oxygen levels prevailing in the preceding permit cycle had a significantly positive effect on both concentration and quantity permits chosen by the regulators.[42]

---

[42] Detailed results of the individual models estimated are presented in the appendix.

Table 2.3: Level and log transformed permits model correcting for serial correlation of errors (FGLS)

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	log avg. concentration permit, by season	log avg. quantity permit, by season
past permit cycle mean (median) downstream DO, by season	1.328** (1.176)**	46.356** (38.593)**		
design flow	-0.100** (0.102)**	2.846 (2.786)		
log past permit cycle mean (median) downstream DO, by season			0.617** (0.556)**	0.322** (0.268)**
log design flow			-0.078* (-0.082)*	0.707** (0.706)**
Observations	175	345	171	341
Number of group(npdes)	38	72	37	71

Note: + denotes significant at 10%; * denotes significant at 5%; ** denotes significant at 1%.

The positive direction of impact means that if mean water quality during the past cycle is lowered by one mg/L, then concentration permits in the subsequent cycle would be reduced by 1.328 mg/L i.e. made more stringent. Alternatively, if median water quality prevailing over the entire span of past cycle is increased by one mg/L, then permit levels in the 'new' cycle would be made less stringent by 1.176 mg/L.

For quantity permits, the coefficient is 46.356 and 38.593 lbs/day depending on whether the mean or median value for preceding cycle water quality is considered. The coefficient can be interpreted as: if mean DO during the entire permit cycle is lowered by one mg/L, then the seasonal (average) quantity measure of quantity limits chosen for the subsequent cycle would be made more stringent by 46.356 lbs/day.

The log-log models in cols (3) and (4) are estimated primarily in order to facilitate interpretation of the estimated impacts in terms of percentages. The positive sign can be interpreted as: if mean water quality during the preceding permit cycle is raised by one percent, then seasonal average concentration permit of BOD5 chosen for the 'new' cycle would be raised by 0.617 percent. Conversely, if median DO prevailing in the previous cycle is reduced by 1 percent, then seasonal permit chosen for the concentration measure of BOD5 is lowered by 0.556 percent.

For the log-transformed quantity permits models, the coefficient can be interpreted as: if mean (median) dissolved oxygen during past cycle is increased by one percent, then average quantity permit chosen in the 'next' cycle and for the same season would be raised by 0.322 (0.268) percent.

Design flow, the other 'core' variable, plays an important but opposite role in the (level and) log concentration and quantity permits models. For the level and log concentration permits models, the coefficient is negative implying it might be capturing the vintage effect i.e. plants with larger design flow are newer with more efficient technology and hence can meet a more stringent limit. Alternatively, it could be capturing the effect on water quality: as will be seen in Chapter 4, ceteris paribus, plants with bigger design flow has an adverse impact on ambient quality, since it gives a bigger weight to the effluent concentration of BOD. Quantity permits, on the other hand, might be capturing the scale effect i.e. bigger plants pollute more in terms of load of BOD.

In order to test for robustness, other measures of past cycle water quality are also tried (separately). Table 2.2 presents the summary statistics of the last three, two, one and excluding last two and one year's seasonal average water quality in the past cycle. Table A 7 and Table A 8 in the appendix compile the results obtained from FGLS estimations of the concentration and quantity permits model (equation (2.6)) using the different measures of water quality.

For concentration permits, the estimated coefficients on various measures of past cycle downstream DO range from 1.129 to 1.381 (1.033 to 1.246) when mean (median) values are included. In general, the estimated impacts seem to be higher for the averages taken over longer time spans (one or two year 'lead' time or the entire permit cycle). For quantity permits, the estimated impact of the various measures of past cycle water quality range between 40.849 and 48.769 (32.642 and 43.324) lbs/day (Table A 8). Unlike the level and log transformed concentration permits models, the mean and median value of the last year of the previous permit cycle assumes the largest magnitude, while average DO excluding last two years of the past cycle has the lowest impact.

Overall, the coefficients on different past cycle dissolved oxygen measures are lower in magnitude for the quantity model(s) as can be seen from the log-transformed concentration and quantity models. The impact of past cycle DO is around 0.3 percent. By contrast, the coefficients on past cycle DO are almost 0.6 percent for the log of concentration permit models. Inference can be drawn on past cycle water quality assuming greater importance when deciding on the concentration effluent limits.

An alternative interpretation could also be given to the apparently higher relevance of past cycle DO when making permitting decisions on concentration of BOD in contrast to quantity. It could also mean that in response to the slow changing downstream water quality during past permit cycle, concentration permits are seldom revised;[43] [44] given that changes in effluent concentration of BOD necessarily implies a change in ambient pollution in the ‘mixing’ zone. In particular, 32 out of the 38 plants in the merged sample for the permit level choice models were already identified as WQBELs. It means that these plants might have been facing WQBELs even before the first year of the time period of this study.

Quantity permits, on the other hand, have more leeway for revisions by the permit writers since ceteris paribus a higher load, in particular, a higher effluent flow implies a bigger weight given to effluent concentration and a lower weight to upstream ambient concentration of BOD. In terms of the results of the permit level choice models, 66 out of the 74 plants with quantity permits were identified as having undergone a change in their effluent limits, which might include changes from a technology to water quality based limits.

Do Regulations Differ By State?

The model in this section tests differences in regulation by state in the context of their response to ambient water quality. In other words, γ, the coefficient on past permit cycle average DO is allowed to vary by state. To implement this framework the measure of past cycle average DO is differentiated according to the state in which the plants are

---

[43] In which case, most of the variability in downstream water quality would be coming from the cross sectional differences, across monitoring locations.

[44] See Table A 21 and Table A 22 in the appendix which show that plant specific seasonal concentration permits (with mean change around –0.92 mg/L) have not varied as much as quantity limits (mean change around 83.38 lbs/day).

located in order to capture any differences on the part of the permit writers by state, in particular, those arising out of variations in regulation itself. As mentioned before, the states are authorized to assign designated uses for their own waterbodies. It means that any measurable differences in the influence of past cycle water quality on current cycle permitting decisions might reflect that uses in Maryland are more refined i.e. require higher standards for ambient quality when compared to Virginia or Pennsylvania. The results below show that indeed regulators in Maryland respond more to past cycle water quality than those in Virginia and Pennsylvania.

If past cycle mean DO is increased by one mg/L in Maryland, then new concentration (load) limits chosen would be made less stringent by 2.282 mg/L (90.379 lbs/day). For plants in Virginia, a similar increase in past cycle DO would lead to (only) concentration limits chosen in the new cycle to be 0.798 mg/L higher. For plants in Pennsylvania, past water quality does not have any influence on permitting decisions in the new cycle for either concentration or load permits. Columns (2) and (4) of Table 2.4 below show that, as expected, the permit writers in Maryland are paying more attention to past cycle water quality than those in Pennsylvania or Virginia even in the log transformed models. A one percent increase in past cycle DO would lead to 1.045 (0.670) percent increase in concentration (load) limits chosen for plants in Maryland.

For plants in Virginia, a one percent increase would lead to a smaller effect on new concentration and load limits chosen; in particular, it would be made less stringent by 0.234 and 0.161 percent (respectively). Inference can be drawn on differences in regulation such as designated use classifications adopted by Virginia being less stringent, in terms of water quality standards, than Maryland.

Table 2.4: Level and log transformed permit models with DO differentiated by state (FGLS)

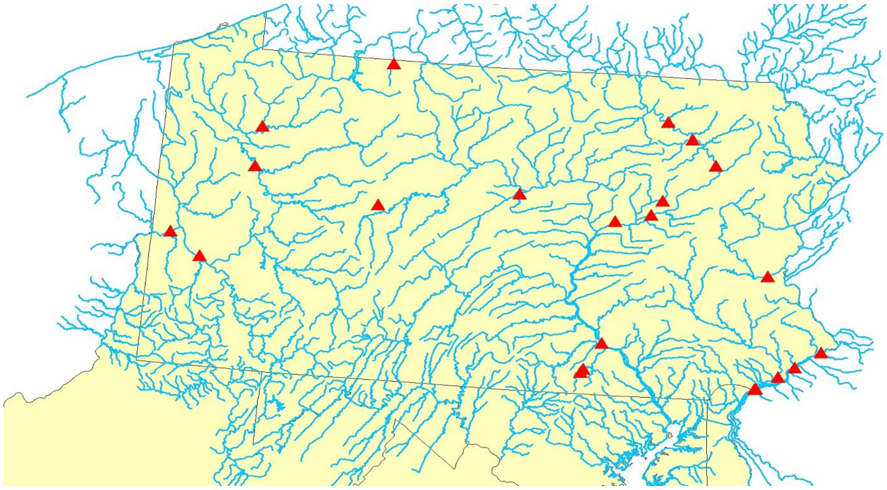
	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	log avg. concentration permit, by season	avg. quantity permit, by season	log avg. quantity permit, by season
past permit cycle seasonal mean DO in MD	2.282 (0.000)**		90.379 (0.006)**	
past permit cycle seasonal mean DO in PA	0.388 (0.198)		37.642 (0.283)	
past permit cycle seasonal mean DO in VA	0.798 (0.055)+		20.167 (0.198)	
design flow	-0.097 (0.000)**		2.841 (0.598)	
log of past permit cycle seasonal mean DO in MD		1.045 (0.000)**		0.670 (0.000)**
log of past permit cycle seasonal mean DO in PA		0.092 (0.244)		0.149 (0.364)
log of past permit cycle seasonal mean DO in VA		0.234 (0.013)*		0.161 (0.023)*
log design flow		-0.080 (0.028)*		0.704 (0.000)**
Observations	175	173	345	343
# of plants	38	37	72	71

Note: p values of one-tailed test in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%.

Tests on the estimated coefficients reveal that Maryland is significantly different from Pennsylvania and Virginia (in terms of the influence of water quality on permitting decisions), but not from each other.[45]

Permit writers in Maryland are found to be more responsive to ambient water quality possibly because all the sampled plants discharge their effluents, ultimately, into the Chesapeake Bay. Given the high priority assigned to the Chesapeake Bay such that uses will be met at a future time period, the permit writers in Maryland might be more aggressive in responding to the water quality of the stream/river, which eventually flows into the Bay. The GIS maps in Figure 2.5, Figure 2.6 and Figure 2.7 show that some of the sampled plants in Pennsylvania and Virginia, particularly the ones in the west, discharge their effluents into streams and rivers that ultimately flow into the Allegheny, Ohio and Mississippi Rivers.

Figure 2.5: Location of Sampled NPDES Plants in Pennsylvania



---

[45] Test statistics and p-values for state level differences in regulation are: concentration: md=pa chi2(1) = 9.42 (Prob > chi2 = 0.0021), md=va chi2(1) = 5.15 (Prob > chi2 = 0.0232), pa=va chi2(1) = 0.37 (Prob > chi2 = 0.5453); log concentration: md =pa chi2(1) = 12.07 (Prob > chi2 = 0.0005), md=va chi2(1) = 12.34 (Prob > chi2 = 0.0004), pa=va chi2(1) = 0.20 (Prob > chi2 = 0.6575); load: md= pa chi2(1) = 0.50 (Prob > chi2 = 0.4807), md =va chi2(1) = 2.69 (Prob > chi2 = 0.1012), pa =va chi2(1) = 0.06 (Prob > chi2 = 0.8007); log load: md=pa chi2(1) = 2.87 (Prob > chi2 = 0.0905); md=va chi2(1) = 3.13 (Prob > chi2 = 0.0766); pa =va chi2(1) = 0.64 (Prob > chi2 = 0.4250).

Figure 2.6: Location of Sampled NPDES Plants in Maryland

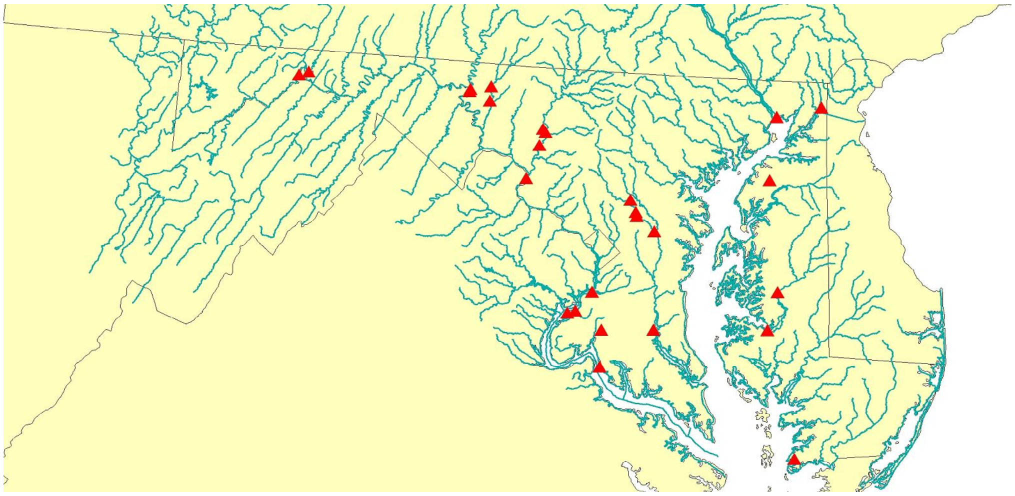
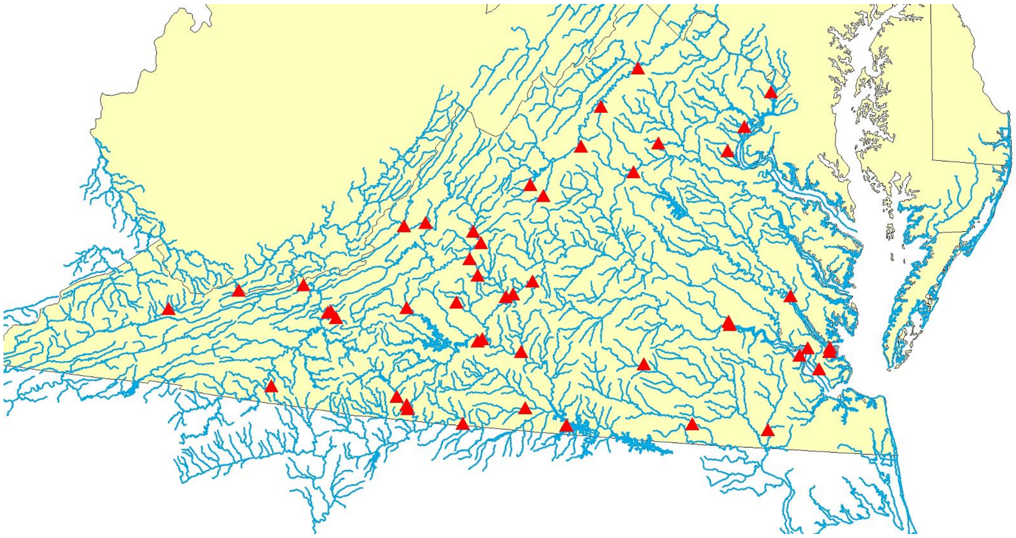


Figure 2.7: Location of Sampled NPDES Plants in Virginia



Conclusions

In this chapter, evidence is presented on the nature of choices made by local regulators, in particular, have the permit writers been using effluent limits as policy instruments to ensure that ambient water quality standards are maintained. Hence, estimating the impact of downstream DO prevailing during past permit cycle on the effluent limits chosen in the new cycle is the focus of this chapter. The empirical results indicate that decentralized water pollution regulation was indeed effective. In other words, permits choices made by the local regulators incorporated water quality in their permitting decisions.

First, past ‘period’ ambient water quality is the most direct channel through which local regulators can influence the ‘status’ of water pollution control. Hence, it is more relevant than related exogenous measures such as stream flow or velocity, or total water area in a state. Given that protecting the designated uses of waterbodies is the fundamental requirement of the CWA, even measures that are more directly related to the benefits of pollution abatement, e.g., total surface water withdrawn are less appropriate. Secondly, previous studies on the regulator’s choice of effluent limits were carried out in a cross-sectional framework. In other words, at a given point in time, what factors influenced the permit choice of the regulator was estimated. The analytical models presented in this chapter, incorporates the time dimension of permit cycles. This is a richer analysis since the permit writer most likely makes her/his decisions based on factors observed in the preceding permit cycle rather than the current cycle.

For the permit level choices, if mean (median) water quality prevailing over the entire span of past cycle is increased by one mg/L, then permit levels in the ‘new’ cycle would be made less stringent by 1.328 (1.176) mg/L. For quantity permits, the coefficient is 46.356 and 38.593 lbs/day depending on whether the mean or median value for preceding cycle water quality is considered. In terms of percentages, the estimated influence of water quality for concentration and quantity permits is 0.617 (0.556) and 0.322 (0.268) percent, respectively. Hence, some evidence is found on the regulators utilizing limits on both effluent concentration and loadings as policy instruments, in particular, to insure that the ambient water quality goals are met. Additionally, inference can be drawn on past cycle water quality assuming greater importance when deciding on the concentration effluent limits. Changes in effluent concentration of BOD imply a direct (one-to-one) change in ambient pollution. On the other hand, ceteris paribus a higher load and hence a higher effluent flow implies a bigger weight assigned to effluent concentration and a lower weight to upstream ambient concentration of BOD (equation (4.5) shows how ambient pollution in the wastewater-river mixing zone is determined). Therefore, it has been verified that permit writers in the states of MD, PA and VA, at least for the current sample of plants, indeed, utilize both concentration and pollutant loadings to insure ambient standards continue to be met.

Chapter 3:Impact of Past Water Quality on Abatement Decisions

Introduction

The primary focus of this chapter is to find evidence on the impact of downstream water quality prevailing in past time periods on pollutant discharges relative to effluent limits, and hence on abatement behavior of plants.[1] Plants pay attention to water quality observed in their receiving waters over and above their effluent limits, since the possibility of an increase in the stringency of its permits, on account of a deteriorating status in ambient quality, exists perpetually.

Occasionally, new information on critical low flow (or actual water quality data) becomes available to a permit writer, not necessarily coinciding with the permit cycle of the relevant plant, that was not available at the time of permit issuance (Kyle I. Winters; email communication). In such a situation, regulators are obligated to conduct a water quality evaluation to find out whether the ambient standard will be violated in the near future. During the 1990s, such evaluations involved simulating the impact of the current level of effluent limits on ambient water quality under critical stream flow (drought) conditions. If the projected water quality was perceived as violating the standard, then a more stringent water quality based effluent limit (WQBEL) was imposed on the polluter. Alternatively, if observed water quality violates the standard, then permit writers have to re-assign a pollutant limit for the plant such that the use will be met subsequently.

---

[1] Heuristically, if past water quality is lower, then this has implications for the benefit of pollution abatement from the same amount of pollutant discharge being higher; ultimately, leading to a higher level of abatement and hence lower pollutant discharges. For example, Gray and Shadbegian (2004) incorporate an estimate of the expected benefits per unit of pollution abatement in their model, and find that it significantly imposes a downward pressure on plant level pollutant discharges.

The above discourse on WQBELs implies that if plants do not incorporate ambient water quality considerations in their (daily) monthly average discharge decisions, then the potential for increased interventions, such as inspections and sanctions, due to deteriorating water quality is raised; eventually, it leads to lower limits on effluent discharge by law. Hence, plant managers of sewage treatment and industrial facilities alike make a strategic choice as part of their abatement strategy to take notice of the water quality downstream to the points of their effluent discharge. Given that abatement technology is inflexible (Bandyopadhyay, 2002), the underlying incentive is to minimize the expected costs of penalties due to non-compliance especially when a more stringent permit level is imposed on them. Concurrently, operators of publicly owned facilities and owners of industrial plants might want to be seen as good citizens exhibiting an active concern for ambient water quality and hence for the environment in general (McClelland and Horowitz, 1999; Houtsma, 2003).

Typically, an empirical model that investigates the abatement behavior of plants entails regressing monthly pollutant discharge on lagged downstream water quality, monthly average permitted amount of the pollutant, and a few other plant specific and location specific variables. Given the evidence on regulators incorporating past cycle water quality considerations when deciding on effluent limits, endogeneity of current (cycle) permit levels is accentuated in the above pollutant discharge model. Specifically, the time span covered multiple permit cycles (usually ranged from 3 to 5 years) implying that the effluent limit is not exogenous or a predetermined factor. For instance, one would expect effluent design flow to play a significant role in the permits choice decision of the regulator as well as in the polluter's decision of actual amount of pollutant to be discharged. Similarly, location specific factors, in particular, ambient water quality might influence the plant's choice of pollutant discharge amount at the same time as exerting its impact on the permits choice decision of the regulator.

Consider the situation that a plant starts its new permit cycle in month "t". Actual discharge in month "t" is then regressed on permit level for month "t" and, among other factors, average ambient water quality prevailing one year ago (for instance). We know that permit level in month "t", in turn, might have been chosen based on average ambient water quality prevailing one year ago. Alternatively, suppose that the plant is not on its new permit cycle in month "t". If its new permit cycle started in any month that was less than one year ago, then average water quality prevailing one year ago might have played a role in determining the permit level in month "t". Endogeneity of limits is sustained due to the assignment of the monthly average permit levels (usually) for a period of five years. To circumvent this problem, the empirical framework of this chapter examines whether past ambient water quality prevailing at downstream locations determine pollutant discharges relative to its effluent limits.

Evidence on the influence of downstream water quality prevailing in past time periods on the abatement decision of polluters will provide an additional clue to the pervasive overcompliance behavior witnessed during the time period of this analysis.[2] This is not to mention its implications for the effective functioning of the water pollution regulation. In particular, if plants have been responding to the underlying cause for its

---

[2] As will be seen later, past studies have found evidence on positive marginal compliance costs (Earnhart, 2007) concurrent with the presence of 'significant' overcompliance, even after, controlling for randomness in the water pollution process (Bandyopadhyay and Horowitz, 2006), and synergistic factors such as jointness in pollution abatement technology (Akobundu, 2004).

effluent limit stringency i.e. ambient water quality,[3] over and above its' permitted levels, then implementing regulations based on fully accounting for the 'contributions' of all sources of pollution (point and non-point) is a step in the right direction.[4]

Lagged rather than contemporaneous downstream water quality is the appropriate explanatory factor. This is primarily because in month 't' a polluter is most likely aware of the ambient water quality prevailing in the stream/river 'receiving' its effluent discharges during a month that is prior to 't'. In fact, downstream quality observed in month 't' is an outcome of the plant's discharge of pollutants in its effluents in month 't' (apart from other sources of pollution and exogenous factors such as season, etc).[5] From a practical standpoint, in month 't' the polluter has knowledge of its receiving water quality up until about 2 to 5 months in the past. Most publicly available databases update their information from 2 to 5 months in the past.[6]

---

[3] If ambient water quality was low in terms of being projected to fall below the standard required to maintain its designated use, then effluent limits on point sources were made more stringent by imposing water quality based permitted levels of pollution.

[4] During most of the 1990-2003 period "major" point sources faced plant-specific effluent limits based on technology or water quality criterion. For instance, water quality based limits for a plant were chosen after assessing the 'contribution' of the plant to adversely impacting downstream water quality. Non-point sources, on the other hand, were regulated under some general (non-enforceable) criteria based on 'cost effective and reasonable' best management practices (BMP) (Brooks and others 1997). Consequently, point sources responding to ambient water quality in addition to its' regulated pollutant limits meant that they 'compensated' for the pollution generated from the effectively unregulated nonpoint sources.

[5] This aspect of the water pollution process is examined in the chapter on formulating an empirical water quality model.

[6] For example, the databases used for gathering data on ambient dissolved oxygen concentration: the USEPA's STORET, the state level databases of Maryland's Chesapeake Bay Program (CBP) and Virginia's Department of Environmental Quality (DEQ) have their most recent data that is 2-3 months in the past.

Review of the Literature

There is a substantial number of studies that examine the compliance and/or abatement behavior of wastewater treatment plants and manufacturing plants such as pulp and paper mills (since they are notably significant polluters of BOD). It is one of the two most “popular” i.e. conventional pollutants (the other one is TSS). The empirical models estimated in most of these papers focus on the regulatory aspects of inspection (others, on enforcement) and community characteristics of the surrounding population, and the resultant BOD discharged by the polluters.

Magat and Viscusi (1990) used quarterly data of 77 pulp and paper mills from the (Permit Compliance System) PCS database of the EPA for the time period 1982 to 1985.

$$\begin{aligned} POLLUTION_{it} &= \alpha + \beta_1 POLLUTION_{it-4} + \sum_{k=1}^{n} \gamma_k IQTR_{t-k} + \beta_2 TONS_i + \beta_3 SIC_i \\ &+ \beta_4 REGN_i + \beta_5 QUARTER_t + v_{it} \end{aligned} \tag{3.1}$$

$POLLUTION_{it}$ is the number of pounds of BOD discharged per day by plant i averaged over the quarter t. $POLLUTION_{it-4}$ is the four-quarter lagged dependent variable which serves as a good proxy for the firm’s stock of capital related to pollution control. High levels of pollution in the past are related with high levels of pollution in the future, because the nature of the abatement technology makes it costly to increase abatement levels. The four quarter lagged value also captures the seasonality effect that often not only determines plant level operations but the permitted levels also vary depending on seasons. The estimated coefficient was 0.98, which implies that past pollution levels was almost a perfect predictor of current levels.

$\sum_{k=1}^{n} \gamma_k IQTR_{t-k}$ is a distributed lag on past EPA inspections. This variable takes on a value of one if the pollution source received an inspection in period $t-k$ i.e. k quarters previous to the pollution measurement in the current quarter t. One quarter lagged inspection reduces mean value of MQAVG by about 20%. $REGN_i$ is the dummy variable for the EPA region in which the plant is located. SIC_i is the dummy variable for the plant's four-digit SIC industry code. $TONS_i$ measure the number of tons of pulp and paper produced daily at the plant i. An OLS model was estimated including a series of twelve quarterly dummy variables. Their results did not show any time pattern, or regional differences, or any 4 digit level industry specific pollution patterns.

Laplante and Rilstone (1996) examined the impact of past 12-month inspections (with equal coefficient on each lag) as well as the threat of an inspection in the current period on absolute discharges and discharges relative to the norm i.e. permitted levels. Their sample comprised of about 60 pulp and paper plants in Canada, for the period 1985 to 1990, and monthly data on BOD and TSS discharges. They estimate a Probit model of the probability of an inspection depending on variables in the basic pollution equation, a variable indicating the total number of inspections that have been conducted at the plant prior to the current period, capacity, and time. The predicted probability of inspection is used as an instrumental variable for the current period threat of inspections in the BOD pollution equation (3.2) below.

$$P_{it} = \alpha + \phi P_{i(t-12)} + \theta_0 INS_{it} + \sum_{j=1}^{12} \theta_j INS_{i(t-j)} + \beta_2 REG_i + \beta_3 PROD_i + \beta_4 CAP_{it} + \gamma t + \varepsilon_{it} \quad (3.2)$$

$$i = 1,2....46; t = 1,2,...60$$

P_{it} is the pollution variable (BOD or TSS measured in lbs/day, absolute or relative to permitted level) associated with plant i in period t. $P_{i(t-12)}$, the plant's 12-month lagged value of pollution, was a very good determinant of current period pollution (coefficient around .87). INS_{it} is the plant's current period threat of an inspection. $INS_{i(t-j)}$ captures whether the plant was inspected in any previous time period. REG and $PROD$ are dummy variables reflecting the plant's location and type of output. CAP indicates the plant's daily productive capacity at time t. Lastly, a time trend t was included. With the exception of BOD relative to the norm, the coefficient estimates on current and lagged inspections were negative and highly significant. Absolute BOD was reduced by 5.4 lbs/day due to past inspections, while the impact of current period probability of inspection was around –193 lbs/day.

In their 2004 paper, Gray and Shadbegian examine the determinants of levels of water pollution (BOD and TSS) of 231 paper and pulp mills in the US for the period 1985 to 1997. EPA's Envirofacts database was used for the monthly water pollution discharges data on BOD and TSS. The model estimated is:

$$Y_{it} = f\left(PLANT_{it}, PEOPLE_{it}, STATE_{it}, YEAR_{t}\right) \tag{3.3}$$

In equation (3.3), the plant specific factors $PLANT_i$ are size (pulp/paper capacity), age (a dummy variable to indicate if plant was established after 1960), firm's financial condition (owning-firm's rate of return on its assets), ownership structure (a dummy variable to indicate if the plant is the only paper and pulp mill owned by the firm), major source (EPA's Majors Rating Database), public health effects (dummy variable to indicate if the plant discharges into a stream segment that is a source of drinking water), and other non-environmental risks (Occupational Safety and Health Agency's (OSHA) violations).

$PEOPLE_i$ are location specific factors such as the percentage of nonwhites in the nearby population (within a 50 mile radius of a plant), and an interaction between county voter turnout in presidential election and membership in conservation organizations. An estimate of expected benefits per unit of pollution reduction is also included. First, the pollutant discharge of the plant in concern was varied, and the resultant water quality was projected using stream flow data. Gray and Shadbegian relied on Carson and Mitchell's formula (based on a survey of consumers' willingness to pay for an average improvement in all rivers in a state), to calculate the estimates of the dollar benefits of improved water quality. Their analysis implied that ceteris paribus benefit of water pollution abatement is greatest when ambient water quality is lowest.

State specific controls $STATE_i$ are environmental attitudes ("green vote": support for environmental legislations by the state's Congressional delegation), average inspection rate for all the other plants in the state across all industry types, unemployment rate in the state for that year, percent of the county designated as urbanized, state level dummy variables, and the effect of political boundaries is accounted by two dummy variables indicating whether the plant is within 50 miles of another jurisdiction (i.e. state or the Canadian border). Annual dummy variables were also included to capture time trends with a base year of 1989.

An OLS model was fitted on the logarithms of pollutant quantities, because of the wide dispersion in discharges across plants. A one standard deviation increase in water pollution abatement benefits was associated with 16% lower BOD and 23% lower TSS discharges. Pulp rather than paper capacity seemed to matter more, and larger plants generated more pollution. Plants in urban areas generated less pollution, but also (surprisingly) faced somewhat less regulatory activity.[7]

Another set of studies focused on the aspect of overcompliance in water pollution abatement behavior of "major" polluting facilities. Earnhart (2004c) examines monthly wastewater discharges by 40 "major" municipal facilities in the state of Kansas for the years 1990 to 1998. The equations estimated are:

$$\begin{aligned}\ln Y_{it} &= \beta_1^{EPA} IP_{it}^{EPA} + \beta_1^{KD} IP_{it}^{KD} + \beta_1^{S} S_{i(t-12)} + \beta_2^{EPA} IL_{it}^{EPA} + \beta_2^{KD} IL_{it}^{KD} \\ &\quad + \beta_3^{EPA} SL_{it}^{EPA} + \beta_3^{KD} SL_{it}^{KD} + \beta_4 X_{it} + \varepsilon_{Yit}\end{aligned} \tag{3.4}$$

$$I_{it}^{EPA} = \theta_1 \breve{Y}_{i(t-6)} + \theta_2 Z_{it} + \theta_3 IP_{it}^{KD} + \varepsilon_{Eit} \tag{3.5}$$

$$I_{it}^{KD} = \alpha_1 \breve{Y}_{i(t-6)} + \alpha_2 Z_{it} + \alpha_3 IP_{it}^{EPA} + \varepsilon_{Kit} \tag{3.6}$$

In equation (3.4), Y_{it} is the measure of plant level monthly BOD/TSS discharges[8] relative to its permitted levels. The current month threat of EPA and state inspections are instrumented by the predicted probability of an EPA and a KDHE inspection, IP_{it}^{EPA} and IP_{it}^{KD}, which are estimated from equations (3.5) and (3.6). $S_{i(t-12)}$ is the 12-month lagged indicator of a state or federal enforcement action taken against a facility.[9] IL_{it}^{EPA} (SL_{it}^{EPA}) and IL_{it}^{KD} (SL_{it}^{KD}) are the annual counts of EPA and KDHE inspections

---

[7] Curiously enough, Gray and Shadbegian do not incorporate effluent limits assigned to each plant, which contradicts their belief that the main motivation for controlling pollution in the US is government regulation, especially for conventional water pollutants. Instead, regulatory pressure captured by the "general" threat of inspections prevailing in the state, excluding the inspections directed at the plant, was not significant in explaining abatement behavior.

[8] For BOD, monthly average concentration measured in mg/L was considered as the appropriate pollutant measure. Facility-specific limits technically restrict both quantity (measured in kg/day or lbs/day) and concentration (milligrams of BOD per liter of water). However, for Kansas's municipal facilities, only the concentration limit was operative. Some industrial facilities on the other hand might face only quantity-based limits.

[9] These three are called specific deterrence since they are directed at the facility itself.

(enforcement actions) conducted on (taken against) other major municipal facilities.[10] X_{it} includes 12-month lagged value of relative discharges, type of abatement technology (secondary as opposed to equivalent to secondary), flow capacity, aspects relating to permit conditions such as final versus interim limits and permit expiration (magnitude in days), county-level community characteristics such as population density, high school graduation rate, voter turnout, Republican voting proportion, unemployment rate, income per capita, and proportion of renter households. Two indicators for replacement of non-reported current and lagged discharges are also included to address the problem of facilities not submitting Discharge Monitoring Reports (DMRs) for each month.[11]

Equations (3.5) and (3.6) model the EPA and KDHE inspections decisions as a function of mean relative discharges in the preceding 6-month period $\breve{Y}_{i(t-6)}$, Z_{it} includes all the factors under X_{it}[12], including EPA's Index of Water Indicators, which categorizes ambient water quality for each watershed according to "better quality", "less serious problems" and "more serious problems".[13] IP_{it}^{KD} and IP_{it}^{EPA}, the predicted probability of a KDHE and an EPA inspection, are included as the instrumental variable for the current

---

[10] These four are the general deterrence measures, since they are aimed at other similar plants, except for the facility in concern.

[11] A Heckman correction procedure was used to validate the assumption that non submission of DMRs is random or at least independent of the true, yet unreported, discharge levels and other important explanatory variables of interest. Given the supported assumption of random non-reporting, missing data points were replaced with facility-specific annual mean values. However, the highly positive coefficient on the non-reported discharges indicator was interpreted as an overestimation since the facility-specific annual average level was higher than the model would have otherwise predicted based on reported discharges.

[12] Except, 12-month lagged relative discharges.

[13] In the state of Kansas, 52.5% of the communities faced "less serious problems" and 45% endured a watershed that had "more serious problems". This index was used as part of the local community characteristics presuming that it would reflect the agencies' decision to modify monitoring protocols in order to protect relatively more pristine waters or improve relatively more polluted waters.

threat of KDHE and EPA inspections in the other authority's decision.[14] Finally, seasonal factors are captured by dummy variables for winter, spring, and summer, and a time trend are included in all the intervention and performance equations.

An OLS model was estimated on a semi-log specification for the pollutant discharge model (equation (3.4)), where relative discharge is log transformed. Some of the results obtained are briefly mentioned here. Aggregate enforcement at other municipal facilities induces plants to lower discharges more than aggregate inspections at other facilities. Significant coefficients on the predicted probability of an EPA inspection and one year lagged enforcement were lost when Random Effects estimation method was used (as expected).[15] Finally, plants with final permit limits (plants with expired permits for a longer time period) had a (positive) negative effect on environmental performance. The significantly positive coefficient on the time-invariant index of ambient water quality in the relative pollutant discharge model (equation (3.4)) was interpreted (by Earnhart) as higher relative discharges resulting in lower ambient water quality, in the same time period.[16]

In addition to the factors identified in equation (3.4), Earnhart (2004a) and Earnhart (2004b) control for differences in the effluent limits faced by each plant in the sample. The mean BOD permit level was around 30 mg/L, and they varied across

---

[14] Lagged facility performance was not found to be a significant determinant in either the EPA or the KDHE inspection decisions. The predicted probability of EPA and KDHE inspections did not exert a significant impact on the other agency's inspections decision.

[15] Earnhart concedes to the possibility of upward bias in the context of interpreting the strong evidence in favor of effectiveness of enforcement actions (both general and specific) targeted towards these public facilities. He mentions that since municipal plants face less enforcement pressure in general, even small increases in aggregate enforcement directed at all the other municipal facilities in the state would result in a substantial increase in the threat of enforcement. Similarly, since these plants face infrequent enforcement activities directed specifically at them, imposition of a sanction may induce a strong reaction by the plant in terms of substantial reductions in its pollutant discharges.

[16] This index did not exert a significant impact in the government intervention equations.

facilities, across years, and within years.[17] Curiously, monthly average effluent limits exerted a negative and significant impact on relative BOD discharges in both these studies. Presence of important unmeasured factors could not be ruled out since the plant level fixed effects was hugely significant. The results obtained were similar except for lower significance in the impact of regulatory deterrence factors (as well as some community features).

Earnhart (2007) estimates the set of equations (3.4), (3.5) and (3.6) including (separately) the differences between the actual effluent limits imposed on a particular facility for a given month and three benchmarks (as opposed to the permits themselves). These are: federally mandated standard, a particular facility's average effluent limit over a specified period (e.g., entire sample period), and preceding monthly limit at a particular facility (e.g., transition to new limit). Earnhart tries to characterize the nature of abatement technology: the underlying hypothesis is that if facilities cannot smoothly or quickly adjust treatment in response to fluctuating limit levels they might choose to overcomply when limits are less stringent, while merely complying when limits are more stringent.

Some of his results are summarized here. If effluent limits that were below federal standards were to be reduced further, i.e., made more stringent, then discharges relative to permit levels would increase implying that compliance levels would fall. Earnhart interprets this as proof of compliance costs rising with limit stringency. If limits

---

[17] According to KDHE officials, BOD limits were sometimes lowered to address ambient surface water quality concerns associated with dissolved oxygen (Earnhart 2004b).

that were above the facility's mean effluent limit over the period of a specific limit type[18] (embedded within a given 5-year permit) were to be increased further, i.e., made less stringent, then relative discharges would be significantly reduced. On the contrary, if limits that were below the facility's average limit over the period of a specific limit type were to be made more stringent, then performance would not be 'affected'. The significantly negative coefficient on permit transitions[19] demonstrated that more stringent limits, at the time of a new permit, increase discharges relative to limit levels, since facilities do (can) not adjust their treatment fully to the new lower limits. These results were interpreted as indications of lumpy treatment since facilities maintained their performance level when limits were relatively more stringent and allowed their performance to improve when limits were relatively less stringent.

Other studies explore factors apart from regulatory 'presence' and community characteristics that explain overcompliance behavior of manufacturing plants as well as municipal facilities. For pulp and paper plants, the possibility of zero marginal abatement cost and stochastic emission patterns are investigated in McClelland and Horowitz (1999) and Brannlund and Lofgren (1996), respectively.

Bandyopadhyay and Horowitz (2006) find evidence of randomness in pollution discharges driving manufacturing and sewage treatment plants to overcomply apart from factors such as public pressure. Their data consists of 764 plant level monthly average BOD concentrations in wastewater over the period 1992 to 1999. They define discharge rate as the ratio of the monthly reported discharge to the permitted level. Plant specific

---

[18] Within the life of a five-year permit, agencies may impose interim limits, which serve as a transition to the final limits, which are generally more stringent, or agencies may impose final limits immediately. The 40 major treatment plants in Kansas faced interim limits only 4% of the time from 1990 to 1998.

[19] Limit type transitions provide the most 'lead time' (1 to 3 years), while permit transitions to a new limit provide the least, with seasonal transitions in between.

median and standard deviation of the log of the discharge rate is then calculated, denoted by m and σ, respectively. Their dependent variable was, $y = m/\sigma$, called the standardized discharge rate, which is median of the log of discharge rate “corrected” for variability. The explanatory factors included dummy variables for EPA vs. state regulated plants, EPA region, manufacturing vs. municipal wastewater treatment plants; size of plant, and zip-code level socioeconomic demographics were the continuous variables included. The standardized discharge rate was found to be higher for sewage treatment plants, smaller size plants and in low median household income and higher minority population areas. They also predicted that plants were over-complying by about 40% on average, even when uncertainty was controlled for.

On a similar note, Houtsma (2003) identified process related reasons for overcompliance amongst POTWs. Houtsma conducted a survey on treatment plant superintendents of Canadian and privately operated WWTPs in the US, and public sector officials. Results of the survey led him to conclude that plant operators did not exert enough control over the wastewater treatment process to allow for higher effluent levels.

Houtsma’s (2003) survey also revealed that WWTPs are expected to meet standards “all the time,” not just on a weekly or monthly basis; over and above the more obvious reasons related to the operator’s desire to keep his/her job or be a good public citizen i.e. plant managers’ concern for water quality and for the environment in general. Other studies have also shown that management preferences for increased environmental protection can lead to reductions in effluent discharges and overcompliance (Doonan et. al., 2002; Kagan et. al., 2003). McClelland and Horowitz (1999) noted that the overcompliance amongst the plants in their dataset may be due to management’s desire to be good neighbors; citing one mill manager's "good neighbor policy" which results in overcompliance at his plant.

Akobundu (2004) estimated a model where BOD discharge is regressed on nitrogen discharge by the plant (*Nitrogen* in equation (3.7) below) since there exists a synergistic relationship between nutrients removal technology and BOD control in the plant's effluent (Stoddard et. al., 2002; Hammer and Hammer, Jr., 1996; Gray, 1999).[20]

$$\begin{aligned} BOD = \beta_0 + \beta_1 Nitrogen + \beta_2 BOD\text{limit} + \beta_3 FLOW + \beta_4 Excesscapa\ city + \\ \beta_5 VarianceBO\ D + \beta_6 SkewnessBO\ D + u \end{aligned} \tag{3.7}$$

Akobundu used monthly data on BOD, $BOD\text{limit}$, plant design flow ($FLOW$), and monthly flow data (to construct the variable *Excesscapacity* i.e. the excess of the plant's design flow over its actual flow) on 171 wastewater treatment plants in Maryland, for the time period 2001-2003. *Variance* of the BOD discharge distribution was included to capture randomness. *Skewness* was included with the expectation that the more skewed the distribution the lower will be the probability of exceedance. The corresponding monthly nitrogen discharge data was available for 133 plants from the CBP office's Point Source database.[21] The variables that entered the empirical model were calculated by collapsing the panel of 2 years of data on 133 plants to form a cross-sectional dataset. The median values of all the explanatory variables (except, excess capacity, variance, and skewness) were log transformed.

The 'BOD elasticity of nitrogen control' was statistically significant at the 1 percent level with an estimate of 0.25 meaning that a 1 percent reduction in nitrogen

---

[20] Specifically, one of the processes of biological nitrogen removal technology involves denitrification of the secondary treated wastewater by introducing denitrifying bacteria, which in turn require anoxic conditions and a suitable organic carbon source in order to convert nitrates to nitrite and finally to nitrogen gas.

[21] Connecticut and Maryland were chosen since these two states suffered from the long-standing problem of excessive nutrient discharges to their surface waters.

effluent discharge would result in a 0.25 percent reduction in the BOD effluent discharge. The coefficient on the variance of BOD output was negative and marginally significant at the 10 percent level with a p-value of 0.07 (only for the Maryland data). The BOD permit limit was statistically significant at the 1 percent level, with a coefficient estimate of 0.78, indicating that a 1 percent reduction in the BOD permit limit translated into a 0.78 percent reduction in the BOD discharge level, on average. Both skewness and excess capacity had a negative influence on BOD discharges.

To summarize, empirical papers (such as the ones discussed above) found evidence on a wide range of factors that influenced the pollutant discharge behavior of sewage treatment as well as manufacturing plants in the US. These were: one year lagged pollutant discharge (as a measure of current abatement technology since proof of its lumpy nature was found), effluent limits, plant capacity or design effluent flow, type of production facility, lagged inspections and enforcement actions directed against the plant, predicted probability (threat) of a current period inspection, general regulatory presence in the state, and location specific i.e. ‘affected’ community features. Moreover, the literature explored a set of determinants mostly related with aspects of abatement technology (e.g. lumpiness, randomness and jointness) that could explain the overarching evidence on plants discharging pollutants much below their permit levels.

A Model of Pollution Abatement

The following two-period model captures the mechanism through which the polluter’s discharge decision is influenced by ambient water quality. Each sewage treatment or manufacturing plant minimizes its costs of abatement $C(.)$ and expected fines of non-compliance $F\left[\Pr(I)\right]$ at a period of time. $\Pr(I)$ is the current period threat or probability of an EPA or a state inspection, F is a constant penalty arising out of non-compliance, e.g., administrative fines as well as legal proceedings. Plant level discharges in period 1 determine ambient water quality in the same time period and hence permit levels in period 2. In practice, such a situation arises when water quality has not been meeting the standard for a long enough period of time. Plant level discharge in period 1 BOD_1 directly affects the probability that ambient water quality might not meet the standard in the next period i.e. $DO_1 + A\left(\sum_{i=1}^{K}\eta_i BOD_{i1}, \sum \alpha_n NPS_n\right) \leq \overline{DO}$. Hence, permit level in period 2 P_2 may be made more stringent in order to ensure standard is met. This adjustment factor essentially raises the future expected costs of non-compliance. DO_1 is the ambient DO level at the start of period 1. Ambient DO level at the start of period 2 is then determined by the additional BOD discharges from all (K) point and non point sources (NPS_n) during period 1. η_i s and α_n s are the transfer coefficients associated with each point and non point source of pollution.

$$\min_{BOD_1}\begin{bmatrix} C(BOD_1, P_1, Z, W) + F[\Pr(I_1[BOD_1 - P_1])] + \\ \\ (1+\delta)^{-1}\{C(BOD_2, P_2, Z, W) + F[\Pr(I_2[BOD_2 - P_2])]\} \end{bmatrix} \tag{3.8}$$

where the adjustment in the permit levels is given by :

$$P_2 = P_1 - \gamma\left[\left\{\Pr\left\{DO_1 + A\left(\sum_{i=1}^{K}\eta_i BOD_{i1}, \sum \alpha_n NPS_n\right) \leq \overline{DO}\right\}\right\}\right]$$

Thus, the regulator's adjustment of each plant's permitted discharge level depends on the probability that ambient water quality does not meet the DO standard. Hence, ambient water quality enters indirectly in the abatement cost and expected fine functions due to its influence on effluent limits. In particular, changes in the status of ambient water quality might lead to a change in the effluent limit, which has an impact on expected costs of non-compliance and threat of regulatory interventions. Z is plant specific characteristics such as size, age and type of plant, and W incorporates location specific attributes.

Ceteris paribus, higher permitted levels of pollution imply that plants are likely to pollute more since abatement is costly. Firms with larger design capacity are expected to pollute more and hence might have higher costs of abatement.[22] The type of the plant is another factor that determines its abatement costs.[23] Ownership structure of the plant might also be a significant factor in driving abatement costs.[24] Lastly, location-specific features such as socioeconomic demographics of the 'neighborhood' population and land use characteristics[25] of the surrounding area are other 'exogenous' factors affecting plant level abatement costs.

Since the role of ambient water quality is the focus of this analysis the justification for including it in the abatement decision of the polluter warrants further

---

[22] Larger plants have to internalize the costs of higher volume of discharges by installing better abatement technology or intensifying O&M processes such that the chances of being in violation with respect to their permitted limits are reduced. On the other hand, there might exist economies of scale with respect to pollution abatement, in particular, newer and bigger plants might have better abatement technology. Hence, age of the plant or the vintage of the abatement technology is another factor that would be influencing abatement costs. Older plants would have less-efficient pollution control infrastructure in place, and hence incur higher costs of abatement for the same level of pollutant discharge in contrast to a newer plant with more efficient technology.

[23] POTWs might have higher abatement costs since they cannot be as responsive as manufacturing firms in terms of adjusting their treatment process to reduce discharges because they treat wastewater generated by the households and some industries (Bandyopadhyay and Horowitz 2006). However, they can also pass on the financial burden of higher pollution control costs to the local community by raising user fees.

[24] Waste treatment plants are not 'profit maximizing'; for example, state revolving funds were used to finance the construction of these projects in order to meet ambient water quality standards. By contrast, the financial burden related to the abatement of pollution for manufacturing plants depends on how much can be passed on to the consumers (in the form of higher prices).

[25] For a primarily rural area, the small number of point sources i.e. sewage treatment and manufacturing plants means that they probably do not contribute as much to low levels of DO relative to agricultural run-offs (both nutrients and BOD). It implies that these plants' abatement costs (ceteris paribus) might be lower than the ones located in densely populated urban (and industrial) areas. A larger number of point sources in urban areas mean that their contribution to low ambient DO is higher relative to urban run offs implying that they would have to abate more (in order to bring down ambient pollution levels) and hence incur higher abatement costs.

discussion. The central idea behind the hypothesis that point source polluters respond to ambient water quality, over and above their existing permit levels, is that plants want to 'minimize' the expected threat of interventions and pecuniary penalties. In other words, plants make a strategic decision to incorporate ambient water quality considerations on a daily (on an average, monthly) basis rather than face the costs of potential noncompliance.

To elaborate, consider a situation when a plant ignores its receiving water quality, which has been poor over a certain period of time (the beginning of its permit cycle e.g.). This means that, ceteris paribus, the same amount of pollutant discharge by the plant would have the potential of producing higher damages. Specifically, poor water quality would have immediate implications for the plant since an increased frequency of inspections is anticipated; thereby, raising its costs of abatement. For instance, Dion et al (1998) find evidence of greater inspection effort being allocated towards plants whose discharges are likely to generate higher level of damages. In their paper higher impact from discharges is measured by the flow of the plant's effluents relative to the flow of the river in which the effluent is discharged.[26]

Essentially, the regulator would want to verify (through monitoring activities), that the plant is not "directly" responsible for the low levels of ambient water quality: by discharging pollutants exceeding its permitted amounts.[27] In the situation that the plant 'happens to be' discharging below its permitted level the potential threat of inspections is

---

[26] The coefficient on relative flow was positive and statistically significant when the variable was lagged 3 or 4 periods. The time lag between the plants submitting their discharge reports and the time when this information becomes available to local regulators for action explained why only lagged relative flow values were significant.

[27] Since point sources are easier to monitor and validate performance (in contrast to non-point sources), designated inspectors would immediately target the major polluters.

raised, nevertheless, for a stream segment with poor water quality (adding to its future abatement costs). In case the plant is found to be non-compliant, a series of enforcement actions is the outcome, which adds to its abatement costs immediately. Subsequently, fines are imposed if the plant cannot take actions to bring itself into compliance with its effluent limits (by further raising its abatement costs). Given the lumpy i.e. time consuming nature of abatement technology, the plant would probably face financial penalties if they cannot ensure a status of compliance within 6 months e.g. Significant Non-Compliance (SNC) status and potentially face closure of operation.

In response to a poor status of ambient quality, and irrespective of the plant's compliance status, the regulator is obligated to undertake a water quality analysis (WQA) or a TMDL.[28] A case in point: during the time period of this analysis, water quality based BOD5 limits that were derived from 'non-TMDL waste load allocations' were actively implemented.[29] Essentially, this meant that the plant would have to face more stringent limits to accommodate higher non-point source pollution, in the situation that the latter might have been accountable for the deteriorating water quality. For example, seasonal limits which were prevalent during the nineties meant that plant operators were obligated to tighten BOD control at certain times of the year, in particular, discharge less BOD in the summer; specially, if they wanted to avoid penalties associated with non-

---

[28] A WQA is completed when current monitoring data support delisting a waterbody as 'impaired' because water quality standards are being attained. TMDLs are completed when a waterbody continues to violate water quality standards. A TMDL establishes the amount of a pollutant, plus a margin of safety, that a waterbody can assimilate and still attain water quality standards.

[29] Given the technology-based effluent limits of a plant, a dissolved oxygen sag analysis was conducted where the lowest concentration of ambient DO realized under 'critical' high temperature, low flow conditions was simulated. If the DO level generated did not meet the ambient standard of 5 mg/L to maintain aquatic life, water quality based limits were invoked.

compliance.[30] Hence, plants would anticipate an increase in its expected costs arising out of non-compliance once the more stringent, seasonal limits would be imposed since they cannot adjust immediately their abatement levels in response to permit level changes.

In reality, a plant is concerned with minimizing its present value of costs of abatement and expected costs of non-compliance; thereby, implying that it would be cost effective in the long run for it to adjust its regular discharge decisions in response to the prevailing ambient water quality. In particular, attenuate a poor 'status' of downstream quality in the previous time periods, as opposed to raising its expected costs arising out of penalties of non-compliance, in the situation that a lower permit is imposed on the plant in a future time period.

Plants incorporate past water quality, most likely, irrespective of the type of permit faced by the plant (technology or water quality based). Specifically, if a more stringent permit based on water quality considerations has not already been imposed on a plant, the operators of the plant will have to be cognizant of the ambient conditions prevailing in the recent past to avoid future enforcement actions and penalties due to noncompliance. Alternatively, WQBELs are implemented only in circumstances when the downstream water quality is low enough to either not meet the stream segment's ambient standard,[31] or is projected to fall below it in the near future.[32] Hence, the polluter is inherently motivated to pay attention to the past periods ambient water quality, especially when these levels have been "low".

---

[30] Existence of seasonal permits indicated the fact that water quality conditions generally deteriorated during the summer due to elevated levels of dissolved oxygen (DO).

[31] This happens when ambient DO falls below the minimum diurnal standard of 5 mg/L at any point in time, even if it is so measured only once in a stretch of one year (for example).

[32] Ambient water pollution process is subject to considerable randomness: the pollutants discharges by point and non-point sources and weather related uncertainties. However, the ultimate ambient water quality recorded in the rivers does not fluctuate as much mostly due to the 'absorptive' capacity of the waterbodies.

Solving for BOD_1:

$$BOD_1 = BOD(P_1, C, \Pr(I_1), Z, W) \tag{3.9}$$

Thus, the abatement/discharge of an industrial/treatment plant depends on its permitted level of BOD discharge (concentration or quantity), costs of abatement, past period ambient DO, current period threat of regulatory intervention, expected enforcement actions and penalties of noncompliance, and other plant specific attributes, local demographics and land use characteristics which are time-invariant.

If ambient standard is met, one can anticipate a positive relationship: all else equal, if lagged downstream DO is reduced by one unit, then the likelihood that a more stringent effluent limit will be imposed are higher, which implies that the polluter's expected costs of non-compliance are raised, in turn, inducing the plant to increase its level of abatement (i.e. lower discharges). Alternatively, if lagged downstream DO is higher by one unit, in a stream segment that meets its designated use, then the likelihood that a less stringent limit will be chosen are higher, since the regulator knows that abatement is costly to plants[33]; thereby, leading plants to lower abatement and discharge higher pollutants.[34]

---

[33] Additionally, evidence on exceptions to anti-backsliding, for streams meeting ambient standards have been found from examining the change in quantity permits decision in the previous chapter. For concentration permits, however, some evidence on anti-backsliding was found (except for average water quality prevailing during the most recent year of past permit cycle) since permits were not revised in response to status of water quality, in particular, less stringent limits were not chosen when water quality improved by a certain amount. Hence, only for quantity discharges, can one anticipate a positive sign on the measure of lagged water quality.

[34] It could be argued, at least theoretically, that even for stream segments meeting their existing uses, better water quality might mean more stringent limits; thereby, implying a non-positive relationship between past water quality and current discharges. In particular, for 'higher' status waterbodies if 'substantial economic impact' on the polluters is not found, stringent permit levels are adopted to maintain the standards that have been achieved. Specifically, if the economic costs of meeting more stringent permits (financial burden to the polluters and social consequences on the local population in terms of jobs lost etc) are not *exorbitantly* higher than the total benefit to the affected community derived from a higher level of pollution abatement, then lower permit levels for the individual polluters are adopted. However, proof of widespread prevalence of such practices, by the regulators, could not be gathered.

An Empirical Pollution Abatement Model

The purpose of this chapter is to investigate if abatement decisions of plants are influenced by ambient water quality considerations. Hence, the reduced form equation (3.9) that identifies the determinants of plant level BOD discharges becomes the starting point of this empirical research. In this chapter, discharges relative to their effluent limits is considered as the appropriate dependent variable in order to allay concerns of endogeneity of limits, specially since the time period covers multiple permit cycles for a plant. The dependent variable, which is monthly average BOD5 discharge measured in concentration or quantity, relative to its respective effluent limit, has been log transformed.[35] Correspondingly, the explanatory variables have also been log transformed in order to be able to interpret the coefficients in percentages. The equation estimated is:

$$\begin{aligned}\ln\left(BOD_{ijt}\big/P_{it}\right) &= \alpha + \beta\ln\left(laggedaverageDO_j\right) + \tau\ln\left(designflow_i\right) + \rho\hat{I}_{it} + \pi sic_i + \\ &\sum_{d=1}^{5}\lambda_{id}\ln demographics_{id} + \nu_y + \theta_s + \sigma_R + \varepsilon_{it}\end{aligned} \qquad (3.10)$$

The polluter i, while deciding on its pollutant discharge BOD_{ijt} relative to its permitted level P_{it} in month t incorporate average dissolved oxygen DO_j prevailing at monitoring location j, downstream to its point of outfall, and in a past time period.[36] It is included because the CWA mandates that if water quality standards will not be met in the

---

[35] BOD5 has been well documented in the literature (Bandyopadhyay and Horowitz, 2006) to follow a log normal distribution.

[36] The underlying assumption behind equation (3.10) is that the elasticity of discharges with respect to permits is equal to one. However, upon regressing absolute discharges (concentration and quantity) on past water quality, permits, etc, the coefficient on log of permits is less than one (0.547 and 0.726 as seen in columns (1) and (3) of Table A 47 in the appendix). Given that abatement technology is inflexible (Bandyopadhyay (2002); Earnhart (2007)), plants cannot adjust their discharges to the same extent as a change in their permit levels. Hence, one expects a coefficient of less than one (also seen in Table A 48 and Table A 49 for the sample of plants with at least 25 and 50 percent monthly observations).

(near) future plant level effluent limits would be made more stringent. In fact, adjustments to permitted effluent levels for BOD5, for existing sources, were made even when the resulting water quality did not fail to meet the fishable/swimmable goal, but was projected to engender a failure to meet its standards.[37] Past ambient water quality is relevant for the current period effluent discharge decision of the plant operator because of the lag with which water quality records are made publicly available. Moreover, even if water quality records were punctually released, the plant, most likely, does not possess sufficient flexibility in its abatement technology to adapt to the downstream DO prevailing in the same month. This is especially true for, the operation of, municipal plants.

Plant size is proxied by $designflow$, which has been found to be an important determinant of monthly abatement choice of a polluter (absolute and relative discharges). The predicted threat of inspection in the current month t is $\hat{I}_{it}$. Since the polluter has no foreknowledge regarding the regulator's inspection decision in the same month t, at best it can incorporate a predicted probability of being inspected in its abatement decisions.[38] The SIC-code of each plant is represented by sic, which denotes the type of the plant: sewage treatment, or manufacturing (food, paper and pulp, chemical etc), or electrical services. Location specific variables captured by the zip code (in which plant i is located)

---

[37] For example, a TMDL for CBOD and NBOD was approved by the MDE for the Antietam Creek watershed in 2002. Based on water quality data collected in Antietam Creek during 1996 and 1997, the DO water quality standard was met at that time (USEPA, 2002). However, increased loadings of CBOD and NBOD beyond the current allowable point source loadings and current background conditions could lead to a violation of the DO water quality standard.

[38] Current period inspection is endogenously determined in the sense that the regulator most likely decides on whether to inspect a facility or otherwise, based on factors that also influence abatement decisions of the plant. Assessing regulatory effectiveness through its enforcement and monitoring activities is not the focus of this chapter; nevertheless, current period threat of regulatory interventions such as inspections are expected to be a significant deterring factor in the relative discharge decisions of polluters, as seen in the evidence presented in past studies. Hence, a simple model is presented in the appendix along with the estimation results.

level socio-demographic characteristics: $demographics_d, d = 1,2..5$.[39] Yearly dummy variables: $v_y, y = 1,2,...13$ are included in order to capture any time specific trends of discharge behavior. Seasonal controls for winter, spring and fall (compared to summer months) are captured by: $\theta_s, s = 1,2,3$. State level dummies $\sigma_R, R = 1,2$ are included in the model to control for potential differences in monthly average performance of plants in MD and PA, in contrast to those in VA.[40] [41]

Finally, omitted variables, and/or unobserved aspects (related to the costs of abatement of the plant, e.g.) are bound to be a specification problem in the above model, since the dependent variable is monthly average (relative) discharges. A case in point: most of the previous studies that model abatement behavior find that one year lagged BOD5 discharges (absolute or relative), which is used as proxy for the abatement/production technology (vintage, etc), almost perfectly determine current month discharges.[42] In the model (equation (3.10)) above, such factors leading to plant specific discharge levels are not captured; thereby, implying that monthly discharges is

---

[39] The characteristics of the "local" population considered are: proportion of non-whites, median household income, proportion of people carpooling to work, proportion of people employed by the manufacturing sector, and total population. However, the level and significance of each one of these coefficients are often credited with the limitations of multicollinearity since these indices are highly correlated with each other. Using broader zip code level data is preferred to considering the demographic characteristics in the immediate vicinity surrounding the plant's location, since concerns of reverse sorting are allayed in the situation that wealthier people chose to live further away, while the poorer chose to live in the 'neighborhood' of the facility.

[40] For instance, state level controls might be capturing the effects of differences in total inspections (enforcement actions) conducted on (taken against) "similar" facilities during the past year and in the same state. In particular, according to authors such as Laplante and Rilstone (1996), "[s]ignificant coefficients on regions [….] indicate that there might be important regional differences in the nature of the relationship between the regulator and the regulatees and/or the monitoring and enforcement procedure across regions" (p.27).

[41] Ownership structure of a plant i.e. public vs. private is not included in the model because once the 'balancing' criteria are applied the number of plants that are private falls substantially: so as to not be able to include this indicator along with the SIC code dummies.

[42] According to Laplante and Rilstone (1996) this variable "reflects the fact that the installation of emissions control equipment typically requires a long time. To this extent, the lagged pollution variable could also be interpreted as a proxy for the production technology" (p.26).

expected to be correlated within the plant and over time. Plant level dummy variables, which would have controlled for (at least) the time-invariant plant specific factors that are not modeled, are not included; thereby, alleviating concerns related to within plants serial correlation of the error terms. The measures of lagged average water quality and design flow are, by and large, plant specific attributes in the current dataset, and hence are not be included along with plant dummies.

Data

The sample is pooled (time series and cross section) over 100 "major" NPDES plants and the years 1990 to 2003, on a monthly basis. The distribution of these plants was 52 in VA, 26 in MD, and 22 in PA.

Plants regulated by effluent limits on their BOD pollution have to monitor and assess the BOD content in their effluent discharges. The results of the BOD-5 day test are then reported on a monthly basis to the EPA or the State. The EPA's Permit Compliance System (PCS) database provides effluent discharge data of BOD5 from the monthly DMR filed by major sewage treatment and manufacturing plants that face a BOD5 limit in terms of concentration (in mg/L) and/or in quantity (in lbs/day).

There are 86 (94) plants with BOD5 concentration (quantity) monthly discharge data.[43] All 54 of the sewage treatment plants[44] face both concentration and quantity

---

[43] Monthly observations on BOD5 discharges that did not have a corresponding monthly limit data automatically dropped out of the sample. Absence of permitted levels could be a consequence of data entry error, or delay in a new permit level being assigned by the regulator (Earnhart, in his various papers, incorporate an indicator of the number of days a permit was expired), or a plant facing an interim limit before the final (usually more stringent) permit becomes effective, or due to regulatory requirements under which plants do not face a limit on their pollutant effluents in a certain month (although, instances of such a requirement could not be confirmed by permit writers such as Kyle I. Winters; email communication).

[44] There are 54 sewage treatment plants, 6 power generation plants, 4 food and kindred products, 4 textile mill products, 7 paper and allied products, 10 chemicals and allied products, 6 petroleum refining and related industries, 1 rubber and miscellaneous plastics products, 2 leather and leather products, 1 fabricated

limits and hence report both discharge measures to the regulator.[45] On an average, these plants seem to be over-complying for relative discharges of both concentration and quantity measures. Table 3.1 below summarizes the data on the dependent variables as well as the explanatory factors included in the estimation exercises.

Table 3.1: Summary Statistics of the Discharge, Past Period DO, Design Flow, and Inspections Data

Variable	Mean	Min	Max
Monthly BOD5 concentration discharge, mg/L	9.04	0.1	713
Monthly BOD5 concentration permit, mg/L	28.52	5	111
Monthly concentration relative to effluent limit, mg/L	0.33	0.002	27.42
Monthly BOD5 quantity discharge, lbs/day	451.69	0.06	40522
Monthly BOD5 quantity permit, lbs/day	1452.62	2.6	42400
Monthly quantity relative to effluent limit, lbs/day	0.27	0.0002	21.06
Past year mean (median) downstream DO, mg/L	9.55 (9.42)	5.64 (4.62)	13.3 (13.6)
Past two years mean (median) downstream DO, mg/L	9.56 (9.42)	6.05 (5.5)	12.37 (12.6)
Past three years mean (median) downstream DO, mg/L	9.56 (9.40)	6.26 (5.45)	12.38 (12.5)
Design flow, million gallons/day	54.89	0.03	2000
Monthly inspections count, by plant	0.24	0	13

The explanatory variable that is the focus of this analysis is a measure of lagged average ambient water quality downstream to the plant i's discharge point and measured in concentration of dissolved oxygen (DO) (mg/L).[46] Different lag lengths have been considered; in the absence of any prior knowledge, e.g., from surveys conducted on the plants that would indicate the appropriate time period. A change in the status of ambient

---

metal products, except machinery and transportation equipment, 1 transportation equipment, 1 justice, public order, and safety, and 3 national security and international affairs.

[45] Abnormally high values, specifically, with respect to the distribution of discharges of each plant over time, have been dropped from the regression sample. 7 and 3 such observations for BOD5 concentration and quantity measures were identified as outliers. Table A 23 and Table A 30 present the summary statistics by each plant whose outlier has been excluded from the estimation. The corresponding figures give a visual picture of the basis on which these observations were replaced as missing, most likely due to data entry error. Table A 33 and Table A 34 in the appendix summarize the BOD5 concentration and load variables across all plants and years, while Table A 35 and Table A 36 show the corresponding values for the concentration and quantity permits.

[46] Details on how the downstream monitoring stations were identified based on the location of the plants, using GIS, is described in the chapter on Water Quality Model.

water quality is most probably not an instantaneous process mainly due to assimilative capacity of the stream; consequently, past period water quality over one,[47] two or three years might be more relevant for plants. In addition, they might be mimicking the regulators who make assessments of status of the stream over similar time phases. Specifically, average (mean/median) DO level prevailing during the past, one, two and three-year time period are included (separately) as a potential determinant of relative discharges.

The past year average water quality is considered because plants might be concerned about the water quality prevailing in the immediate past months. However, instead of looking at a segment of the past year, e.g., most recent past six months the annual average is calculated to allay seasonality issues of the underlying DO data.[48] The past two and three years lag lengths are included based on the historical 303(d) list of impaired waters which are updated every two years (in Maryland e.g.); hence, water quality problems would ideally need be addressed by the regulators (by either a WQA or a TMDL) within the 2 years listing cycle. The Chesapeake Bay Monitoring Program (CBP) also tracks water quality status for a period of two years.[49]

---

[47] For example, previous studies have found empirical evidence that polluters can adjust their abatement levels in response to average regulatory presence e.g. inspections conducted during the past quarter. The general hypothesis behind this finding was that plants tend to reduce discharge levels and enhance compliance through immediate attention to better plant operation and maintenance rather than long-term capital investments. Thus, empirical evidence suggests that plants are flexible enough to incorporate changes in frequency of inspections in the immediate past quarter. Appealing to the same logic, plants would be able to respond to, and hence incorporate changes in the average water quality prevailing during the past year, through immediate changes in their operations and maintenance procedures.

[48] For example, BOD5 monthly discharges for the month of January 1991 is regressed on median DO prevailing during the months of January 1990 through December 1990 (a lapse of 12 months in total). Correspondingly, BOD5 discharges for August 1992 is regressed on median DO prevailing during the months of January 1991 through December 1991.

[49] Water quality trends, on the other hand, are assessed over a much longer period: usually around twenty years, or in some cases short-run trends over ten years are looked at (MDDNR, 2008; its documentation can be found in Ebersole and others (2002)). For a specific application of detection of trend and analysis over a period of 40 years, see Forrester (2000).

Table 3.1 above describes the mean and median values (as part of the robustness checks on the estimated impact) of downstream water quality prevailing during the past one, two and three years.[50] The similarity in the averages lead one to infer that ambient water quality (records) did not change drastically, over the time period that is examined (at least), as long as they are reasonably close to one another namely one, two or three years. Second, the averages were well above 5 mg/L (4-6 depending on temperature): the minimum standard required for maintaining aquatic life.

Average design flow, which did not vary over time, for any one of the 97 plants, with data available, is 55 mgd.[51] On an average, a plant was inspected at least once every 5 months. The socio-demographic variables are discussed in the previous chapter.

Results

Equation (3.10) repeated below is the model that is estimated, where log of the monthly average discharge relative to the effluent limit of each plant is regressed on log of average water quality in past time periods, log of design flow, predicted threat of an inspection, type of plant or industry, socio-demographic controls, and year, season and state level dummy variables. Concentration and quantity measures of BOD are estimated separately.

---

[50] For details on the sources of monthly dissolved oxygen data recorded at various monitoring stations across the three states of Maryland (MD), Virginia (VA) and Pennsylvania (PA) please refer to Chapter 1. Six observations on downstream water quality were dropped from the sample; for details on identifying the outliers please refer to Chapter 4.

[51] It is the average flow in millions of gallons per day that a permitted facility was designed to accommodate. This information is available from the EPA’s PCS database.

$$\ln\left({BOD_{ijt}}\big/{P_{it}} \right) = \alpha + \beta \ln\left(laggedaverageDO_j\right) + \tau \ln\left(designflow_i\right) + \rho \hat{I}_{it} + \pi sic_i + \sum_{d=1}^{5} \lambda_{id} \ln demographics_{id} + \nu_y + \theta_s + \sigma_R + \varepsilon_{it}$$

The estimation method is FGLS i.e. allowing the error terms within plants to be serially correlated (as well as heteroskedastic) along with annual dummy variables.[52] The tests for these consistency and efficiency issues are reported in rows (7), (8), (9) and (10) of Table 3.2 which present a summary of the results of the relative concentration and quantity discharge models.

Missing data on monthly average discharge measure of concentration (or quantity), and/or effluent limits is another problem with the data. Hence, in addition to estimating the model on the entire sample (with available data) it is subsequently fitted on two sub-samples with at least 25% and 50% of monthly observations on relative discharges. Essentially, such a 'balancing' of the panel of plants tests the robustness of the estimates. Previous studies have dealt with this problem either by testing the randomness of non-reporting behavior of plants and by 'instrumenting' the missing data points with facility level annual average discharges along with a no discharge indicator. Plants have been found to miss reporting randomly, and controlling for missing data has often yielded a highly significant and positive coefficient on the non-reporting indicator (see Earnhart's discussion).

---

[52] Curiously enough none of the previous empirical studies, discussed in this chapter, address or test for serial correlation in their monthly pollution discharge models. Moreover, one year lagged absolute or relative discharge is used frequently as a proxy for plant level abatement technology in the absence of readily available information on abatement costs of plants.

Table 3.2: Log of Ratio of Monthly BOD5 Concentration to Effluent Limit Models Controlling For Serial Correlation (Estimation Technique: FGLS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	***Lagged DO measures***	Past three years mean (median)			Past two years mean (median)			Past one year mean (median)		
(2)	***Sampling Criterion***	Entire sample	At least 25 percent monthly observations	At least 50 percent monthly observations	Entire sample	At least 25 percent monthly observations	At least 50 percent monthly observations	Entire sample	At least 25 percent monthly observations	At least 50 percent monthly observations
(3)	# of plants	77	60	41	77	60	41	77	60	41
(4)	***Log of downstream DO***	1.301** (1.257**)	1.411** (1.411**)	2.001** (1.903**)	1.132** (1.163**)	1.221** (1.269**)	1.779** (1.795**)	0.986** (0.797**)	1.067** (0.897**)	1.219** (0.997**)
(5)	***Log of design flow***	0.078** (0.075**)	0.019 (0.013)	0.022 (0.014)	0.077** (0.075**)	0.020 (0.018)	0.029 (0.021)	0.079** (0.079**)	0.027 (0.027)	0.059** (0.061**)
(6)	***Predicted inspections***	-1.401* (-1.312*)	-0.831 (-0.651)	-0.588 (-0.268)	-1.125+ (-0.969)	-0.608 (-0.394)	-0.861 (-0.466)	-0.858 (-0.883)	-0.368 (-0.332)	-1.352 (-1.338)
(7)	Wooldridge test for autocorrelation in panel data	111.279	93.884	74.426	122.954	104.973	79.819	127.071	107.883	80.938
(8)	*Prob > F*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*
(9)	Breusch-Pagan test for heteroskedasticity	1.86	5.93	0.05	3.98	10.20	0.33	4.41	10.69	0.91
(10)	*Prob >chi2*	*0.1730*	*0.0148*	*0.8255*	*0.0461*	*0.0014*	*0.5676*	*0.0358*	*0.0011*	*0.3411*
(11)	Joint test for log of Sociodem.	69.82	119.73	54.73	68.46	122.29	65.05	69.98	125.66	87.49
(12)	*Prob >chi2*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*

Note: The test statistics for autocorrelation, heteroskedasticity, and joint test of log of demographics are reported for the mean values of past water quality.

Table 3.2 above presents the results of regressing log of relative discharges of BOD5 concentration on (the three measures of) log of lagged average values of ambient DO, design flow, and current period predicted number of inspections, along with other plant and location specific controls.

The measures of past water quality exert a positive impact on monthly relative concentration discharges. If the status of downstream water quality declines by one unit, implying that it is closer to the ambient standard, then the likelihood of a water quality evaluation, conducted by the permit writer, revealing that the standard will not be met in the future is higher. For the polluter it means that the chances of facing higher intervention activities ultimately resulting in more stringent effluent limits are higher. In response, the plant reduces its monthly concentration discharge, relative to its permit, to reduce its (adverse) impact on ambient water quality. Conversely, given that ambient standards are generally met by the sampled locations, if lagged downstream DO is higher by one unit, then the likelihood of facing less stringent permits are higher, since abatement is costly to plants.

The coefficient on three year lagged average DO measure(s) in cols (2), (3) and (4), of row (4), is bigger in magnitude than the past two (cols (5), (6) and (7), of row (4)) and one (cols (8), (9) and (10), of row (4)) year averages (with a smaller difference from the past two years averages). The coefficient in col. (3) and row (4) can be interpreted as: if past three years mean (median) water quality is reduced by one percent, then BOD5 concentration discharge relative to its limit would be reduced by 1.411 percent (for the sub-sample with at least 25 percent monthly observations). By contrast, for the same sub-sample if past year mean (median) downstream DO is reduced by one percent, then concentration discharge relative to permit level is reduced by 1.067 (0.897) percent (col. (9) and row (4)).

Another noteworthy observation is that the estimated impact of past period water quality is larger, for the more ‘balanced’ sub-samples of the monthly observations on relative concentration discharges for each plant. For example, the coefficient in col. (5) and row (4) can be interpreted as: if past two years mean (median) DO is reduced by 1 percent, then relative concentration is reduced by 1.132 (1.163) percent, when the entire sample of plants is considered. This estimate is smaller than the coefficients for the same measure of DO, but for the sample with at least 50 percent of monthly observations on relative concentration discharges: 1.779 (1.795) percent (col. (7) and row (4)).

Size of abatement technology captured by design flow does not exert a consistently significant (positive) impact on relative concentration discharges. It has a coefficient around 0.078 percent only when the entire sample of plants is considered (and irrespective of the measure of average DO prevailing during past time periods). The only other sub-sample where it is significant is the sample of 41 plants with at least 50 percent monthly observations over the 14 years in the sample: 0.059 (0.061) percent, col. (10) and row (5).

Predicted number of inspections has a negative impact on relative concentration discharges, but the coefficient is seldom statistically significant. For the entire sample of 77 plants, inspections exert a negative (significant at 5 percent) impact on monthly relative concentration when the past three years average water quality is considered: -1.401 (-1.312) percent, col. (2) and row (6).

The socio-economic ‘variables’ were jointly significant across all three sub-samples (as reported in row (11) of Table 3.2, with their corresponding p-values in row (12)). BOD5 concentration relative to its limit also has a seasonal pattern, as seen in the controls for season, with plants discharging significantly higher during winter and spring, as opposed to the summer months, on an average. Overall, only lagged average water quality seems to exert a consistently positive (and significant) influence on the current period relative concentration discharges of a plant.

Table 3.3 below presents the results of estimating the monthly average relative discharge model measured in quantity (lbs/day) in the plant level effluents. Similar to the findings of the relative concentration model the estimated coefficients are higher when the past three years average DO is included, as opposed to the past two and one year mean/median values. For example, the coefficient in col. (4) and row (4) can be interpreted as, if past three years mean DO is reduced by one percent then relative quantity discharge would be reduced by 1.395 percent (for the sub-sample with at least 50 percent monthly observations). By contrast, if past two and one year(s) mean water quality is reduced by one percent (for the same sub-sample), then relative quantity discharges would be reduced by 1.259 and 1.025 percent, respectively (as seen in cols (7) and (10) of row (4)).

Unlike the relative concentration model(s), the marginal impact of past period average water quality actually deteriorates in magnitude (though by a small amount) as one moves across the different ‘balancing’ criteria applied to the original sample, and irrespective of which measure of lagged DO is considered. For instance, when the full sample of plants is considered, if past year mean (median) water quality is reduced by one percent, then its effect on relative quantity discharges is that it would be reduced by 1.214 (1.037) percent (col. (8) and row (4)). The above estimate is bigger than the coefficients obtained when only the plants with at least 25 and 50 percent monthly observations are included. For the above sub-samples, if past year mean (median) DO is reduced by one percent, then relative quantity is reduced by 1.079 (0.927) (col. (9) and row (4)) and 1.025 (0.985) (col. (10) and row (4)) percent, respectively.

Log of design flow is rarely significant; e.g., when the past two years of DO but only the 'most balanced' sample is considered, then relative quantity discharge would be actually be reduced by 0.039 percent (significant at the 10 percent; col. (7) and row (5)). This result is contradictory to the relative concentration model(s), where bigger plants are found to pollute more, though the coefficient is rarely significant in either model.

Unlike the relative concentration models, predicted inspections seem to be influencing quantity discharges, relative to limits, decision in all the different water quality lags and sub-samples considered. For the past three years mean (median) values of downstream DO, the threat of an additional inspection reduces relative quantity discharges by 93.38 (93.05) percent (col. (2) and row (6). The (huge) negative impact (as one would expect) means that regulatory pressure is effective, through the monitoring activities conducted by the designated authorities.

Seasonal pattern in relative quantity discharges is again (similar to the concentration models) evident with higher relative discharges during winter and spring (compared with summer).

Table 3.3: Log of Ratio of Monthly BOD5 Quantity Discharge to Effluent Limit Models Controlling For Serial Correlation (Estimation Technique: FGLS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	***Lagged DO measures***	Past three years mean (median)			Past two years mean (median)			Past one year mean (median)		
(2)	***Sampling Criterion***	Entire sample	At least 25 percent monthly observations	At least 50 percent monthly observations	Entire sample	At least 25 percent monthly observations	At least 50 percent monthly observations	Entire sample	At least 25 percent monthly observations	At least 50 percent monthly observations
(3)	# of plants	85	67	46	85	67	46	85	67	46
(4)	***Log of downstream DO***	1.558** (1.657**)	1.390** (1.538**)	1.395** (1.593**)	1.402** (1.498**)	1.272** (1.406**)	1.259** (1.629**)	1.214** (1.037**)	1.079** (0.927**)	1.025** (0.985**)
(5)	***Log of design flow***	0.024 (0.023)	0.018 (0.018)	-0.045+ (-0.049)*	0.026 (0.029)	0.018 (0.021)	-0.039+ (-0.039+)	0.027 (0.024)	0.016 (0.014)	-0.034 (-0.035)
(6)	***Predicted inspections***	-2.715** (-2.667**)	-3.090** (-3.033**)	-2.681** (-2.408**)	-2.522** (-2.360**)	-2.978** (-2.846**)	-2.684** (-2.221*)	-2.631** (-2.560**)	-3.218** (-3.105**)	-2.837** (-2.700**)
(7)	Wooldridge test for autocorrelation in panel data	150.026	169.617	164.839	131.220	132.017	180.196	130.816	128.909	92.704
(8)	*Prob > F*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*
(9)	Breusch-Pagan test for heteroskedasticity	160.83	206.32	180.25	165.51	205.17	198.52	210.42	217.37	222.91
(10)	*Prob >chi2*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*
(11)	Joint test for log of Sociodem.	56.20	74.43	133.52	51.19	68.31	143.80	56.47	73.04	175.81
(12)	*Prob >chi2*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*

Note: The test statistics for autocorrelation, heteroskedasticity, and joint test of log of demographics are reported for the mean values of past water quality

Conclusions

In this chapter, plant managers and operators are found incorporating past period downstream water quality in their (daily) monthly average abatement decisions (after controlling for other factors that might also influence plant level discharges). The underlying incentive is that they want to avoid future (expected) penalties of non-compliance and other enforcement actions such as sanctions and closure of operation. In particular, if plants do not pay attention to water quality prevailing in their receiving waters, a poor status of ambient quality might result in a more stringent effluent limit for the plant. Given that abatement technology is inflexible, the potential for stricter permits imply that the facility might run up against increased frequency of inspections, enforcement actions and pecuniary penalties until the time it can bring itself into compliance with its lower permit. Plants responding to ambient water quality, over and above their effluent limits, might also explain why they overcomply, in general.

In the relative discharge models, estimated for both BOD5 concentration and quantity measures, positive evidence is found on lagged ambient water quality at downstream locations exerting a significant influence on abatement decisions. Polluters seem to be incorporating longer-term measures of past water quality, in particular, two or three year lagged averages, as opposed to only the preceding year. This inference conforms with the regulators, themselves, making assessments on the ‘status’ of ambient water quality at specific locations over two to three years. The FGLS estimates (for the entire sample) show that if past three years mean (median) DO is reduced by one percent, then relative concentration discharges is reduced by 1.301 (1.257) percent, while relative quantity discharges is reduced by 1.558 (1.657) percent. On the other hand, adopting a ‘balancing’ approach, by estimating the model with at least 25 and 50 percent of the total possible monthly observations, improved the marginal impacts of the past water quality measures only for the relative concentration models.

Among the other determinants included, design effluent flow does not seem to exert a consistent impact on either concentration or quantity relative discharges. Predicted inspections, on the other hand, exert a consistently significant and negative impact for relative quantity discharges, while it is rarely significant (but negative, as expected) for the concentration measure.

Chapter 4:A Model of Water Quality and BOD discharges

Introduction

In this chapter the impact of total BOD discharges from the effluents of one or more plants, on one dimension of in-stream water quality, namely dissolved oxygen (DO), is examined. The total amount of pollutant discharges and physical-chemical aspects of a stream such as assimilative capacity[1], at a certain point in time, determine how contemporaneous abatement decisions and hence pollutant discharges affect ambient water quality. The ability of a waterbody to absorb pollution, in turn, depends on temporal variables such as pressure, stream flow, velocity, salinity, temperature (i.e. season), etc., not to mention other sources contributing to ambient pollution, arising due to differences in land use, for example.

The Streeter-Phelps (S-P) equation describes the underlying procedure of how pollutant discharge from a plant, at a certain point on a stream, determines ambient water quality at some distance downstream. The functional relation defined by this mathematical model predicts the ambient water quality (DO) resulting from the discharge of a pollutant (consuming DO) (along with the other factors mentioned in the previous paragraph). Hence, S-P is used extensively for predictive purposes; in particular, to simulate the effect of regulations i.e. effluent limits on downstream DO.

The model estimated in this chapter differs from the S-P. It is employed in an ex-post analysis examining the impact of actual BOD discharges on observed water quality.

---

[1] The assimilative capacity of a stream (also known as the natural attenuation process) "is defined as the breakdown of wastes by bacteria, fungi, and invertebrates; and the utilization of these waste products for the growth of algae and other forms of aquatic life" (Forrester, 2000).

The empirical framework of this chapter assesses the impact of pollutant discharges on downstream water quality utilizing actual ambient water quality (and hence ex-post) and self-reported pollutant discharge data.

One might perceive the analysis in this chapter as apparent since the inherent characteristics of "major"[2] facilities imply that these polluters most likely exert a significant impact on local downstream water quality. For example, Earnhart (2004c) did not include ambient water quality in his pollutant discharge model for municipal facilities based on the presumption that ambient water quality and discharges "are most certainly endogenously determined, especially given the "major polluter" status of the studied facilities" (p. 422; footnote 43). Nevertheless, it is important to study the impact of pollutant inputs on water quality in order to improve the effectiveness of policy since it is the channel through which policy affects ambient pollution, or at least understand the impact of 'major' point sources (if any).

Better information linking planned expenditures (costs of abatement) to ambient quality improvements (expected benefits) is necessary to plan more effective programs under the CWA. For example, to estimate benefits of pollution abatement one has to trace how policy changes alter pollutant discharges, and how this in turn affects ambient environmental quality. Hence, policy variables (such as effluent limits) are, at best, indirect determinants of downstream water quality, since they exert an influence only through the pollutant discharges of the plants that are regulated by these policies. In particular, whether pollutant discharges from major plants actually had an impact on

---

[2] Major municipal dischargers include all facilities with design flows greater than one million gallons per day, or facilities serving populations greater than 10,000, or facilities with EPA/State approved industrial pretreatment programs i.e. they receive industrial process wastewater. Major industrial facilities are determined based on specific ratings criteria developed by the EPA or the authorized State (USEPA, 1996).

actual downstream water quality has never been examined. However, there does exist some limited evidence on the simulated impact of BOD discharges by major plants where only large increases in upstream pollutant input generated ‘noticeable’ changes in downstream water quality.

Review of the Literature

Ambient water quality has not been studied extensively in the water pollution and its regulation literature primarily on account of unavailability of continuous time series data for a large enough geographical area (e.g. a sub-watershed in the least). McConnell and Schwarz (1992) and Schwarz and McConnell (1993) estimate cross-section permits and discharge models (respectively), which consider factors such as mean flow and velocity as proxies for ambient water quality. They mention that consistent water quality data within even 20-50 miles of the plant locations were not available during the early 1980s.

With the availability of data, most of the conventional[3] pollutants namely ambient BOD (e.g. in Sigman 2002 and 2004), fecal coliform, total suspended solids, phosphorus and nitrogen (e.g. in Sigman 2005) have been modeled as measures of water quality. However, none of the studies that model ambient water quality estimate the impact of pollutant inputs on in-stream water quality. Instead, exogenous factors such as stream flow and temperature that influence the impact of pollutants on water quality have been

---

[3] Conventional pollutants have been the focus of most regulatory efforts specially since toxics such as heavy metals are attributed almost solely to industrial activities. By contrast, manufacturing and sewage treatment plants as well as agricultural farms and urban runoffs discharge these pollutants.

used (e.g. in Sigman 2005). Ceteris paribus, higher stream flow means that the adverse impact of a given BOD discharge is lower due to reaeration and dilution. Ceteris paribus, higher temperature means that the impact of a certain amount of BOD input on ambient DO is higher since breakdown of organic matter is higher which requires DO.

Luken et al. (1992) utilize the S-P formulation with the purpose of simulating ambient water quality as a consequence of the total pollutant inputs in a free-flowing stream segment. They do not estimate a relationship between actual pollutant discharges and ambient water quality. Their sample included cross-section data of 60 mills on 66 reaches where pulp and paper mills were the dominant point source pollution within 10 miles upstream and downstream. Field data from EPA's STORET water quality database was compared with the predictions of the S-P model to validate the outcomes of the simulations.

Sigman (2002) is one of the first studies that estimate an empirical ambient water quality model using data on observed BOD in rivers and streams. In the United Nations' Global Environmental Monitoring System (GEMS) database annual mean concentrations of BOD at 291 stations, from around the world, are recorded from 1979 to 1990, while triennial means are available from 1991 to 1996. The model was:

$$\log b_{it} = g(S_i, POP_{it}, GDP_{it}, f_i, k_{it}, c_i, t) + \varepsilon_{it} \tag{4.1}$$

b_{it} is the mean pollution concentration at station i in year t, S_i is a vector of the measures of spillover[4], POP_{it} is upstream population, GDP_{it} is annual per capita GDP, f_i is river

---

[4] The term "spillover" refers to the physical effect of upstream countries' pollution on downstream pollution, and it is distinct from any elevated pollution attributable to free riding. When pollution crosses national (or state) borders polluting countries (states) do not experience the full benefits of pollution control. As a result, upstream countries (states) free ride by moving polluting activities near their downstream borders implying that pollution levels would be elevated at stations downstream of the border. BOD is identified as a good candidate since it travels reasonably far downstream allowing significant

flow, k_{it} is the deoxygenation rate, c_i is a country effect, and t is a time trend. Spillover effects were captured by categorizing the set of stations as upstream, downstream or on an international border, within or outside the EU. A weighted least squares model was estimated with the number of monthly measurements used to calculate the annual BOD measure as the weights.

For stations upstream of borders external to the EU pollution was 52 percent higher. For stations upstream of borders, within the EU, pollution was lower implying less free riding, but the coefficient of this indicator of station location was not significant. Thus the results provide some evidence that international spillovers cause degradation of water quality. Among the other determinants, upstream population had a positive and significant coefficient. The river flow rate had a negative coefficient, which was consistent with dilution, but dropped out of significance upon including country effects. The deoxygenation rate had positive but insignificant coefficient.

The focus of Sigman's 2004 paper was the influence of bilateral trade on pollution in rivers that cross international borders. The equation estimated is:

$$\log b_{it} = g(TI_{it}, Pop_i, CountryChar_{it}, RiverChar_{it}, t) + \varepsilon_{it} \tag{4.2}$$

b_{it} is the mean pollution concentration at station i in year t, TI_{it} (which is $\frac{Exports_{ud} + \mathrm{Im}\,ports_{ud}}{GDP_u}$, where the numerator captures the exports and imports from the upstream country to the downstream country divided by the GDP of the upstream country, for the upstream and downstream countries at station i in year t) is a matrix of

---

spillovers at many stations on international rivers. Sigman estimates that approximately 34% of pollution at an average upstream station will reach the border with the remainder getting attenuated naturally.

trade intensity variables. Dummy variables are included for the location of the station (upstream, downstream or on a border). Pop_i is upstream population for the station in 1994, $CountryChar_{it}$ include income, openness, and political rights, $RiverChar_{it}$ include flow and de-oxygenation rate, and t is a time trend.

Higher trade between adjacent countries meant significantly lower BOD pollution at downstream stations. Evidence of free riding was also found in the positive coefficients on the downstream station dummies. Combining these two effects, pollution was still slightly higher at downstream stations, but not significant. Sigman concluded that sufficient trade (coordination effect) overcomes free riding. The coefficients on river flow and de-oxygenation rates were negative and significant at 5% and 10% respectively.

Upon estimating the model including station-level fixed effects, rather than country dummies, all the trade intensity variables dropped out of significance. Sigman interprets this result as support for the hypothesis that the observed effect of trade on pollution is a coordination effect rather than a scale effect. Higher trade flows might have the short run effect of higher production of tradable goods resulting in more pollution (scale effect). However, higher trade eventually leads to greater collaboration efforts to reduce cross-country pollution (long run, coordination effect). Hence, short run fluctuations in trade due to changes in production related factors, e.g., are unlikely to influence the ability of countries to coordinate, and thus the purely time-series identification of these coefficients did not yield any results.

In 2005, Sigman goes on to estimate an ambient water quality model using USGS's National Stream Quality Accounting Network (NASQAN) data. The analysis examined the effect of free riding on water quality downstream of authorized states in the

US. NASQAN had data on water quality for a panel of about 500 monitoring stations across the US from 1973 to 1995 (with substantially fewer observations at the beginning and end of this period). The equation that was estimated:

$$WQ_{it} = G(S_{it}, S_{ht,} y_{it}, y_{ht}, g_{it} g_{ht}, L_{it}^{h}, f_{it}, m_{it}, A_{i}) \qquad (4.3)$$

Water quality index WQ_{it} [5] at a station with location i and in year t was a function of local and upstream variables (upstream location denoted by h). The factors considered were: upstream and downstream states' authorization status S_{ht} and S_{it}, level of economic activity y_{ht} and y_{it}, green preferences g_{ht} and g_{it}, land use (that is characterized by both upstream and downstream conditions at the watershed level) L_{it}^{h}, the river flow f_{ht} and temperature m_{it}, and lastly, station-specific fixed effects A_{i}. States that were vested with the responsibility of implementing and enforcing regulations are referred to as "authorized" as opposed to the ones where the EPA assumes the lead role in implementing policies. "Free riding"[6] was captured by a combination of authorization status and the station's location. For stations located downstream of borders, the measure was whether the upstream state was authorized. For stations located upstream of a border, the measure was an interaction of the upstream dummy with own-state authorization status. For stations on the border, the measure was a dummy variable that equals one if either state is authorized.

---

[5] This index was based on five common pollutants: dissolved oxygen, fecal coliform, total suspended solids, phosphorus and nitrogen.

[6] Free riding is synonymous with the ''shifting'' hypothesis which suggests that upstream state authorization would cause river water quality to fall more dramatically as a river flows downstream than it would with federal authority. Basically, upstream state has the incentive to shift its polluting activities to very near the border resulting in overall good water quality, except in the miles near the border.

Stations that were downstream from an "authorized" state witnessed significantly higher pollution. Sigman interpreted this outcome as free riding, which resulted in about 4% deterioration in the water quality index at downstream to border stations, on an average. As for border stations, the rivers were about 6% dirtier if at least one of the adjacent states was authorized. Next, the location variables were altered to reflect distance from a station to the nearest upstream or downstream border, to capture the natural attenuation phenomenon.[7] At downstream stations within 50 miles of the border[8], the upstream state's authorization had a significantly negative effect.

Earnhart (2004c) found indirect evidence of the significant impact of pollutant discharges from "major" municipal facilities on watershed level ambient water quality. A time-invariant ambient water quality index, published by the EPA, for each watershed (in the state of Kansas) was used as a measure of water quality. The significant positive coefficient on the water quality index was interpreted as greater relative discharges (ratio of discharges to permitted levels) leading to a decline in the ambient water quality.

Gray and Shadbegian (2004) also found indirect evidence of the potentially small (in magnitude) but "significant" impact of the pollutant discharges from point sources on water quality. They used EPA's National Water Pollution Control Assessment Model (NWPCAM)[9] to simulate water quality as a result of the BOD and TSS discharges of 231 paper and pulp mills in the US (for the period 1985 to 1997). Pollutant discharge data are combined with stream and river flow data to calculate the transport of pollutants

---

[7] "Far downstream of a border, the pollution endowment from upstream free riding dwindles with natural attenuation; far upstream of a border, the polluting state experiences almost all the damage" (p.93).

[8] Sigman mentioned that the 50 miles criterion was chosen based on the physical rates of attenuation for oxygen depletion.

[9] It includes pollutant discharge data for over 50,000 industrial and 13,000 municipal water polluters, nationwide.

downstream and the resulting water quality on a mile-by-mile basis for every stream. For most plants, only large increases in their pollution discharges generated measurable impacts on downstream water quality.

The discussion in this section reveals that empirical evidence on the impact (if any) of BOD pollution from ‘major’ polluters on actual downstream water quality does not exist. Previous studies that dealt with observed data on ambient water quality did not incorporate upstream pollutant inputs in the absence of previously available information linking specific polluters to downstream monitors.[10] Other studies that were not based on examining actual ambient water quality used the S-P mechanism to predict downstream water quality as a consequence of BOD discharges by point sources. Such simulations indicate that BOD discharges by ‘major’ polluters might have a significant but small (in magnitude) impact on actual downstream water quality.

The next section presents the S-P model that outlines the physical-chemical processes that determine downstream dissolved oxygen content in the water as a function of BOD inputs and a number of river/stream characteristics (such as flow) since they influence the assimilative capacity of the stream.

The Theoretical Water Quality Model

In this chapter, the ***Streeter-Phelps oxygen-sag curve*** is presented not to predict the ambient dissolved oxygen that would be observed in-stream, but to gain an understanding of the technical relationship between observed ambient dissolved oxygen and the BOD effluent discharges of the “major” point source polluters. It describes the

---

[10] The analysis in this current chapter models the upstream-downstream relation by using GIS (pre-mapped data available through the EPA’s BASINS software).

evolution of ambient dissolved oxygen levels as a consequence of pollution inputs of BOD, upstream ambient concentrations, flow, temperature and assimilative capacity of the stream, etc.

Oxygen is essential for the survival and propagation of aquatic organisms. If the amount of oxygen dissolved in water, falls below the minimum requirements for survival, aquatic organisms or their eggs and larvae may die. A severe example is a fish kill. Hence, surface waters protected for warm-water fish and aquatic life must meet the minimum dissolved oxygen standard of 5 mg/l.

Oxygen enters the water by photosynthesis of aquatic biota and by the transfer of oxygen across the air-water interface (reaeration). Different forms of pollution cause declines in DO. Matter containing carbon or nitrogen uses dissolved oxygen from the water as it decomposes, which can result in a dissolved oxygen decline. Nitrogenous demand for oxygen (NBOD) arises due to the presence of nitrifying bacteria, which oxidizes ammonia to nitrite first, then to nitrate. Non-point sources of pollution (agriculture primarily) are the predominant factors giving rise to significant NBOD. Carbonaceous organic matter present in the effluent discharges of point sources also create a demand for oxygen (CBOD), since bacteria oxidizes organic carbon into carbon dioxide and water.

Dissolved oxygen (DO) also varies greatly due to natural phenomena resulting in daily[11] and seasonal cycles. Seasonally, DO concentrations are greater in the colder winter months and lower in the warmer summer months. It is because gas solubility increases with decreasing temperature (colder water holds more oxygen), decreasing

---

[11] The natural diurnal (daily) cycle of DO concentration is well documented. Dissolved oxygen concentrations are generally lowest in the morning, climbing throughout the day due to photosynthesis and peaking near dusk, then steadily declining during the hours of darkness.

salinity (freshwater holds more oxygen than saltwater), and decreases with decreasing pressure (higher altitude waters have less oxygen because of the decrease in relative pressure). High temperatures encourage the microbial breakdown of organic matter a process that requires dissolved oxygen. In addition, stream flow (in freshwater) that is generally lower during late summer and early fall (in MD and VA) greatly affects flushing (dilution of pollutant inputs), re-aeration (mixing at the air-water interface), and the extent of saltwater intrusion, all of which affect dissolved oxygen values. The low-flow and high-temperature period is referred to as the critical condition since it has the potential to produce the worst effect on water quality.

The differential equation that outlines the process of evolution of deficit (D) in ambient dissolved oxygen is:

$$\frac{dD}{dt} = k_D L_0 e^{-k_D t} - k_R D \tag{4.4}$$

Equation (4.4) is essentially a balance between DO consumption due to BOD expression and stream reaeration. The first term captures the rate of deoxygenation i.e. consumption of DO, while the second expression captures the reaeration process. k_R and k_D are the reaeration time constant (depends on stream velocity and depth) and the de-oxygenation constant, respectively. t is time, and L_0 is the initial DO deficit in the stream, at the point of discharge of effluents from a point source. The oxidation of carbonaceous (and nitrogenous) substances present in the wastewater of the municipal and industrial plants creates an initial oxygen deficit at the point of outfall of the effluents. $L_{0,}$ also known as the "ultimate" BOD, is the BOD measured in the river/wastewater mixture, which is given by:

$$L_0 = \frac{Q_r L_r + Q_w L_w}{Q_r + Q_w} \tag{4.5}$$

Where :

L_0 = BOD (in mg/L) at the point of wastewater discharge

Q_r = Flow of the river, upstream of the waterwater discharge

L_r = BOD (in mg/L) measured in the river, upstream of the wastewater discharge

Q_w = Flow of the wastewater discharge

L_w = BOD (in mg/L) measured in the wastewater discharge

It is also known as the mass balance formula for mixing.

The solution to the differential equation (4.4) gives the S-P dissolved oxygen sag curve (Streeter and Phelps, 1925), which was originally developed for use on the Ohio River in 1914. It is a steady-state model relating dissolved oxygen concentration in a free flowing stream to BOD.[12]

$$D = \frac{k_D L_0}{k_R - k_D}(e^{-k_D t} - e^{-k_R t}) + D_0 e^{-k_R t} \tag{4.6}$$

D = Dissolved Oxygen deficit = DO_s-DO i.e. the difference between the equilibrium concentration DO_S and the actual concentration DO is the oxygen deficit. DO_s is the maximum amount of dissolved oxygen that can be held in the water. It depends on the water temperature, salinity, and pressure. Given u, which is the average water velocity, the ambient water quality at a distance d downstream (to wastewater discharges) can be determined.

---

[12] The S-P dissolved oxygen deficit equation with modifications to account for the oxygen demand resulting from nitrification of ammonia (nitrogenous oxygen demand) and the organic demand found in the water body sediment is shown in the equation below.

$$D = \frac{K_1 L_0}{K_2 - K_1}\left(e^{-K_1 t} - e^{-K_2 t}\right) + \frac{K_3 N_0}{K_2 - K_3}\left(e^{-K_3 t} - e^{-K_2 t}\right) + \frac{SOD}{K2H}(1 - e^{-K_2 t}) + D_0 e^{-K_2 t}$$

Where: D = dissolved oxygen deficit at time t, mg/l; L0 = initial CBOD, mg/l; N0 = initial NBOD, mg/l (NBOD = NH3-N x 4.57); D0 = initial dissolved oxygen deficit, mg/l; K1 = CBOD decay rate, 1/day; K2 = reaeration rate, 1/day; K3 = nitrification rate, 1/day; SOD=sediment oxygen demand, g O2/ft2/day; H=average stream depth, ft; and t = time, days.

To start with, DO is being depleted faster than it can be replenished (in equation (4.4)). As long as this occurs, the DO of the stream will continue to drop. Since the BOD is decreasing as time goes on, at some point, the rate of deoxygenation decreases to just the rate of reaearation. At this point (called the critical point) the DO reaches a minimum. Downstream of the critical point, reaeration occurs faster than deoxygenation, so the DO increases. Using calculus we get:

$$t_c = \frac{1}{k_R - k_D} \ln\left(\frac{k_R}{k_D}\left[1 - \frac{D_0\left(k_R - k_D\right)}{k_D L_0}\right]\right) \tag{4.7}$$

The above is the critical sag distance (time) equation for dissolved oxygen, which depends on the values of the reaeration and deoxygenation constants, the DO deficit and BOD at the outfall of the wastewater. "DO might reach a minimum outside the limits of the municipal or industrial boundary of the discharge and indeed… in another state or even another country" (p. 308; Thomann and Mueller 1987). Hence, the maximum impact of oxidation of organic material is generally not at the location of the point source discharge, but at some distance downstream where the maximum DO deficit occurs. For instance, Ntelekos (2005) simulated critical distances of 14.7, 21.3 and 23.7 miles using three different reaeration rate constants when a BOD discharge of 13 mg/L was introduced in the Delaware River, at Trenton, under specific conditions such as temperature ($15^\circ C$) and streamflow etc. The point of maximum DO deficit is also called the DO sag. For example, the TMDL for DO is based on a daily average of not less than 5.0 mg/l at the DO sag during critical conditions in the water quality limited segment (WQLS).

Substituting for initial dissolved oxygen deficit D_0 (= DO_s - DO_0), dissolved oxygen deficit D (DO_s - DO), L_0 and DO_0, using the respective (mass balance) formulas:

$$\begin{aligned} DO &= DO_S(1-e^{-k_R t}) - \left(\frac{k_D}{k_R - k_D}\right)\left(\frac{Q_w}{Q_r + Q_w}\right)L_w(e^{-k_D t} - e^{-k_R t}) \\ &- \left(\frac{k_D}{k_R - k_D}\right)\left(\frac{Q_r}{Q_r + Q_w}\right)L_r(e^{-k_D t} - e^{-k_R t}) + \frac{Q_r}{Q_r + Q_w}DO_r e^{-k_R t} + \frac{Q_w}{Q_r + Q_w}DO_w e^{-k_R t} \end{aligned} \tag{4.8}$$

Hence, a negative relationship is predicted between concentration of BOD in the wastewater discharge, L_w, and downstream ambient water quality DO. Equation (4.8) predicts the DO level in a single reach of stream as a result of the addition of a point source "loading" of waste at the upstream end of the reach. This equation can be used to determine the DO concentration in several successive reaches by applying the deficit at the downstream end of one reach as the initial deficit of the succeeding reach. Thus, equation (4.8) can be applied iteratively to determine the DO profile of an entire stream system (see Liebman and Lynn, 1966; Thomann and Mueller 1987; and USEPA 1991 for DO simulation in consecutive (multiple) stream segments).

Regulators undertaking water quality evaluations every permit cycle have used the S-P model[13] of a free flowing stream, and hence to determine whether (and what level of) water quality based effluent limits for BOD were needed. For example, in Maryland

---

[13] Using the basic concept of S-P many increasingly complex mathematical models have been developed to simulate DO dynamics in streams (Vellidis and others, 2006). "Most were developed to simulate parameters associated with [the NPDES] permits" (p. 1007), while some specifically simulated DO, others were broader in-stream water quality models, and yet others were watershed-scale transport models incorporating the contribution of non-point sources to water quality degradation. QUAL2E (Enhanced Stream Water Quality Model) is one of the two most popular (one-dimensional, steady-state etc) models for developing DO TMDLs (USGS, 2005), while HSPF (Hydrological Simulation Program- Fortran) is a dynamic model. Vellidis and others (2006) list numerous examples of TMDLs implemented across the nation that were based on these models. Soil and Water Assessment Tool (SWAT) is another example of a river basin model that quantifies the impact of land management practices in large watersheds, at the same time as simulating in-stream processes such as DO.

(MD) and Virginia (VA) water quality based limits for BOD were derived from non-TMDL waste load allocations. Given the effluent limits of a plant determined by technology-based standards, a dissolved oxygen sag analysis is conducted where the lowest concentration of ambient DO realized under critical low flow condition[14] is simulated. If the DO level generated does not meet the ambient standard required to meet the designated use of the stream, water quality based limits that are more stringent than technology based ones are invoked.

As the water pollution regulation moves towards the TMDL regime (in order to incorporate the loadings of non-point sources) the S-P equation continues to be used as the underlying model.[15] In particular, it is now being extensively used as an evaluation tool for the implementation and adoption of TMDLs for stream segments with low dissolved oxygen problems[16]. The S-P uses background and point source loadings of BOD, and simulates oxygen addition through atmospheric re-aeration and photosynthesis. It determines how much more load allocations from all non-point sources and waste-load allocations from all point sources[17] could be permitted so that the ambient water quality standard is met in future time periods. Alternatively, it is used to

---

[14] For MD, critical low flow condition is representative of a drought condition, and is defined as the minimum 7 consecutive day average stream discharge having a recurrence interval of 10 years (7Q10). It is so called because the ability of a water body to assimilate pollutants without exhibiting adverse impacts is at a minimum.

[15] TMDLs utilize a steady-state model similar to a modified Streeter-Phelps DO sag equation. The in-stream DO target for a TMDL is a daily average of not less than 5.0 mg/l for surface water.

[16] Either on account of increases in point sources pollution or excessive algal growth due to high dissolved nitrogen levels. Excessive inputs of nutrients (nitrogen and phosphorus) can lead to over-enrichment and eutrophication of the water-body. The nutrients act as fertilizer leading to excessive growth of aquatic plants, which eventually die and decompose leading to bacterial consumption of dissolved oxygen (DO) implying that ambient concentrations might fall below what is necessary to support the designated use.

[17] Load Capacity is calculated using the formula:
Permit limit average daily load =(Design flow of facility in cfs) X (effluent pollutant concentration in mg/L) X (the constant 5.395 to convert to pounds/day.)

determine how much pollutant inputs from all relevant sources have to be reduced in order to maintain the water quality standard for a stream segment.

Estimating the Streeter-Phelps Equation As A Typical Empirical Model

The only two "coefficients" in the S-P model, in equation (4.8), are the reaeration (k_R) and deoxygenation (k_D) constants. However, these two are not parameters in the conventional sense since they are functions of stream characteristics, environment, and waste discharges from point (and non-point) sources observed at each monitoring location on a stream/river. Consequently, the parameters in the S-P formulation are location specific, and hence cannot be generalized across distinct monitoring locations: to answer the question of how important is the influence of BOD5 pollutants of a (typical) plant's discharge on ambient quality, on an average. Keeping this important qualification in mind, if one were to still try to estimate the S-P equation as an "empirical" model, i.e., a common set of reaeration and deoxygenation constants across different stations, some of the challenges posed are discussed below.

To empirically estimate the S-P (equation (4.8)) applied to observed water quality (DO) and pollutant discharges (L_w) data, one would require additional information on saturated dissolved oxygen concentration (DO_S) of all the relevant stream segments defined by each pair of monitoring stations, flow of the wastewater effluent (Q_w), stream flow (Q_r), stream velocity (u) given that distance d is known, concentration of BOD upstream to the point of discharge of the plant's effluents (L_r), dissolved oxygen concentration in the river upstream of the point of outfall of the plant (DO_r), and lastly, the dissolved oxygen concentration in the wastewater(DO_w).

First, there is no data on equilibrium (saturated) concentration of dissolved oxygen DO_S for each segment of river. There is, however, data on monthly concentration of BOD5 and flow of the wastewater discharge which are L_w, and Q_w, respectively in equation (4.8). Third, the only source of data on stream flow at the monitoring stations Q_r is the USGS "Water Data for the Nation".[18] However, data on stream velocity u is not recorded, even, in this USGS database. Data on the concentration of BOD in the river prior to the wastewater discharge (L_r) are much less likely to be monitored.[19] The concentration of dissolved oxygen in the wastewater DO_w is also much less likely to be reported by the NPDES plants.[20] Additionally, water quality data immediately upstream to the point of discharge of a plant (DO_r) is not available for almost all the plants in the sample. Instead, DO data was available at monitoring stations located at a certain distance further upstream to the point of outfall. Lastly, a unique pair of upstream and downstream stations could not be identified for each plant in the sample. Incorporating multiple polluters in between the same pair of upstream and downstream

---

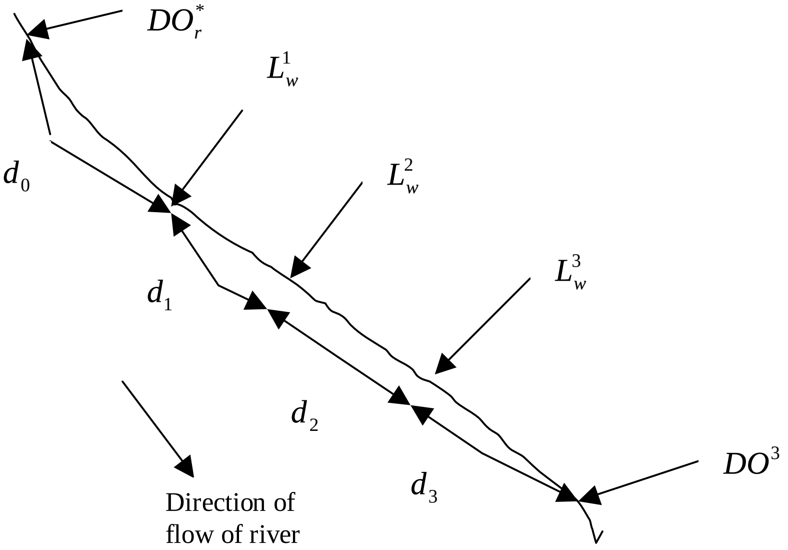
[18] The monitoring stations in the current sample (from EPA's STORET etc.) would have to be matched with the USGS gage station ids since it is the only source of stream flow data.

[19] For instance, ambient BOD concentration (L_r) have not been recorded by the USEPA STORET, or the USGS water quality monitoring systems for the stations in the current sample during the period 1990 to Feb. 2004. State level sources such as the Virginia Department of Environmental Quality (VADEQ) has stopped monitoring for BOD5 from 2001 onwards, since over the last twenty years a steady decline was witnessed and most of the BOD values (> 90%) were less than the detection limit.

[20] Plants are more likely to face permits for allowable levels of BOD5 in their effluent discharges rather than monthly average DO concentration limits.

monitors (Figure below), either due to lack of water quality data or monitoring stations, the equation that could be estimated is:[21]

Figure 4.1: A Stream Segment with Multiple Polluters



$$
\begin{aligned}
DO_{jt} &= DO_{ut}e^{-k_R(d_0+d_1+d_2+d_3)} - \left(\frac{k_D}{k_R-k_D}\right)e^{-k_R(d_2+d_3)}(e^{-k_Dd_1}-e^{-k_Rd_1})BOD_{1jt} \\
&-\left(\frac{k_D}{k_R-k_D}\right)e^{-k_Rd_3}(e^{-k_Dd_2}-e^{-k_Rd_2})BOD_{2jt} - \left(\frac{k_D}{k_R-k_D}\right)(e^{-k_Dd_3}-e^{-k_Rd_3})BOD_{3jt} \qquad (4.9) \\
&+\delta_j + \gamma_T + \varepsilon_{jt}
\end{aligned}
$$

DO_{jt} = concentration of DO in the river at the monitoring location j downstream to all the 3 point source polluters

DO_{ut} = concentration of DO in the river, $d_0 + d_1 + d_2 + d_3$ miles upstream to the downstream monitoring location j

BOD_{1jt} = concentration of BOD5 in the effluent discharge reported by plant 1, $d_1 + d_2 + d_3$ miles upstream to the downstream monitoring location j

---

[21] Details on how equation (4.9) is derived are included in the Appendix (Deriving the Streeter-Phelps equation for multiple polluters).

BOD_{2jt} = concentration of BOD5 in the effluent discharge reported by plant 2,
$d_2 + d_3$ miles upstream to the downstream monitoring location j
BOD_{3jt} = concentration of BOD5 in the effluent discharge reported by plant 3,
d_3 miles upstream to the downstream monitoring location j

δ_j = dummy variable for monitoring location j
γ_T = dummy variable for year T, with $T = 1,2....13$
ε_{jt} = error term for monitor j in month t

Equation (4.9) shows that the BOD input by each plant enters additively in the model that determines downstream DO. However, it does not give a simple coefficient on each one of the BOD5 discharges L_w. Instead, it is a complex function of the re-aeration and de-oxygenation “constants” and the different distance components. Therefore, a non-linear relationship could be estimated using least squares method across monitoring stations, keeping in mind that the constants of S-P are really location specific.

But, first, a simplified linear relationship (across monitoring locations), between ambient DO at a certain distance downstream and total (sum of the) BOD pollution from all the relevant point sources, is estimated in the next section. This basic framework suffices to meet the objective of an empirical evaluation of the impact of BOD5 discharges from point sources’ effluents on ambient water quality. Forrester (2000) also noted that studies such as El-Shaarawi, Esterby and Kuntz (1983), Bodo (1992) and Esterby (1996) report using linear, robust and multivariate regression analysis extensively for determining water quality trends. Other empirical papers such as Hirsch and Slack

(1984) have noted that among the common water quality variables, only temperature, pH and DO can be considered close to normal[22].

A Change in Downstream Net of Upstream Water Quality Model

In equation (4.10), below, an alternative model of the impact of pollutant inputs from point source dischargers on ambient water quality is estimated. The dependent variable is dissolved oxygen measured at location *j* downstream to the point of outfall of plant i in month t (DO_{jt}) net of water quality measured at an upstream location u in month t (DO_{ut}). Upstream water quality is certainly a determinant of water quality observed at a certain distance downstream on the same stream/river. In the current sample, correlation between contemporaneous upstream and downstream water quality is about 75%. Upstream water quality is endogenous in the sense that it is determined based on the same factors that explain downstream quality (except of course, the pollutant discharges). In particular, seasonal variations (flow and temperature) are expected to affect ambient DO at both upstream and downstream locations similarly (even if they were not ‘close’ to each other).

$$\begin{aligned} DO_{jt} - DO_{ut} = \alpha + \beta_1 \sum_{i=1}^{3} w_1 BOD_{ijt} + \beta_2 \sum_{i=1}^{3} w_2 BOD_{ijt} + \beta_3 \sum_{i=1}^{3} w_3 BOD_{ijt} \\ + \delta_j + \gamma_T + \varepsilon_{jt} \end{aligned} \tag{4.10}$$

---

[22] Most other water quality parameters such as nutrients, BOD, and biological indicators such as biomass and bacterial counts have been found to be non-normally (in particular, log normally) distributed (USEPA, 1991; Gilliom and Helsel, 1986).

DO_{jt} = concentration of DO in the river at monitoring location j and month t,
downstream to the point source polluters
DO_{ut} = concentration of DO in the river at location u,
upstream to the monitoring location i and for month t
$w_1 = 1$ for $0 < d_{ij} < d1$; $= 0$ otherwise
$w_2 = 1$ for $d1 < d_{ij} < d2$; $= 0$ otherwise
$w_3 = 1$ for $d2 < d_{ij}$; $= 0$ otherwise
BOD_{ijt} = concentration or quantity of BOD5 reported in plant i's wastewater in month t,
d_{ij} miles upstream to monitoring location j, $i = 1,2,3$
δ_j = dummy variable for monitoring location j
γ_T = dummy variable for year T, with $T = 1,2....13$
ε_{jt} = error term for monitor j in month t

The primary explanatory variable of interest is the weighted sum of concentration (or quantity) of BOD5 measured in each plant i's wastewater in month t (BOD_{ijt}) with its point of outfall between locations j and u. In the current sample there are at most 3 plants in between monitoring locations j and u (i ={1,2,3}). The weights used are indicators of the distance of each plant from its downstream monitoring location j (d_{ij}). In essence, the sample of plants discharging in the same stream segment is grouped according to their location relative to the downstream monitor. The underlying principle behind this weighting system (discussed in detail below) is that if the plants are located "far enough", represented by $d1$, upstream from station j then its pollutants have traveled the longest distance on the stream/river, and hence are likely to have undergone the most attenuation. At the same time if the downstream monitor is located "too near", represented by $d2$, the plants, then the adverse impact of BOD (on DO) might not have manifested itself yet.

Among the controls, time invariant[23] location specific effects that have not been captured by water quality at the upstream location[24] are captured by station level dummy variables. Specifically, it controls for non-point source pollution for the segment of the stream in between the upstream and downstream station. Regulators while conducting the TMDL analysis extensively use field data on water quality in order to approximate the impact of pollution from non-point sources. Yearly dummy variables have been included to capture any possible time trends in downstream water quality net of upstream condition across the three states from 1990 to 2003.[25] Seasonal indicators are not included since variations in temperature and rainfall are expected to affect downstream and upstream water quality in an identical manner, and hence season is not anticipated to influence the change in downstream from upstream DO.[26] Lastly, annual dummy variables are included to control for any possible time trend over the 14 years of data.[27]

Serial correlation of error terms within the same monitoring location and from one month to the next is also anticipated, since there might be un-captured factors that are

---

[23] For example, differences in land use at the downstream location since this is unlikely to change 'substantially' in a short period of time.

[24] In particular, location specific "physical" effects such as velocity and depth (determining natural attenuation rates), pressure (and topography) and salinity (determining saturated oxygen levels) might be captured reasonably well by upstream water quality if the two monitors are closer to each other.

[25] Even though the BOD discharge data was collected from 1990 to Feb. 2004, the observations from 2004 are not included in the regression sample since a yearly trend for the year 2004 cannot be identified from observations on the first two months.

[26] Moreover, monthly BOD discharged by the plants also exhibits seasonal variability. Seasonal average sum of BOD5 concentration was 12 mg/L during winter, 10.2 in spring, 8.5 in summer, and 10.1 in fall. In particular, plants face a higher (less stringent) permit level during the cooler winter and fall seasons and lower (more stringent) permissible amount during the warmer summer and spring seasons. One anticipates this seasonal behavior of the polluters even if it has not been formally incorporated in their permitting system. Plant managers have the foreknowledge of the well-documented seasonal pattern of DO and hence might decide to pollute less (in terms of concentration of BOD) during the summer months in contrast to the winter months.

[27] It is worth mentioning that detecting a trend in the water quality data is not what this Chapter aims for; nevertheless, the yearly dummies are included to differentiate between observations coming from different years. Studies such as Esterby (1996) have found that data records of at least 10 to 20 years are needed to model short-term trends. Similarly, Zipper and others (1998) undertake trend analysis over a period of 30 years early to mid 1970s to 1997, while VADEQ (2004) does it over twenty years from 1985-2004.

specific to a monitoring location and a certain point in time. For instance, assimilative capacity of a stream segment is calculated at a point in time, which depends on temporal variables such as pressure, stream flow, velocity, salinity, temperature etc., not to mention other sources contributing to ambient pollution. In particular, “minor” point sources that also discharge BOD or even other “major” facilities, which might have changed ‘status’ during the time period of this study.[28] Hence, water quality observed at a certain monitoring location in month ‘t’ would be correlated with water quality in month ‘t+1’, at the same location, and irrespective of whether the two months being compared are within the same season.

From S-P we know that the impact of BOD discharged on DO depends on the distance traveled downstream from the point of outfall of the wastewater. Specifically, its impact is increasing till the critical distance is reached, at which point DO is minimum. Beyond the critical distance, the impact of BOD declines till DO reaches levels that are observed before the wastewater discharge. Hence, if the plants discharging in the same stream segment are located too near their downstream monitor, then its impact on DO might not be observed yet. Alternately, if the set of plants are located too far upstream relative to the downstream monitor, then the impact of their BOD discharges might have undergone substantial attenuation.

The downstream stations (with available DO data) that are considered in the current sample might not be in the ‘range’ where the impact of BOD on DO can be measured. S-P tells that natural attenuation of BOD i.e. how far BOD discharged in a stream can travel before its effect essentially gets “wiped out”, i.e. DO gets back to the

---

[28] As Forrester (2000) reports various studies that undertake trend analysis of water quality data have dealt with the serial correlation problem, e.g., Hirsch, Slack and Smith (1982) and Hirsch and Slack (1984).

level prevailing before the wastewater discharge, varies case-by-case. In other words, it can be 50 miles as Sigman pointed out in her papers on water quality (no references were cited), or close to 40 miles as seen in the graphs presented by Ntelekos (2005), or 7-10 miles as seen in the case studies in pages 318 and 366 of Thomann and Mueller (1987). For example, these distances would depend on the initial BOD and DO deficit, and stream conditions determining its assimilative capacity, etc. In the absence of this (fore) knowledge on "the impact zone" of BOD (on DO), the sample of relevant plants are distinguished by their location relative to the downstream monitor as: 'too near', 'in the impact zone' and 'too far'.

Data

In the current sample, there are 79 pairs of downstream and upstream stations with data on monthly dissolved oxygen over the 14-year period. The choice (and hence location) of downstream as well as upstream monitoring stations, for one or two or even three facilities, was primarily driven by availability of water quality data. GIS maps by EPA, named BASINS, were used to place the point source on a stream/river reach and then identify the relevant upstream and downstream monitoring stations. Hence, the distance between the point source outfall and upstream/downstream monitoring location was determined on a case-by-case basis depending on how far upstream/downstream there were stations with ambient water quality data records. However, two pairs of stations and plants did not have GIS data. A third plant had to be dropped since an appropriate upstream station did not exist.[29] By states, the distribution of stations is 42

---

[29] The plant discharged its effluents near the point of confluence of two tributaries.

for VA, 20 for MD, and 14 for PA. The distribution of the 97 major facilities was 51 plants in VA, 26 plants in MD, and 20 plants in PA.

There are 5 plants (with concentration discharge data) and stations combinations, which had at least one plant on a tributary (total 3) whereas the others were on the main river. For these plants, the upstream and downstream monitoring stations and hence water quality data is considered from the main river i.e. before and after the tributary joins it.[30] The 'choice' of monitors is based on ensuring that water quality data from the most appropriate pair of monitoring stations are considered: for purposes of assessing the impact of pollutant inputs on downstream water quality. Admittedly, this method of approximating the location of the effluent discharge point with the point of confluence of the tributary with the main river is inaccurate.[31]

A unique pair of upstream and downstream stations could be identified for 59 of the 97 "major" manufacturing and sewage treatment facilities sampled. For the remaining 38, 26 of them had one other major facility discharging into the same stream segment and hence they had the same pair of upstream and downstream stations. The last 12 plants had two other plants polluting into the same stream segment i.e. there were three plants discharging "in between" the same pair of upstream and downstream monitoring stations.[32] Average upstream and downstream distance for the current sample

---

[30] Water quality data at either an upstream and/or a downstream station on these tributaries is not available.

[31] If the 3 plants on the tributary are dropped, then the station level fixed effects would be capturing the effect of other major polluters in the relevant stream segment that are not incorporated in the explanatory variable i.e. total pollutant discharge. The latter is bound to be an issue even for other plants and station combinations since the entire universe of major polluters are not included in the current sample of 100 plants.

[32] Average distance between the upstream station and the first/only plant in the present sample was 8.1675 miles with 75% of the data points below 13 miles (Table A 60). Average distance between first and second plant (only plant and downstream station) was 8.645192 miles with 75% of the measurements less than 11 miles (Table A 61). Average distance between second and third plants was 10.86678 miles (Table A 62) and between third and downstream station was 7.865 miles (Table A 63).

of plants, keeping in mind all of the above approximations utilized, are 10 and 11 miles respectively with 75% of these observations under 14.5 and 16.5 miles (Table A 64 and Table A 65).

Average DO across stations and years at the downstream (upstream) locations was 12.1, 8.9, 7.5 and 10.1 (12.1, 8.9, 7.6 and 10.4) for winter, spring, summer and fall, respectively (Table A 72). Hence, DO levels are higher during winter and fall seasons when low temperatures prevail. Average DO levels during spring are also lower (in comparison with fall specially) despite similar milder conditions and higher stream flow than fall. Additionally, there is more than one record of DO for a month measured either on different days or at different times in the same day. For this present estimation, monthly data points are calculated by averaging across different times in the same day and/or different days in the same month.[33]

The min and max of DO data[34] at stations that are downstream to at least one NPDES plant in the current sample is 1.41 and 42.1 mg/L (Table A 69).[35] By contrast, the range for upstream stations is 1.24 to 18.79 mg/L (Table A 70). Upon examining the station level data, it was found that DO values are recorded at remarkably high levels for

---

[33] Time of measurement of DO for each observation i.e. data record is another potentially important aspect because of the well-established diurnal nature of ambient DO. One line of further investigation could be to take differences in the time of measurement into account when calculating the monthly data by station. Ensuring that the time at which the data was measured is similar (if not identical) across all the data points would be a possible refinement of the empirical model that is estimated here.

[34] For details on data sources please see page 4 of Chapter 1.

[35] The data on monthly water quality observations at both the upstream and downstream locations is an unbalanced panel. Table A 66 shows that the data on monthly dissolved oxygen at downstream and upstream stations by year and by state is most balanced for MD. For VA, the number of DO records was declining in 2001 and later, whereas for PA the decline started in 1999. Table A 67 present the summary statistics of DO data at downstream and upstream locations by state, and Table A 68 report the downstream and upstream DO across monitoring stations by year. No interesting pattern emerges upfront from these ‘layers’ of analyzing the DO data.

two monitoring stations on the Delaware River in PA.[36] After dropping the six observations the new range for downstream stations becomes 1.41 to 19 mg/L (Table A 71). The mean and median of DO recorded at the upstream and downstream stations now become 9.5-9.7 mg/L, much above the ambient standard of 4-6 mg/L (depending on temperature) required for maintaining "aquatic life support". Only the bottom 10% of the DO data from these downstream and upstream stations had some observations lower than this range (Table A 70 and Table A 71).

Table 4.1: Summary Statistics of Change in Downstream DO and Sum of Concentration/Load Data

Variable	Mean	Min	Max
Change in downstream DO from upstream DO, mg/L	-0.08	-10	9.4
Sum BOD5 concentration, mg/L	10.07	0.1	199.6
Sum BOD5 quantity, lbs/day	547.45	.0616	40522

Note: Detailed summary statistics of the dependent and primary explanatory variables are presented in Table A 74, Table A 75 and Table A 76.

Data on five-day biochemical oxygen demand (BOD5) measured in the effluent discharges of each plant facing a BOD5 limit, are reported in the monthly Discharge Monitoring Report (DMR), to the relevant regulatory authority (EPA or State). Out of the 97 facilities in the sample, 83 of them reported concentration of BOD (5-day) discharged in their effluents on a monthly basis. Table A 73 reports that the range was from 0.1 to 178.5 mg/L, and the average was around 8.5 (5.7 median).[37] Monthly data might not be available due to reporting/monitoring requirements, or non-reporting by the

---

[36] In 2002, DO levels recorded at these two stations changed dramatically from 6.45 (and 7.31) mg/L in May to 30.85, 32.45 (and 34.1, 42.1) mg/L in June (and July). Similarly, DO values changed from 37.75 (and 39.75) mg/L in September to 8.8 (and 9.1) mg/L in November. These observations are most likely data entry errors because moving from spring towards summer cannot dramatically increase DO levels, neither does moving from summer to fall drastically reduce the DO levels.

[37] Table A 75 reports the summary statistics of the sum of BOD5 concentration, for one or more plants if they were identified as discharging in between the same pair of upstream and downstream monitors. There were 67 first/only plants reporting BOD5 concentration. These included 54 'single' plants between two monitoring stations, and 13 'first' plants in a multiple polluter segment. Out of the latter 13, there were 10 that were the second (or last) in a two-polluter segment. And, 3 others had their discharge points on a 3-polluter segment.

facility, or even data entry error. In order to maintain consistency with the sample of plants used in the permits and abatement models, where observations without a numeric effluent limit drop out, the results presented in this chapter focus on those monthly discharges where limits were populated by numerical values.[38] Even though, for purposes of estimating water quality models all (daily and monthly average) discharges by a plant, reported or otherwise and irrespective of the status of the effluent limits, might be relevant.

As with any data over a certain time series and across a set of plants addressing the issue of unbalanced data is a crucial concern. Typically, one can address this problem by 'balancing' the sample e.g. by restricting attention to the sample with at least a certain number of monthly observations across the years.[39] In this current chapter, such methods are not applied because irrespective of the total number of time periods (months in this case) for which data on change in water quality as well as BOD discharges are available, downstream DO is influenced by upstream pollution in an 'instantaneous' manner (as seen in the Streeter-Phelps equation). In other words, total BOD discharges from all sources of pollution 'aught' to affect downstream dissolved oxygen (if at all) for any or every instant in time and hence any or every month (on an average). Consequently, the issue of non-randomness of non-reporting of monthly BOD discharges by the plants is not a central problem for the analytical framework of this chapter.[40] On the other hand,

---

[38] In most of these situations the regulatory agency might have monitored the facility without limiting their discharges. The only inference that could be drawn from when a polluting facility does not face a numerical limit is that most likely the regulatory agency deems these discharges "insignificant" as "major" polluters.

[39] Well over 80% of the monthly observations of water quality and pollutant discharge had at least 10 out of the total 12 months data (as seen in Table A 77,Table A 78,Table A 79 and Table A 80).

[40] In particular, let us consider the situation that plants do not report their discharges in the monthly DMR when its discharges are higher relative to other months (irrespective of its level of compliance). This implies that ceteris paribus the adverse impact of BOD pollution on ambient dissolved oxygen would be

multiple monthly observations from the same location and across years identify any time pattern exhibited by the DO data over the 14 years. Moreover, missing observations on DO are due to differences in the monitoring frequency at different stations, and hence is largely determined by the resources available to water quality monitoring networks to facilitate sampling at all stations on a regular basis (at least once a month).

Estimating the Change in Downstream Water Quality Model

In equation (4.10) (repeated below), downstream net of upstream DO is regressed on sum of BOD5 concentration (or quantity) of all relevant plants i.e. discharging in the same stream segment. The pollutant inputs are categorized into three groups. The first and third groups include the plants that are located too near to and too far from the downstream monitor, respectively. The second groups are those that are most likely in the range where the impact of BOD on DO can be evident. The equation also includes fixed effects for monitoring station and year.

$$DO_{jt} - DO_{ut} = \alpha + \beta_1 \sum_{i=1}^{3} w_1 BOD_{ijt} + \beta_2 \sum_{i=1}^{3} w_2 BOD_{ijt} + \beta_3 \sum_{i=1}^{3} w_3 BOD_{ijt} + \delta_j + \gamma_T + \varepsilon_{jt}$$

The first step in estimating is identifying a reasonable "impact zone" for the BOD discharges in the sample (captured by the dummy variable w_2). In turn, this involves

---

higher during the months with non-reported discharge data. On the other hand, suppose that positive evidence is found on BOD discharges from major point sources determining downstream ambient quality for those time periods with available data (on both variables) and hence presumably lower BOD discharges from these plants. In which case, inference can be drawn on the estimated impact, which would be bigger if data was available for those months during which plants did not report their discharge. Moreover, conversations with VA permitting officials revealed that missing BOD discharge data due to changes made in the processing of the PCS data as opposed to non-reported monthly DMRs might be a more accurate explanation for majority of the unavailable BOD discharge data (Kyle I. Winters; email communication).

using a best approximate lower (w_1) and upper (w_3) limits for the downstream distance of each relevant plant that can be applied across distinct plant(s)-monitor combinations. In particular, for distances longer than the upper limit the impact of BOD is not felt i.e. attenuated completely, and for distances closer than the lower limit the impact of BOD is not felt either, this time, due to deoxygenation rates that are starting out from zero (see S-P discussion in the appendix). Different combinations of such upper and lower limits are tried given that the exact distance of how far BOD can travel, i.e., how far downstream the impact of BOD discharge can still be felt is, or how soon, in terms of distance downstream to the point of discharge, its effect can be observed, is not known. Best guesstimate for the "impact zone" is chosen when sum of BOD does not have any impact for either below (above) the lower (upper) limit, but exerts a (statistically) significant impact within the consequential impact range. The expected direction of impact of sum of BOD discharges, within the impact range, on net downstream DO is negative, since by definition BOD is a pollutant that consumes dissolved oxygen.

The coefficient estimates for β_1, β_2 and β_3 are found in Table 4.2 below shows that the lower and upper limits of 2 and 26 miles and the consequential range of >2 and <26 miles serves as a good set of candidate of the impact zone that can be generalized across monitors (and plants). The first set of upper and lower limit tried is less than 6 miles and bigger than 19 miles (column (1) of Table 4.2). It is not a good candidate for approximately identifying a common impact range for the current sample. In particular, sum of BOD exerts a significantly negative impact on downstream net of upstream DO for the sub-sample of plants that are located less than 6 miles from their downstream monitor. The second set is the lower and upper limit of 4 and 23 miles. This too fails to meet the criterion of a “good” impact zone. Specifically, for plants that are located less than 4 miles from its monitors, sum of BOD has a significant, negative influence on downstream net of upstream DO. The final set tried is the impact zone of more than 2 and less than 26 miles. It is a good candidate because for plants that are located at distances lower than 2 miles and higher than 26 miles, sum of BOD does not exert a statistically significant impact. Consequently, for plants located within the higher than 2 and lower than 26 miles range, if sum of BOD is increased by one mg/L, then downstream relative to upstream DO is reduced by 0.005 mg/L.

The table also shows that the data exhibits autocorrelation problem even after controlling for monitor level variations by including dummy variables. Hence, the FGLS estimation method is appropriate where the error terms moving from one month to the next, but within the same monitor are correlated (as well as heteroskedastic). Station level fixed effects are highly significant, while there is no apparent trend in net downstream water quality over the 14 years sampled. The joint test for station specific effects is accepted with a significant chi2-statistic and very small p-value.

Table 4.3 below reports the results obtained from undertaking the impact-zone analysis using BOD loads as opposed to concentration discharges. Plants located at distances longer than 4 miles and less than 23 miles seem to be the relevant range for BOD loads to exert a statistically significant impact on downstream net of upstream DO. The coefficient can be interpreted as: if sum of BOD5 loads is increased by one lbs/day, then downstream net of upstream DO would be reduced by 0.000(2) mg/L (column (3)).

Table 4.2: FGLS Results of Net Downstream Water Quality with Sum of BOD5 Concentration

		(1)	(2)	(3)
Variables	Coefficient in equation (4.10)	downstream - upstream, DO	downstream - upstream, DO	downstream - upstream, DO
sum BOD5 concentration, downstream distance of all plants <6 miles	β_1	-0.017 (0.000)**		
sum BOD5 concentration, downstream distance of all plants >6 and <19 miles	β_2	0.005 (0.237)		
sum BOD5 concentration, downstream distance of all plants <19 miles	β_3	0.006 (0.487)		
sum BOD5 concentration, downstream distance of all plants <4 miles	β_1		-0.009 (0.023)*	
sum BOD5 concentration, downstream distance of all plants >4 and <23 miles	β_2		-0.012 (0.011)*	
sum BOD5 concentration, downstream distance of all plants >23 miles	β_3		0.009 (0.298)	
sum BOD5 concentration, downstream distance of all plants <2 miles	β_1			-0.007 (0.143)
sum BOD5 concentration, downstream distance of all plants >2 and <26 miles	β_2			-0.005 (0.045)*
sum BOD5 concentration, downstream distance of all plants >26 miles	β_3			0.016 (0.276)
Observations		6322	6322	6322
Test for autocorrelation (Prob >F)		9.189 (0.0036)	8.854 (0.0042)	8.960 (0.0040)
Test for heteroskedasticity (Prob > chi2)		9.00 (0.0027)	8.81 (0.0030)	8.32 (0.0039)
Joint test for station FE (Prob > chi2)		1942.36 (0.0000)	1942.19 (0.0000)	1931.15 (0.0000)
Joint test for annual FE (Prob > chi2)		14.97 (0.3094)	16.06 (0.2461)	15.12 (0.3002)

Note: p values in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%; joint test statistics are reported for FGLS (p-values in parentheses); and significance of estimates of the “impact zone” are based on one tailed test.

Table 4.3: FGLS Results of Net Downstream Water Quality with Sum of BOD5 Loads

		(1)	(2)	(3)
Variables	Coefficient in equation (4.10)	downstream - upstream, DO	downstream - upstream, DO	downstream - upstream, DO
sum BOD5 quantity, downstream distance of all plants <6 miles	β_1	0.000 (0.677)		
sum BOD5 quantity, downstream distance of all plants >6 and <19 miles	β_2	-0.000 (0.475)		
sum BOD5 quantity, downstream distance of all plants >19 miles	β_3	-0.000 (0.822)		
sum BOD5 quantity, downstream distance of all plants <4 miles	β_1		0.000 (0.342)	
sum BOD5 quantity, downstream distance of all plants >4 and <23 miles	β_2		-0.000 (0.025)*	
sum BOD5 quantity, downstream distance of all plants >23 miles	β_3		0.000 (0.758)	
sum BOD5 quantity, downstream distance of all plants <2 miles	β_1			0.000 (0.336)
sum BOD5 quantity, downstream distance of all plants >2 and <26 miles	β_2			-0.000 (0.453)
sum BOD5 quantity, downstream distance of all plants >26 miles	β_3			0.001 (0.571)
Observations		7179	7179	7179
Number of group(downstreamstation)		72	72	72

Note: p values in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%; and significance of estimates of the "impact zone" are based on one tailed test.

Comparing the Sum of BOD Concentration Results to S-P Estimates

For comparison, a "partial" S-P model, equation (4.9), is estimated given the constraints on availability of data on different components required to estimate the "full" S-P equation (4.8). Downstream water quality at location j is regressed on upstream water quality at location u, BOD5 concentration of plant '1' i.e. one furthest upstream of downstream monitor, concentration discharges of plant '2' the one located in an intermediate range from the downstream monitor, and concentration discharges from plant '3' which is located nearest to the downstream station. k_R and k_D are the coefficients that will be estimated; given that d_0 is the distance of plant '1' from the upstream station, d_1 is the distance between plant '1' and plant '2', d_2 is the distance of plant '3' downstream from plant '2', and d_3 is the distance between plant '3' and the downstream station j. Station level fixed effects are included to control for the monitoring location specific features, other than BOD discharges by point sources, that also determine downstream water quality (along with annual dummy variables).

$$\begin{aligned} DO_{jt} &= DO_{ut} e^{-k_R(d_0+d_1+d_2+d_3)} - \left(\frac{k_D}{k_R-k_D}\right) e^{-k_R(d_2+d_3)} (e^{-k_D d_1} - e^{-k_R d_1}) BOD_{1jt} \\ &\quad - \left(\frac{k_D}{k_R-k_D}\right) e^{-k_R d_3} (e^{-k_D d_2} - e^{-k_R d_2}) BOD_{2jt} - \left(\frac{k_D}{k_R-k_D}\right) (e^{-k_D d_3} - e^{-k_R d_3}) BOD_{3jt} \\ &\quad + \delta_j + \gamma_T + \varepsilon_{jt} \end{aligned}$$

In this modified S-P model, it is the concentration measure of BOD5 that is relevant. Consider equation (4.5)[41], which calculates the "ultimate" BOD that can be

---

[41] $L_0 = \dfrac{L_r}{1+{Q_w}/{Q_r}} + \dfrac{L_w}{1+{Q_r}/{Q_w}}$

observed in the wastewater- river mixture after a point source's effluent outfall and controlling for upstream ambient pollution. Ceteris paribus, a plant with a higher load i.e. higher effluent flow has a bigger weight on the effluent concentration and a smaller weight on the ambient upstream concentration. Since the concentration of BOD measured in the wastewater is anticipated to be greater than the ambient BOD concentration before the point of wastewater discharge, on account of natural attenuation, "ultimate" BOD is actually higher.[42] [43] In the limit (i.e. as design effluent flow approaches an infinitesimally large number, "ultimate" BOD is determined only by the effluent concentration L_w (in equation (4.5)).[44] Consequently, regulators (EPA and State) focus their permitting, monitoring and compliance activities towards "major" polluters where one of the criteria used is design flow greater than 1 mgd. Essentially, when the planned volume and hence flow of wastewater is 'large enough' the chances that it will not be 'insignificant' when compared to actual stream flow are greater.

Evidence on the concentration rather than the load being utilized as the relevant measure for purposes of 'determining' downstream water quality can be found from the practical standpoint of how policies are implemented. For instance, regulators assign TMDLs for stream segments that are "impaired" in terms of limits on the concentration of BOD5 under critical 7Q10 drought-like stream flow conditions. Water quality models

---

[42] Unless of course there is massive non-point source pollution problem just before the point of effluent discharge, for instance, that would imply that upstream concentration of BOD is higher than effluent concentration.

[43] For example, Summers, Kazyak, and Weisberg (1991) utilize the QUAL2E-UNCAS water quality model to simulate the impact of a reduction in the effluent discharge rate (flow) by 55 percent of a large kraft paper mill discharging its effluents to the Pigeon River in NC. They find that simulated BOD in the river is reduced by 20 ppm in the vicinity of the discharge.

[44] On the other hand, as stream flow approaches an infinitesimally big number "ultimate" BOD is determined solely by the upstream concentration of BOD L_r (in equation (4.5)). Hence, the concerns for non-point sources of pollution contributing to the majority of ambient pollution during high stream flow seasons.

such as QUAL2E were utilized to derive the "ultimate" BOD that would meet the ambient standards (accounting for a margin of safety). Given that background pollution is zero under 7Q10 conditions, BOD in the river/wastewater mixture was solely captured by effluent concentration. The corresponding maximum allowable load is determined by calculating the product of the effluent limit of BOD5 and the design effluent flow of a plant. See, for example, numerous TMDLs implemented by various states and the EPA across the US: MODNR (1999), MODNR (2005), MDE (1999), MDE (2000), MDE (2002) and SCDHEC (1998).

In Table 4.4 below the non-linear estimation is conducted on the identical sub-sample of plants i.e. the best candidate for capturing the impact range for BOD (on DO). For the sub-sample with downstream distance of plants greater than 2 and less than 26 miles, the estimates of impact of BOD of each plant are obtained by evaluating the coefficients on the respective BOD terms in equation (4.9) for a certain set of distances. This approach is taken primarily because the distances for the current sample of plants are fixed factors i.e. exogenous. Consequently, plugging in the estimates of k_R and k_D at the mean values of the different distance components (e.g.) for the relevant sample of plants gives an estimate of the impact of plant '1', '2' and '3' discharging in between the same upstream and downstream stations, on an average.

Given the estimates of k_R and k_D in Table 4.2[45], consider three plants that are located in the same stream segment. The distance between plant 1 and plant 2, d_1 in equation (4.9), is 7.68 miles; the distance between plant 2 and plant 3, d_2 in equation (4.9), is 11.05 miles; the distance between plant 3 and the downstream monitor, d_3 in

---

[45] Appendix Table A 81 presents the NL results of estimating the "partial" S-P on the full sample and plants >2 and <26 miles for robustness checks on the estimates of kR and kD.

equation (4.9), is 8.27 miles. Hence, if plant 1's BOD is increased by one mg/L then downstream DO would be reduced by 0.008 mg/L (distance between plant 1 and downstream monitor is 27 miles); if plant 2's BOD is increased by one mg/L then downstream DO would be reduced by 0.011 mg/L (distance between plant 2 and downstream monitor is 19.32 miles); if plant 3's BOD is increased by one mg/L then downstream DO would be reduced by 0.008 mg/L (distance between plant 3 and downstream monitor is 8.27 miles).

Table 4.4: Comparison of FGLS and NL Estimates

Variables (Estimation Method)	downstream - upstream, DO (FGLS)	downstream DO (NL)	downstream DO (NL)
Sampling Criterion for downstream distance	Distance of plants >2 and <26 miles	Distance of plants >2 and <26 miles	Full sample
Sum of BOD5 concentration	–0.005* (0.045)		
Concentration of the plant furthest upstream from downstream monitor		–0.0077** (0.0055)	–0.0063** (0.0005)
Concentration of the plant second-most upstream of downstream monitor		–0.0111** (0.0055)	–0.0080** (0.0005)
Concentration of the plant closest to downstream monitor		–0.0083** (0.0060)	–0.0058** (0.0005)
# of downstream stations	43	43	67

Note: p values of one-tailed test in parentheses; + significant at 10%; * significant at 5%; ** significant at 1%.

Proof of the DO-sag is found from the estimates as seen in the Table 4.4 below. On an average, BOD discharge of the plant located furthest upstream has traveled the longest; consequently, its adverse impact is smaller than the 'intermediate' plant. Secondly, the plant closest to the downstream monitor has a smaller impact than the one 'in between' because the impact of BOD is maximized at a certain distance downstream and not immediately near the point of effluent outfall. Overall, the NL coefficients of the impact of BOD from plants 1,2 and 3 are over one and half times bigger than the FGLS coefficient of –0.005 mg/L.

Now, given the estimates of k_R and k_D for the full sample of monitors and plants (column (1) of Table 4.4), the coefficients on the respective BOD discharges are evaluated at the average distances. The distance between plant 1 and plant 2, d1 in equation (4.7), is 8.65 miles; the distance between plant 2 and plant 3, d2 in equation (4.7), is 10.87 miles; the distance between plant 3 and the downstream monitor, d3 in equation (4.7), is 7.87 miles. Hence, if plant 1's BOD is increased by one mg/L then downstream DO would be reduced by 0.006 mg/L (distance between plant 1 and downstream monitor is 27.39 miles); if plant 2's BOD is increased by one mg/L then downstream DO would be reduced by 0.008 mg/L (distance between plant 2 and downstream monitor is 18.74 miles); if plant 3's BOD is increased by one mg/L then downstream DO would be reduced by 0.006 mg/L (distance between plant 3 and downstream monitor is 7.87 miles). Since the coefficients for the full sample are smaller than the NL estimates for the sub-sample of plants that are >2 and <26 miles inference is drawn on the validity of the impact zone identified (albeit arbitrarily identified).

However, Table 4.3 shows that the corresponding impact zone for BOD loads is >4 and <23 miles.

Conclusions

The chapter investigated if the pollutant inputs of ‘major’ point sources had a significant impact on downstream water quality net of that prevailing at an upstream location (and by how much). Studying this causality is important because it traces the channel through which changes in policy has an impact on ambient water quality (the goal of regulation itself), and hence on benefits on pollution abatement. Specifically, abatement and hence pollutant discharge behavior of the plants in response to the limits they face on pollutants, discharged in their effluents, is the mechanism through which policy affects water quality.

There has been no previous empirical study on ambient water quality that examines the impact of pollutant discharges (BOD5 in this case) by ‘major’ point source polluters. Hence, the positive evidence found in this chapter means that regulation has most likely been effective, in terms of influencing ambient water quality, through variations in effluent limits of polluters and its effect on abatement behavior. Moreover, it also has implications for potential channels of improvements in efficacy of regulation since it links costs of pollution abatement with expected benefits.

Positive empirical evidence is found on the effect of pollutant discharges, from the plants, determining observed ambient water quality at downstream locations, controlling for upstream DO. A simple linear model is first estimated given that only a "partial" S-P equation can be estimated since data on the different components of equation (4.8) is not available. However, knowledge of the DO-sag is applied to categorize the sample of plants and monitors as too far or too near their respective downstream monitors. Admittedly, the 'impact zone' criterion applied is arbitrary since the upper and lower limits for distances where the impact of a given BOD discharge is not 'observed' on a particular stream segment, is case specific i.e. it cannot be generalized across distinct locations and BOD inputs. The coefficient obtained from the best candidate for capturing the impact of BOD is the range of greater than 2 and less than 26 miles. If sum of BOD5 concentration is increased by one mg/L, then downstream net of upstream DO is reduced by 0.005 mg/L (significant at 5%).

The NL estimates of the "modified" S-P equation for the identical sub-sample of plants and stations is bigger, however, tiny too. In particular, on an average, an increase in the BOD discharge of the plant located furthest upstream relative to the downstream monitor, affects downstream DO by approximately –0.008 mg/L. For the plant located 'in between' and the one closest to the downstream monitor, the coefficient is -0.011 and -0.008 mg/L respectively. Hence, policy is found to have an impact on its target (ambient water quality), indirectly, through the significant impact of abatement behavior of plants on downstream DO.

Appendix

Table A 1: Avg. concentration permit, by season

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%            5              5
 5%           10              5
10%           10              5       Obs                 481
25%         22.5              5       Sum of Wgt.         481

50%           30                      Mean           27.77957
                        Largest       Std. Dev.      13.32743
75%           30            100
90%           30            100       Variance       177.6205
95%         40.2            111       Skewness        2.66219
99%           88            111       Kurtosis        15.9011
```

Table A 2: Avg. quantity permit, by season

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%           22           9.68
 5%        39.82           9.68
10%       116.16       17.56111       Obs                 583
25%        312.4       18.04423       Sum of Wgt.         583

50%          622                      Mean           1868.636
                        Largest       Std. Dev.      4372.907
75%       1788.6          42400
90%         3800          42400       Variance       1.91e+07
95%         5621          42400       Skewness       6.725831
99%        22979          42400       Kurtosis       56.58848
```

Table A 3: Summary statistics of different DO measures

Variable	Obs	Mean	Std. Dev.	Min	Max
past permit cycle mean downstream DO, by season	450	9.635721	1.811466	4.722857	13.616
past permit cycle median downstream DO, by season	451	9.568149	2.009915	3.8	13.3
mean DO of last three years of preceding cycle, by season	440	9.594718	1.819835	4.746154	13.83846
median DO of last three years of preceding cycle, by season	440	9.516409	1.995797	3.41	13.7
mean DO of last two years of preceding cycle, by season	439	9.573254	1.857321	4.654	13.66
median DO of last two years of preceding cycle, by season	439	9.497642	2.0358	2.875	13.6
mean DO of last one year of preceding cycle, by season	436	9.548951	1.908215	4.098	14.8
median DO of last one year of preceding cycle, by season	436	9.457053	2.086402	2.875	14.8
two years before past cycle ends mean DO, by season	358	9.66558	1.806465	4.761111	13.96667
two years before past cycle ends median DO, by season	358	9.615363	2.027056	2.65	13.7
one year before past cycle ends mean DO, by season	421	9.677049	1.865312	4.77	14.66667
one year before past cycle ends median DO, by season	421	9.610036	2.074908	2.65	14.4

Table A 4: Design flow, mgd

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%          .03            .03
 5%          .34          .1081
10%           .6            .12       Obs                  97
25%          1.5           .325       Sum of Wgt.          97

50%         2.89                      Mean           54.89512
```

```
                        Largest       Std. Dev.       246.761
75%           12            368
90%        40.57            621       Variance       60890.97
95%          210           1234       Skewness       6.414689
99%         2000           2000       Kurtosis       46.44419
```

Table A 5: Summary statistics of SIC codes of plants

Variable	Obs	Mean	Std. Dev.	Min	Max
Electric Services	100	0.06	0.238683	0	1
Food And Kindred Products	100	0.04	0.196946	0	1
Textile Mill Products	100	0.04	0.196946	0	1
Paper And Allied Products	100	0.07	0.256432	0	1
Chemicals And Allied Products	100	0.1	0.301511	0	1
Petroleum Refining And Related Industries	100	0.06	0.238683	0	1
Rubber And Miscellaneous Plastics Products	100	0.01	0.1	0	1
Leather And Leather Products	100	0.02	0.140705	0	1
Fabricated Metal Products, Except Machinery And Transportation Equipment	100	0.01	0.1	0	1
Transportation Equipment	100	0.01	0.1	0	1
National Security And International Affairs	100	0.03	0.171447	0	1
Justice, Public Order, And Safety	100	0.01	0.1	0	1
Sewerage Systems	100	0.54	0.500908	0	1

Table A 6: Zip-code level sociodemographic characteristics

Variable	Obs	Mean	Std. Dev.	Min	Max
percent non-white	97	16.80697	17.23821	0	79.09908
median household income, in thousands of dollars	97	29.54049	9.608291	13.131	60.59
percent car-pooling	97	16.59107	4.627341	6.746733	30.12658
percent employed in the manufacturing sector	97	23.01013	12.56915	3.773585	53.75837
total population in the zipcode, in thousands	97	14.65946	14.5768	0.492	68.444
percent of residents living in urban areas	97	29.32448	41.96988	0	100

Table A 7: Level and log transformed concentration permits model (FGLS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	***Variables***	Permit cycle average ***concentration*** limit, by season (mg/L)					
(2)	***Past cycle seasonal DO measures***	Entire permit cycle mean (median)	Last three years of past cycle mean (median)	Last two years of past cycle mean (median)	Last one year of past cycle mean (median)	Excludes last two years of past cycle mean (median)	Excludes last one year of past cycle mean (median)
(3)	*DO (mg/L)*	1.328** (1.176**)	1.235** (1.137**)	1.176** (1.075**)	1.129** (1.033**)	1.381** (1.246**)	1.326** (1.174**)
(4)	Design flow (mgd)	-0.100** (-0.102**)	-0.095** (-0.096**)	-0.095** (-0.096**)	-0.092** (-0.094**)	-0.109** (-0.109**)	-0.100** (-0.103**)
(5)	***Variables***	***Log*** of permit cycle average ***concentration*** limit, by season					

(6)	*Log of DO*	0.617** (0.556**)	0.582** (0.539**)	0.563** (0.521**)	0.577** (0.529**)	0.625** (0.553**)	0.590** (0.534**)
(7)	Log of design flow	-0.078* (-0.082*)	-0.076* (-0.079*)	-0.076* (-0.078*)	-0.076* (-0.078*)	-0.102* (-0.105*)	-0.075* (-0.077*)
(8)	Chi2 stat. sociodem.	34.98	36.44	36.78	37.02	25.37	31.04
(9)	*Prob.>chi2*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0003*	*0.0000*
(10)	Chi2 stat. log socio.	32.78	33.60	33.78	34.31	17.73	26.96
(11)	*Prob.>chi2*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0033*	*0.0001*

Note: + significant at 10%; * significant at 5%; ** significant at 1%, and significance of one tailed test for coefficients on level and log BOD. Chi2 statistics and p-value for joint test of significance of socio-demographics are shown only for the mean values of past cycle water quality measures.

Table A 8: Level and log transformed quantity permits model (FGLS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	***Variables***	Permit cycle average ***quantity*** limit, by season (lbs/day)					
(2)	***Past cycle seasonal DO measures***	Entire permit cycle mean (median)	Last three years of past cycle mean (median)	Last two years of past cycle mean (median)	Last one year of past cycle mean (median)	Excludes last two years of past cycle mean (median)	Excludes last one year of past cycle mean (median)
(3)	*DO (mg/L)*	46.356** (38.593**)	48.769** (41.238**)	44.560** (38.358*)	48.193* (43.324*)	40.849** (32.642**)	44.746** (36.644**)
(4)	Design flow (mgd)	2.846 (2.786)	2.991 (2.885)	2.401 (2.310)	2.518 (2.407)	2.518 (2.515)	2.781 (2.662)
(5)	***Variables***	***Log*** of permit cycle average ***quantity*** limit, by season					
(6)	*Log of DO*	0.322** (0.268**)	0.344** (0.290**)	0.315** (0.255**)	0.357** (0.314**)	0.276** (0.207*)	0.293** (0.233**)
(7)	Log of design flow	0.707** (0.706**)	0.716** (0.715**)	0.719** (0.717**)	0.720** (0.719**)	0.711** (0.710**)	0.700** (0.700**)
(8)	Chi2 stat. sociodem.	17.61	18.51	18.08	18.56	14.51	14.75
(9)	*Prob.>chi2*	*0.0073*	*0.0051*	*0.0060*	*0.0050*	*0.0244*	*0.0223*
(10)	Chi2 stat. log socio.	21.77	21.92	19.67	18.68	24.14	18.63
(11)	*Prob.>chi2*	*0.0006*	*0.0005*	*0.0014*	*0.0022*	*0.0002*	*0.0023*

Note: + significant at 10%; * significant at 5%; ** significant at 1%, and significance of one tailed test for coefficients on level and log BOD. Chi2 statistics and p-value for joint test of significance of socio-demographics are shown only for the mean values of past cycle water quality measures.

FGLS estimations of permits model using different measures of past cycle water quality

Table A 9: Permits model with entire permit cycle mean DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	ln avg. concentration permit, by season	ln avg. quantity permit, by season
past permit cycle mean downstream DO, by season	1.328 (0.000)**	46.356 (0.003)**		
		(0.006)**		
avg. lag. size, by season	-0.100 (0.000)**	2.846 (0.600)		
ln past permit cycle mean downstream DO, by season			0.617 (0.000)**	0.322 (0.000)**
ln avg. lag. size, by season			-0.078 (0.030)*	0.707 (0.000)**
Constant	11.524	582.558	0.764	6.450
	(0.021)*	(0.300)	(0.328)	(0.000)**
Observations	175	345	171	341
Number of group(npdes)	38	72	37	71

Table A 10: Permits model with entire permit cycle median DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	ln avg. concentration permit, by season	ln avg. quantity permit, by season
past permit cycle median downstream DO, by season	1.176 (0.000)**	38.593 (0.005)**		
avg. lag. size, by	-0.102 (0.000)**	2.786 (0.607)		

season				
ln past permit cycle median downstream DO, by season			0.556 (0.000)**	0.268 (0.000)**
ln avg. lag. size, by season			-0.082 (0.023)*	0.706 (0.000)**
Constant	12.991	659.176	0.861	6.562
	(0.007)**	(0.234)	(0.265)	(0.000)**
Observations	175	345	171	341
Number of group(npdes)	38	72	37	71

Table A 11: Permits model with last three years of past cycle mean DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	ln avg. concentration permit, by season	ln avg. quantity permit, by season
mean DO of last three years of preceding cycle, by season	1.235 (0.000)**	48.769 (0.008)**		
avg. lag. size, by season	-0.095 (0.000)**	2.991 (0.579)		
ln mean DO of last three years of preceding cycle, by season			0.582 (0.000)**	0.344 (0.000)**
ln avg. lag. size, by season			-0.076 (0.034)*	0.716 (0.000)**
Constant	12.526	561.877	0.846	6.400
	(0.012)*	(0.318)	(0.282)	(0.000)**
Observations	175	343	171	339
Number of group(npdes)	38	72	37	71

Table A 12: Permits model with last three years of past cycle median DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	ln avg. concentration permit, by season	ln avg. quantity permit, by season
median DO of last three years of preceding cycle, by season	1.137 (0.000)**	41.238 (0.009)**		
avg. lag. size, by season	-0.096 (0.000)**	2.885 (0.593)		
ln median DO of last three years of preceding cycle, by season			0.539 (0.000)**	0.290 (0.000)**
ln avg. lag. size, by season			-0.079 (0.029)*	0.715 (0.000)**
Constant	13.597	644.394	0.935	6.530
	(0.006)**	(0.243)	(0.231)	(0.000)**
Observations	175	343	171	339
Number of group(npdes)	38	72	37	71

Table A 13: Permits model with last two years of past cycle mean DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	ln avg. concentration permit, by season	ln avg. quantity permit, by season
mean DO of last two years of preceding cycle, by season	1.176 (0.000)**	44.560 (0.009)**		
avg. lag. size, by season	-0.095 (0.000)**	2.401 (0.662)		
ln mean DO of last two years of			0.563 (0.000)**	0.315 (0.000)**

preceding cycle, by season				
ln avg. lag. size, by season			-0.076 (0.035)*	0.719 (0.000)**
Constant	13.026	533.247	0.868	5.968
	(0.011)*	(0.348)	(0.270)	(0.000)**
Observations	175	342	171	338
Number of group(npdes)	38	71	37	70

Table A 14: Permits model with last two years of past cycle median DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	ln avg. concentration permit, by season	ln avg. quantity permit, by season
median DO of last two years of preceding cycle, by season	1.075 (0.001)**	38.358 (0.012)*		
avg. lag. size, by season	-0.096 (0.000)**	2.310 (0.673)		
ln median DO of last two years of preceding cycle, by season			0.521 (0.000)**	0.255 (0.001)**
ln avg. lag. size, by season			-0.078 (0.031)*	0.717 (0.000)**
Constant	14.098	601.201	0.945	6.109
	(0.005)**	(0.283)	(0.228)	(0.000)**
Observations	175	342	171	338
Number of group(npdes)	38	71	37	70

Table A 15: Permits model with last one year of past cycle mean DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by	avg. quantity permit, by	ln avg. concentration permit, by	ln avg. quantity permit, by

	season	season	season	season
mean DO of last one year of preceding cycle, by season	1.129 (0.000)**	48.193 (0.012)*		
avg. lag. size, by season	-0.092 (0.000)**	2.518 (0.645)		
ln mean DO of last one year of preceding cycle, by season			0.577 (0.000)**	0.357 (0.000)**
ln avg. lag. size, by season			-0.076 (0.033)*	0.720 (0.000)**
Constant	13.671	496.194	0.881	5.913
	(0.006)**	(0.388)	(0.247)	(0.000)**
Observations	175	340	171	336
Number of group(npdes)	38	71	37	70

Table A 16: Permits model with last one year of past cycle median DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	ln avg. concentration permit, by season	ln avg. quantity permit, by season
median DO of last one year of preceding cycle, by season	1.033 (0.001)**	43.324 (0.011)*		
avg. lag. size, by season	-0.094 (0.000)**	2.407 (0.660)		
ln median DO of last one year of preceding cycle, by season			0.529 (0.000)**	0.314 (0.000)**
ln avg. lag. size, by season			-0.078 (0.029)*	0.719 (0.000)**

Constant	14.803	555.598	1.014	6.029
	(0.002)**	(0.325)	(0.173)	(0.000)**
Observations	175	340	171	336
Number of group(npdes)	38	71	37	70

Table A 17: Permits model excluding last two years of past cycle mean DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	ln avg. concentration permit, by season	ln avg. quantity permit, by season
two years before past cycle ends mean DO, by season	1.381 (0.000)**	40.849 (0.005)**		
avg. lag. size, by season	-0.109 (0.000)**	2.518 (0.721)		
ln two years before past cycle ends mean DO, by season			0.625 (0.000)**	0.276 (0.009)**
ln avg. lag. size, by season			-0.102 (0.014)*	0.711 (0.000)**
Constant	13.175	818.960	1.315	6.615
	(0.020)*	(0.254)	(0.146)	(0.000)**
Observations	135	263	131	259
Number of group(npdes)	37	71	36	70

Table A 18: Permits model excluding last two years of past cycle ends median DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	ln avg. concentration permit, by season	ln avg. quantity permit, by season
two years before past cycle ends median DO, by season	1.246 (0.000)**	32.642 (0.007)**		
avg. lag. size, by season	-0.109 (0.000)**	2.515 (0.721)		

ln two years before past cycle ends median DO, by season			0.553 (0.000)**	0.207 (0.012)*
ln avg. lag. size, by season			-0.105 (0.012)*	0.710 (0.000)**
Constant	14.790	903.918	1.488	6.771
	(0.007)**	(0.204)	(0.096)+	(0.000)**
Observations	135	263	131	259
Number of group(npdes)	37	71	36	70

Table A 19: Permits model excluding last one year of past cycle ends mean DO

	(1)	(2)	(3)	(4)
	avg. concentration permit, by season	avg. quantity permit, by season	ln avg. concentration permit, by season	ln avg. quantity permit, by season
one year before past cycle ends mean DO, by season	1.326 (0.000)**	44.746 (0.006)**		
avg. lag. size, by season	-0.100 (0.000)**	2.781 (0.624)		
ln one year before past cycle ends mean DO, by season			0.590 (0.000)**	0.293 (0.001)**
ln avg. lag. size, by season			-0.075 (0.041)*	0.700 (0.000)**
Constant	12.400	584.543	1.068	6.522
	(0.021)*	(0.350)	(0.198)	(0.000)**
Observations	158	319	154	315
Number of group(npdes)	38	72	37	71

Table A 20: Permits model excluding last one year of past cycle median DO

	(1)	(2)	(3)	(4)
	avg. concentration	avg. quantity	ln avg. concentration	ln avg. quantity

	permit, by season	permit, by season	permit, by season	permit, by season
one year before past cycle ends median DO, by season	1.174 (0.000)**	36.644 (0.003)**		
avg. lag. size, by season	-0.103 (0.000)**	2.662 (0.641)		
ln one year before past cycle ends median DO, by season			0.534 (0.000)**	0.233 (0.001)**
ln avg. lag. size, by season			-0.077 (0.038)*	0.700 (0.000)**
Constant	13.902	661.035	1.148	6.641
	(0.008)**	(0.292)	(0.162)	(0.000)**
Observations	158	319	154	315
Number of group(npdes)	38	72	37	71

Table A 21: Change in concentration permits, by season

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%          -58            -70
 5%           -5            -70
10%            0            -58       Obs                 308
25%            0            -58       Sum of Wgt.         308

50%            0                      Mean          -.9149485
                        Largest       Std. Dev.      7.951825
75%            0             10
90%            0             14       Variance       63.23153
95%     1.666666             20       Skewness      -6.631193
99%           10             20       Kurtosis       55.22439
```

Table A 22: Change in quantity permits, by season

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%      -1744.6        -2182.4
 5%         -286        -2182.4
10%    -82.27271        -1744.6       Obs                 390
25%            0        -1744.6       Sum of Wgt.         390

50%            0                      Mean           83.38129
                        Largest       Std. Dev.      672.2914
75%           40           3714
90%        290.4           3714       Variance       451975.7
95%        600.6           5010       Skewness       3.557245
99%         3714           5010       Kurtosis       26.40324
```

Table A 23: BOD concentration outlier of "MD0020524"

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%            1              1
 5%            2              1
10%          2.4              1       Obs                 135
25%          3.2              1       Sum of Wgt.         135

50%            5                      Mean           7.015778
                        Largest       Std. Dev.      12.08183
75%            7             23
90%         11.3             26       Variance       145.9706
95%         15.7             30       Skewness        9.17225
99%           30          135.8       Kurtosis       97.25202
```

Figure A 1: Monthly BOD concentration discharges of "MD0020524"

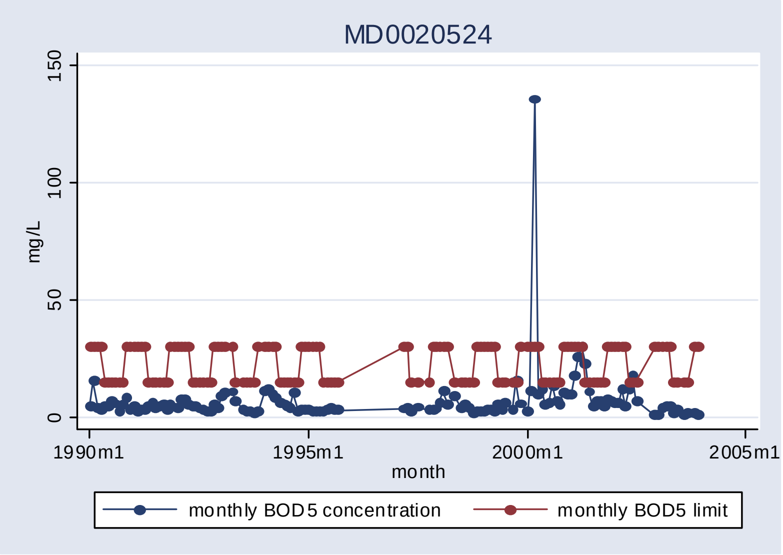


Table A 24: BOD concentration outlier of "PA0002551"

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%     1.900001       1.900001
 5%     2.700001       2.200001
10%     3.099999       2.500002       Obs                  99
25%     3.599998       2.599999       Sum of Wgt.          99

50%     5.400002                      Mean           10.70404
                        Largest       Std. Dev.      23.71032
75%     9.100002           30.4
90%         22.4       43.79999       Variance       562.1791
95%         29.6       48.09998       Skewness       8.080826
99%          229            229       Kurtosis       74.18888
```

Figure A 2: Monthly BOD concentration discharges of “PA0002551”

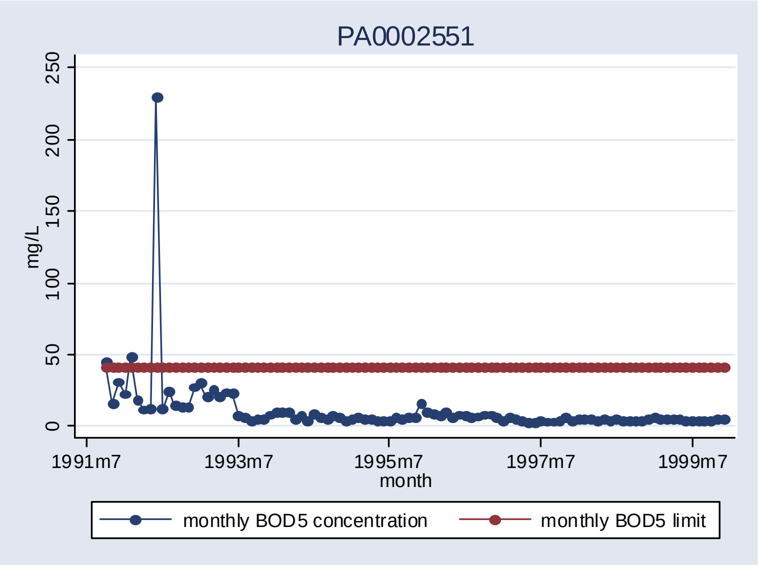


Table A 25: BOD concentration outlier of “VA0003077”

```
--------------------------------------------------------------
      Percentiles      Smallest
 1%            2            .28
 5%            3              2
10%            4              2       Obs                 161
25%            7              2       Sum of Wgt.         161

50%         13.4                      Mean           22.42451
                        Largest       Std. Dev.      34.48305
75%           23         115.75
90%         48.5         126.35       Variance        1189.08
95%     60.41667          178.5       Skewness       6.189763
99%        178.5         347.25       Kurtosis       53.18477
```

Figure A 3: Monthly BOD concentration discharges of “VA0003077”

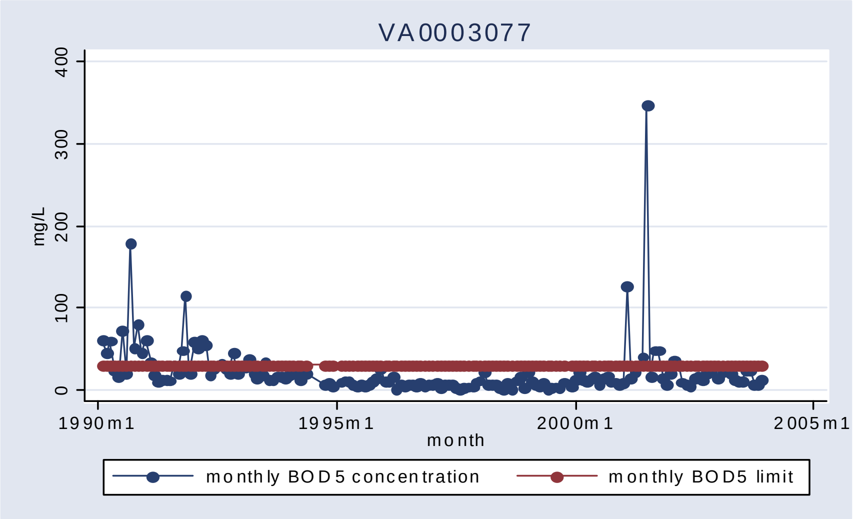


Table A 26: BOD concentration outlier of “VA0020362”

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%            3              2
 5%            4              3
10%          4.8              3       Obs                 161
25%            6              3       Sum of Wgt.         161

50%            8                      Mean           9.634161
                        Largest       Std. Dev.      7.514312
75%           11             27
90%           17             29       Variance       56.46489
95%           20             32       Skewness       5.232339
99%           32             78       Kurtosis       44.51351
```

Figure A 4: Monthly BOD concentration discharges of “VA0020362”

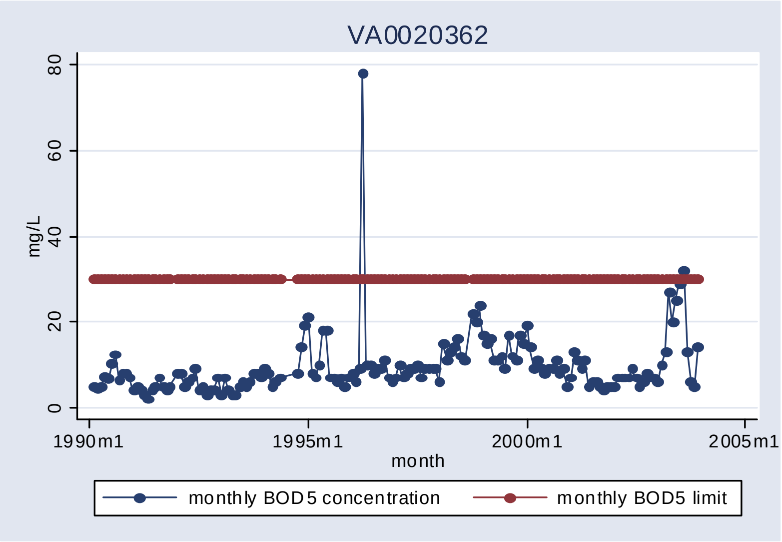


Table A 27: BOD concentration outlier of “VA0025194”

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%         8.23           8.23
 5%         10.3            8.6
10%         11.8           9.17       Obs                  95
25%           14           10.2       Sum of Wgt.          95

50%         19.3                      Mean           19.79589
                        Largest       Std. Dev.      8.388003
75%           24           30.6
90%         27.9           30.8       Variance        70.3586
95%           30           32.2       Skewness       3.335608
99%           77             77       Kurtosis       23.79144
```

Figure A 5: Monthly BOD concentration discharges of “VA0025194”

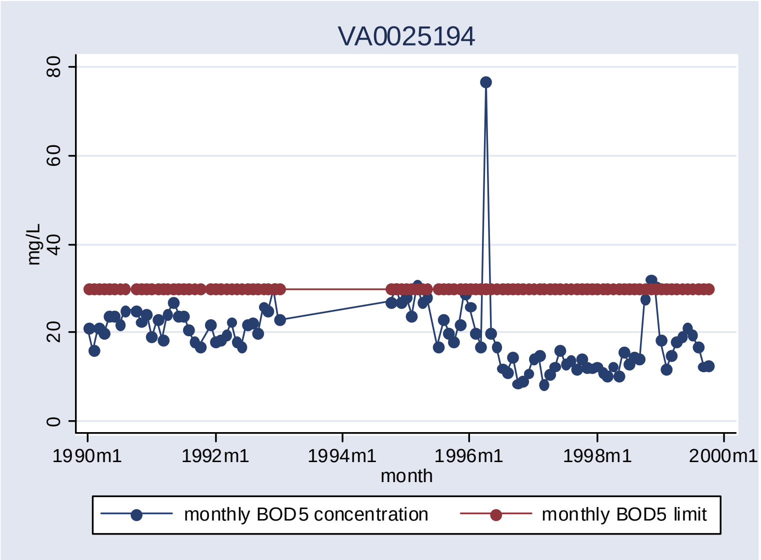


Table A 28: BOD concentration outlier of “MD0021610”

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%            2              2
 5%          2.6              2
10%          3.3            2.1       Obs                 164
25%          4.5            2.1       Sum of Wgt.         164

50%            7                      Mean           11.59878
                        Largest       Std. Dev.      31.79629
75%        13.95             27
90%           20           31.1       Variance       1011.004
95%           22             32       Skewness       11.92421
99%           32            408       Kurtosis       149.1136
```

Figure A 6: Monthly BOD concentration discharges of “MD0021610”

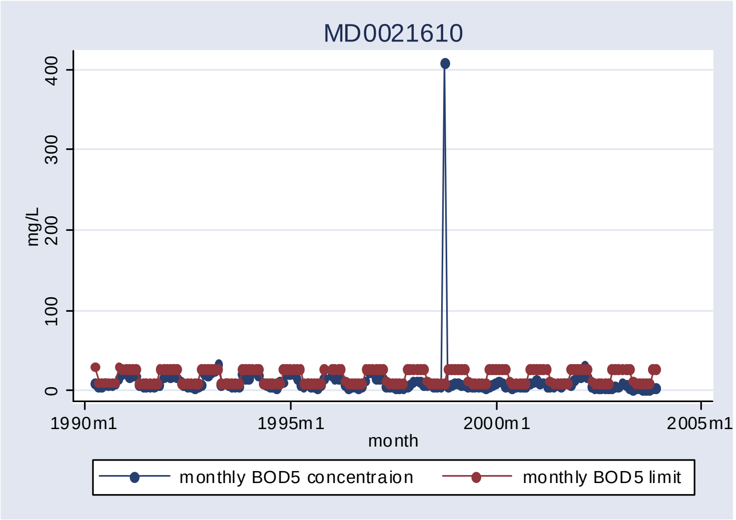


Table A 29: BOD concentration outlier of “VA0078484”

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%            7              7
 5%           10              7
10%           12              7       Obs                 129
25%           14              8       Sum of Wgt.         129

50%           17                      Mean           19.37984
                        Largest       Std. Dev.      9.548555
75%           22             40
90%           29             40       Variance        91.1749
95%           34             44       Skewness       3.670213
99%           44             91       Kurtosis       26.17601
```

Figure A 7: Monthly BOD concentration discharges of “VA0078484”

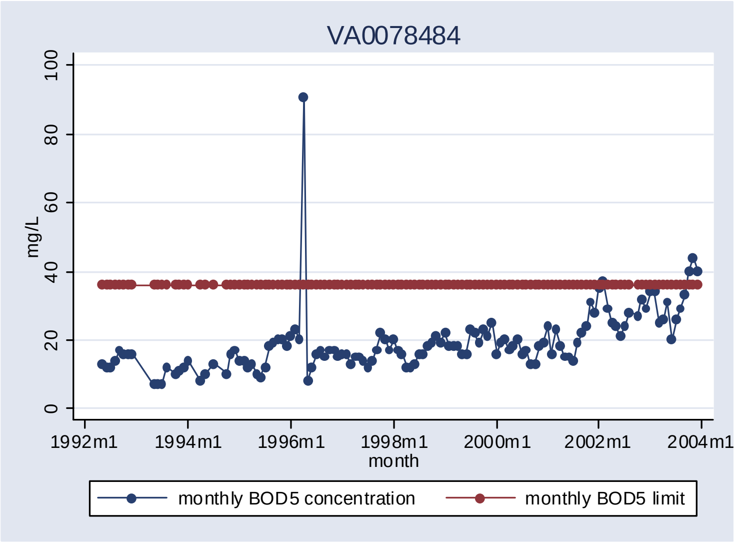


Table A 30: BOD quantity outlier of “MD0020524”

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%           .6            .25
 5%            6             .6
10%         12.3              1       Obs                 139
25%         17.6              3       Sum of Wgt.         139

50%         26.3                      Mean           79.59273
                        Largest       Std. Dev.      410.9134
75%           45            159
90%         85.1            162       Variance       168849.8
95%          119           1305       Skewness       10.58363
99%         1305           4708       Kurtosis       118.0625
```

Figure A 8: Monthly BOD quantity discharges of “MD0020524”

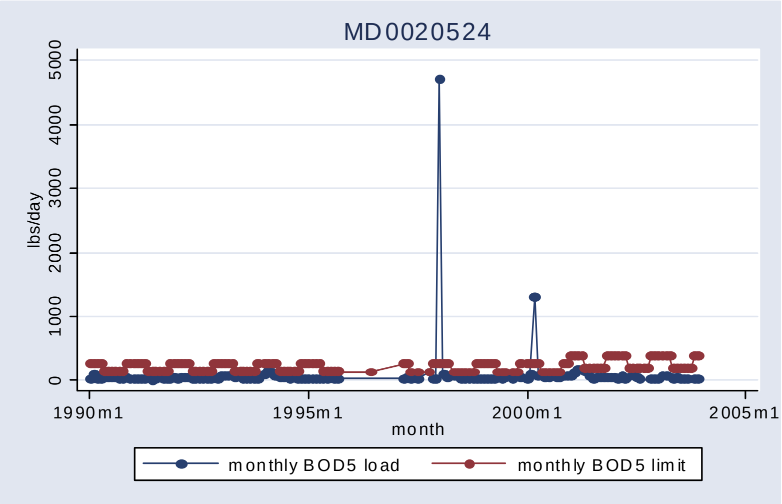


Table A 31: BOD quantity outlier of “PA0008419”

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%          249            238
 5%          329            249
10%          383            278       Obs                 165
25%          501            287       Sum of Wgt.         165

50%          775                      Mean           1270.212
                        Largest       Std. Dev.      5035.186
75%         1083           2670
90%         1582           2709       Variance       2.54e+07
95%         2171           3082       Skewness       12.52527
99%         3082          65214       Kurtosis       159.5619
```

Figure A 9: Monthly BOD quantity discharges of “PA0008419”

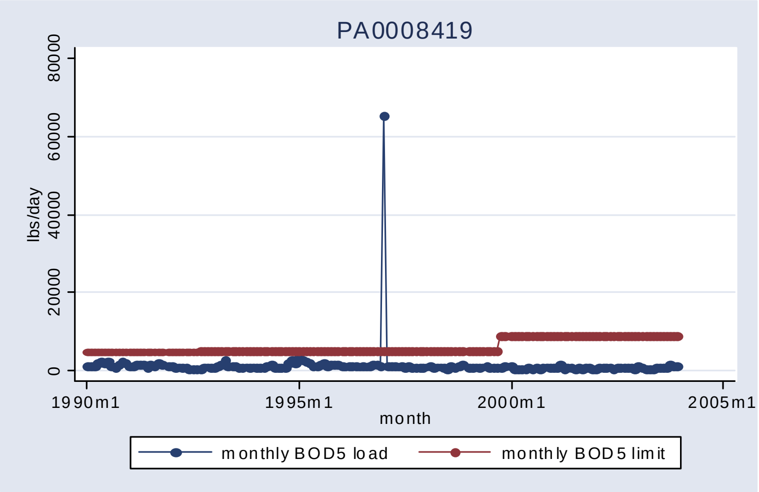


Table A 32: BOD quantity outlier of “VA0069345”

```
-------------------------------------------------------------
      Percentiles      Smallest
```

```
 1%            1              1
 5%            2              1
10%            5              1       Obs                  84
25%         32.5              2       Sum of Wgt.          84

50%     216.3601                      Mean           610.2518
                        Largest       Std. Dev.      3795.993
75%     308.7386       452.1683
90%     384.0999       469.2826       Variance       1.44e+07
95%     439.0409       534.8228       Skewness        8.98129
99%     34962.82       34962.82       Kurtosis       81.78234
```

Figure A 10: Monthly BOD quantity discharges of “VA0069345”

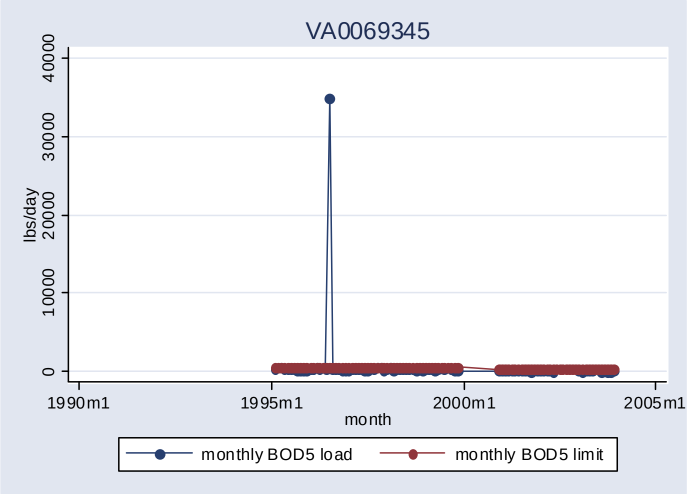


Table A 33: BOD5 concentration, mg/L

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%           .9             .1
 5%          1.4             .1
10%            2             .1       Obs               10329
25%            3            .17       Sum of Wgt.       10329

50%          5.7                      Mean            9.03746
                        Largest       Std. Dev.      14.76022
75%           11            396
90%           20            454       Variance        217.864
95%         25.5            472       Skewness       21.25007
99%           46            713       Kurtosis       774.5878
```

Table A 34: BOD5 quantity, lbs/day

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%         1.14          .0616
 5%         4.34           .065
10%         8.71             .1       Obs               12426
25%           24           .195       Sum of Wgt.       12426

50%           75                      Mean           451.6932
                        Largest       Std. Dev.      1883.381
75%          249          32491
90%          799          34802       Variance        3547124
95%         1533          39914       Skewness       10.89712
99%         5967          40522       Kurtosis        149.178
```

Table A 35: Monthly concentration permit, mg/L

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%            5              5
 5%           10              5
10%           13              5       Obs               10819
25%           24              5       Sum of Wgt.       10819

50%           30                      Mean           28.52151
                        Largest       Std. Dev.      13.68418
75%           30            111
90%           30            111       Variance       187.2567
95%           45            111       Skewness       2.891968
99%          100            111       Kurtosis        16.8978
```

Table A 36: Monthly quantity permit, lbs/day

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%        12.65            2.6
 5%         36.3            3.3
10%         87.6            4.4       Obs               12782
25%          227            4.4       Sum of Wgt.       12782

50%          454                      Mean           1452.622
                        Largest       Std. Dev.      4109.136
75%     1404.833          42400
90%         2498          42400       Variance       1.69e+07
95%         5020          42400       Skewness       8.213693
99%        10445          42400       Kurtosis       79.96706
```

Table A 37: Ratio of BOD5 concentration to effluent limit

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%     .0333333       .0019231
 5%     .0588235       .0038462
10%     .0769231       .0066667       Obs               10329
25%     .1296296       .0066667       Sum of Wgt.       10329

50%     .2272727                      Mean           .3277256
                        Largest       Std. Dev.      .5266216
75%     .4166667       15.23077
90%     .6833333       17.46154       Variance       .2773304
95%     .8506173       18.15385       Skewness       26.11969
99%     1.326667       27.42308       Kurtosis       1046.568
```

Table A 38: Ratio of BOD5 quantity to effluent limit

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%     .0080972       .0001585
 5%     .0253618       .0002859
10%     .0400837         .00034       Obs               12426
25%     .0809211       .0004323       Sum of Wgt.       12426

50%     .1731698                      Mean           .2681092
                        Largest       Std. Dev.      .4271887
75%     .3487143       10.84375
90%     .5871212           11.5       Variance       .1824902
95%      .755627          15.25       Skewness       20.07116
99%     1.270548        21.0625       Kurtosis       723.7531
```

Table A 39: Summary statistics of past 1, 2, and 3 years downstream DO

```
Variable                                 Obs      Mean Std. Dev.       Min        Max
past year mean downstream DO           14892  9.555377  1.159611     5.642       13.3
past year median downstream DO         14892  9.420012  1.330624      4.62       13.6
past two years mean downstream DO      13920  9.565388  1.102955  6.046087   12.37083
past two years median downstream DO    13920  9.419422  1.227003       5.5       12.6
past three years mean downstream DO    12924  9.559568  1.085171  6.269143   12.37857
past three years median downstream DO  12924  9.403561  1.177163      5.45       12.5
```

Table A 40: Non-zero monthly inspections count, by plant

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%            1              1
 5%            1              1
10%            1              1       Obs                3047
25%            1              1       Sum of Wgt.        3047

50%            1                      Mean           1.334755
                        Largest       Std. Dev.      .6539313
75%            2              5
90%            2              6       Variance       .4276262
95%            3              7       Skewness       3.768726
99%            3             13       Kurtosis       41.56491
```

Table A 41: Log ratio of BOD5 concentration to effluent limit

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%    -3.401197      -6.253829
 5%    -2.833213      -5.560682
10%    -2.564949      -5.010635       Obs               10329
25%    -2.043074      -5.010635       Sum of Wgt.       10329

50%    -1.481604                      Mean          -1.486689
                        Largest       Std. Dev.      .8515072
75%    -.8754688       2.723318
90%    -.3807725       2.860001       Variance       .7250644
95%     -.161793       2.898882       Skewness       -.071097
99%     .2826695       3.311385       Kurtosis       3.313881
```

Table A 42: Log ratio of BOD5 quantity to effluent limit

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%    -4.816241      -8.749587
 5%     -3.67451      -8.160018
10%    -3.216785      -7.986565       Obs               12426
25%    -2.514281      -7.746424       Sum of Wgt.       12426

50%    -1.753483                      Mean          -1.833837
                        Largest       Std. Dev.      1.087184
75%    -1.053502       2.383589
90%     -.532524       2.442347       Variance       1.181969
95%    -.2802074        2.72458       Skewness      -.6177647
99%     .2394483       3.047494       Kurtosis       4.433562
```

Table A 43: Log design flow

```
-------------------------------------------------------------
```

```
      Percentiles      Smallest
 1%    -3.506558      -3.506558
 5%     -1.07881      -2.224699
10%    -.5108256      -2.120264       Obs                  97
25%     .4054651       -1.12393       Sum of Wgt.          97

50%     1.061257                      Mean           1.445924
                        Largest       Std. Dev.      1.877767
75%     2.484907       5.908083
90%     3.703029       6.431331       Variance        3.52601
95%     5.347107       7.118016       Skewness       .8166391
99%     7.600903       7.600903       Kurtosis       4.554573
```

Table A 44: Summary statistics of log of past 1, 2, and 3 years downstream DO

Variable	Obs	Mean	Std. Dev.	Min	Max
ln past year mean downstream DO	14892	2.249279	0.12732	1.730239	2.587764
ln past year median downstream DO	14892	2.232173	0.149072	1.530395	2.61007
ln past two years mean downstream DO	13920	2.251064	0.121183	1.799411	2.515342
ln past two years median downstream DO	13920	2.23362	0.138259	1.704748	2.533697
ln past three years mean downstream DO	12924	2.250653	0.119565	1.83564	2.515967
ln past three years median downstream DO	12924	2.23263	0.132917	1.695616	2.525729

Inspections model

$$Freq.(I_{it}) = I\left(\frac{1}{6}\sum_{l=1}^{6}\left(BOD_{ij(t-l)}\Big/P_{i(t-l)}\right), designflow_i, sic_i, pri_i, demographics_{ik}, \nu_y, \sigma_R\right)$$

I_{it} = number of inspections conducted in month t, against plant i

$\frac{1}{6}\sum_{l=1}^{6} BOD_{ij(t-l)}\Big/P_{i(t-l)}$ = past 2 quarter average relative discharge

sic_i = dummy variable for SIC - code of plant

pri_i = private versus public ownership

ν_y = dummy variable for year y, with $y = 1,2....13$

σ_R = state level dummy variable, with $R = 1,2$

The regulator is most likely unaware of the polluter's performance or its discharge levels during the same month as when he/she decides to inspect the facility. Hence, the inspection decision in month t is determined by the facility's discharge behavior in the past months rather than the facility's performance in the current month. On the other hand, a polluter in month t cannot be aware of the inspection decision by the EPA or the State regulator, in the very same month. In this sense, inspections and effluent discharges are not simultaneously determined in month t. Average discharges relative to its corresponding permit level, past two quarters ago is considered as the indicator of plant level performance. This lag length is chosen because the EPA or the authorized states report noncompliance on a quarterly basis (Quarterly Noncompliance Reports (QNCRs)). For example, Earnhart in his 2004c paper justifies inclusion of performance over the six-month time period based on the EPA tracking facilities with its QNCRs, which is used to guide inspection and enforcement decisions.

Lagged average relative discharges (concentration or quantity), is expected to have a positive effect on current period probability of inspection. Ceteris paribus, the higher is average discharges relative to effluent limit, the higher is the ‘probability’ that the plant is either noncompliant, or its ‘degree’ of compliance is lower. This implies that the chances that an inspector will be visiting the plant are higher, than if it had lower lagged average relative discharges. Other plant specific factors such as design flow (capturing size of the abatement technology), type of plant, and private versus public are also considered as potential determinants of current period frequency of inspections. Ceteris paribus, plants with larger design flow are expected to pollute more, which doesn’t necessarily mean that they will be in noncompliance since they face higher permit levels.[1] Location specific factors such as attributes of the neighboring (zip-code level) community are also included as potential factors influencing inspection probability in any given month t. Year dummies are included in order to capture any trends in the inspections over time. Lastly, state level differences e.g. in regulatory policy as well as efficiency is controlled for by state dummies. Finally, lagged ambient water quality is not included on account of its influence on lagged average relative discharges, the measure of performance.

A Negative binomial model is used to estimate frequency of monthly inspections for the 77 and 85 plants with relative concentration and quantity discharges. Robust standard errors are reported to account for (unobserved) heterogeneity in the model specification, since it is likely that there are some ‘omitted’ variables that determine why a plant faces more monitoring activities than others. Results presented in indicate that two quarters-lagged average performance measured by relative discharges of both BOD5 concentration and quantity, actually does not have a significant impact on monthly inspections. The substantial degree of ‘overcompliance’ that characterizes the present sample of plants for concentration and quantity discharges might explain why past performance is not significant in inspections decisions, ex-post. Table A3 through A6 show that plants facing BOD5 concentration limits are polluting monthly 6 mg/L, on an average, while they are allowed 30 mg/L (on an average). Similarly, average quantity discharges are 75 lbs/day compared to average limits around 454 lbs/day. Design effluent flow, on the other hand, has a significant impact on actual inspections conducted, but only for relative quantity discharges. Bigger plants seem to be facing less monitoring activities thereby, indicating that plants that can accommodate more pollution, might be more efficient in controlling pollution, not to mention that they also face higher effluent limits. Community specific factors at the zip-code level exert a significant influence in both models. As expected, the type of the plant seems to explain differences in the average frequency of inspections significantly (more so, for relative concentration than quantity). Type of ownership is not significant in either model. State level differences in frequency of monitoring activities are significant (for the current sample of plants), in both models with plants in MD and PA, facing higher inspections than those in VA. Whereas, no clear time pattern in inspections frequency emerges.

---

[1] In addition, if younger plants are also larger due to technology improvements, then the size variable might be capturing the vintage of the technology, and hence might be inspected less than the older and smaller plants with less efficient abatement equipment.

Table A 45: Monthly inspections count, by plant

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0              0
 5%            0              0
10%            0              0       Obs               16800
25%            0              0       Sum of Wgt.       16800

50%            0                      Mean           .2420833
                        Largest       Std. Dev.      .5848682
75%            0              5
90%            1              6       Variance       .3420708
95%            1              7       Skewness        3.35017
99%            2             13       Kurtosis       25.89265
```

Table A 46: Negative binomial estimation of monthly inspections counts regressed on performance, design flow etc.

	Monthly inspections count	Monthly inspections count
Average performance two quarters ago for plants with concentration limits	0.015	
	(0.13)	
Average performance two quarters ago for plants with quantity limits		-0.002
		(-0.03)
Design flow	0.000	-0.003
	(0.18)	(-3.37)**
rnwhite	0.003	0.003
	(1.81)+	(1.47)
mhhi	-0.009	-0.006
	(-2.54)*	(-1.72)+
carpl	0.005	0.003
	(0.74)	(0.45)
manuf	0.008	0.007
	(2.90)**	(2.65)**
popt	0.010	0.008
	(5.23)**	(4.51)**
elec	-0.762	-0.401
	(-2.59)**	(-0.53)
mill	-0.444	0.241
	(-1.24)	(0.45)
paper	-0.285	0.447
	(-0.83)	(0.84)
chem	-0.191	0.574
	(-0.77)	(1.10)
petro	-1.250	-0.119
	(-4.19)**	(-0.22)
rubber	-0.962	
	(-2.07)*	
leather	0.232	1.022
	(0.77)	(1.90)+
metal	-0.970	0.005

	(-2.34)*	(0.01)
transp	-1.010	
	(-2.55)*	
secu	-0.440	-0.438
	(-3.39)**	(-3.38)**
just	-0.853	-0.780
	(-4.22)**	(-3.85)**
pri	0.354	-0.645
	(1.39)	(-1.25)
MD	0.462	0.390
	(6.37)**	(5.45)**
PA	0.230	0.478
	(2.27)*	(5.48)**
yeardum1	-0.185	-0.184
	(-0.88)	(-0.93)
yeardum2	-0.061	0.083
	(-0.33)	(0.47)
yeardum3	0.045	0.106
	(0.25)	(0.61)
yeardum4	-0.043	0.069
	(-0.23)	(0.39)
yeardum5	0.118	0.233
	(0.63)	(1.31)
yeardum6	0.070	0.204
	(0.38)	(1.15)
yeardum7	-0.132	0.065
	(-0.70)	(0.36)
yeardum8	0.046	0.190
	(0.25)	(1.08)
yeardum10	0.125	0.224
	(0.68)	(1.28)
yeardum11	0.123	0.268
	(0.65)	(1.49)
yeardum12	-0.080	0.016
	(-0.39)	(0.08)
yeardum13	-0.190	-0.150
	(-0.85)	(-0.70)
yeardum9	0.276	0.459
	(1.52)	(2.67)**
food		0.438
		(0.83)
N	8376	9817
chi2	211.87	218.15
p	0.00	0.00

Table A 47: Absolute discharges model for full sample

	(1)	(2)	(3)	(4)
	log monthly BOD5 concentration	log monthly BOD5 concentration	log monthly BOD5 quantity	log monthly BOD5 quantity
log monthly concentration permit	0.547	0.552		
	(0.000)**	(0.000)**		

log monthly quantity permit			0.726 (0.000)**	0.734 (0.000)**
log past three years mean downstream DO	0.942 (0.000)**		1.311 (0.000)**	
log past three years median downstream DO		0.946 (0.000)**		1.425 (0.000)**
log design flow	0.073 (0.000)**	0.071 (0.000)**	0.223 (0.000)**	0.217 (0.000)**
Test for coefficient on log permits = 1	chi2(1) = 158.98	chi2(1) = 155.63	chi2(1) = 76.88	chi2(1) = 72.26
Prob > chi2	0.0000	0.0000	0.0000	0.0000
Observations	6439	6439	7497	7497
# of plants	77	77	85	85

Table A 48: Absolute discharges model with at least 25% monthly observations

	(1)	(2)	(3)	(4)
	log monthly BOD5 concentration	log monthly BOD5 concentration	log monthly BOD5 quantity	log monthly BOD5 quantity
log monthly concentration permit	0.480 (0.000)**	0.491 (0.000)**		
log monthly quantity permit			0.718 (0.000)**	0.728 (0.000)**
log past three years mean downstream DO	0.932 (0.000)**		1.054 (0.000)**	
log past three years median downstream DO		1.009 (0.000)**		1.249 (0.000)**
log design flow	-0.007 (0.741)	-0.011 (0.587)	0.210 (0.000)**	0.204
Test for coefficient on log permits = 1	chi2(1) = 153.21	chi2(1) = 148.40	chi2(1) = 68.82	chi2(1) = 64.17

Prob > chi2	0.0000	0.0000	0.0000	0.0000
Observations	5600	5600	6590	6590
Number of group(npdes)	60	60	67	67

Table A 49: Absolute discharges model with at least 50% observations

	(1)	(2)	(3)	(4)
	log monthly BOD5 concentration	log monthly BOD5 concentration	log monthly BOD5 quantity	log monthly BOD5 quantity
log monthly concentration permit	0.460 (0.000)**	0.474 (0.000)**		
log monthly quantity permit			0.856 (0.000)**	0.852 (0.000)**
log past three years mean downstream DO	1.318 (0.000)**		1.410 (0.000)**	
log past three years median downstream DO		1.347 (0.000)**		1.605 (0.000)**
log design flow	-0.009 (0.711)	-0.016 (0.488)	0.057 (0.148)	0.055 (0.167)
Test for coefficient on log permits = 1	chi2(1)= 132.04	chi2(1)= 127.08	chi2(1)= 13.00	chi2(1)= 13.57
Prob > chi2	0.0000	0.0000	0.0003	0.0002
Observations	4164	4164	4713	4713
Number of group(npdes)	41	41	46	46

Table A 50: FGLS estimation of log of relative BOD discharge with past three years average water quality, etc (full sample)

	(1)	(2)	(3)	(4)
	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 quantity to effluent limit	ln ratio of BOD5 quantity to effluent limit
ln past three years mean downstream DO	1.301 (7.46)**		1.558 (7.43)**	
ln past three years median downstream DO		1.257 (8.21)**		1.657 (8.76)**

ln design flow	0.078	0.075	0.024	0.023
	(4.49)**	(4.34)**	(0.99)	(0.95)
predicted inspections	-1.401	-1.312	-2.715	-2.667
	(2.24)*	(2.11)*	(4.18)**	(4.09)**
lrnwhite	-0.067	-0.062	-0.032	-0.023
	(2.97)**	(2.79)**	(1.17)	(0.83)
lmhhi	-0.513	-0.481	-0.301	-0.248
	(5.28)**	(4.94)**	(2.79)**	(2.29)*
lcarpl	0.012	-0.018	-0.335	-0.359
	(0.14)	(0.22)	(3.35)**	(3.59)**
lmanuf	0.241	0.236	0.325	0.329
	(4.98)**	(4.92)**	(5.08)**	(5.16)**
lpopt	0.155	0.149	0.187	0.186
	(4.06)**	(3.91)**	(4.74)**	(4.69)**
elec	-0.316	-0.294	-1.945	-1.903
	(2.42)*	(2.27)*	(4.11)**	(4.06)**
food	0.385	0.313	-0.341	-0.374
	(1.86)+	(1.53)	(2.09)*	(2.29)*
mill	-0.211	-0.228	-0.090	-0.139
	(1.43)	(1.56)	(0.84)	(1.29)
paper	0.792	0.781	0.871	0.851
	(5.24)**	(5.23)**	(9.30)**	(9.08)**
chem	-0.115	-0.139	-0.383	-0.405
	(0.84)	(1.03)	(3.50)**	(3.71)**
petro	-0.756	-0.767	-1.078	-1.096
	(3.36)**	(3.43)**	(4.93)**	(4.99)**
rubber	0.399	0.375	-2.762	-2.802
	(1.28)	(1.21)	(8.58)**	(8.71)**
leather	-0.022	-0.034	0.019	0.048
	(0.09)	(0.14)	(0.09)	(0.23)
metal	-1.109	-1.125	-0.832	-0.871
	(6.06)**	(6.19)**	(3.37)**	(3.53)**
transp	-0.565	-0.619	0.000	0.000
	(1.75)+	(1.92)+	(.)	(.)
secu	-0.009	0.012	-0.162	-0.145
	(0.08)	(0.11)	(1.19)	(1.06)
just	-1.272	-1.244	-1.592	-1.579
	(5.77)**	(5.68)**	(7.06)**	(7.00)**
winter	0.198	0.199	0.337	0.338
	(9.21)**	(9.24)**	(13.32)**	(13.36)**
spring	0.102	0.102	0.152	0.152
	(5.43)**	(5.44)**	(6.92)**	(6.95)**
fall	0.036	0.036	0.030	0.030
	(1.89)+	(1.89)+	(1.37)	(1.36)
MD	0.208	0.192	0.254	0.240
	(1.98)*	(1.84)+	(2.54)*	(2.40)*
PA	-0.154	-0.122	0.164	0.210
	(1.39)	(1.11)	(1.28)	(1.64)
year== 1990.0000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)
year== 1991.0000	0.000	0.000	0.000	0.000

	(.)	(.)	(.)	(.)
year== 1992.0000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)
year== 1993.0000	-0.211	-0.220	-0.189	-0.207
	(2.47)*	(2.59)**	(1.82)+	(2.00)*
year== 1994.0000	-0.053	-0.063	-0.049	-0.067
	(0.81)	(0.96)	(0.59)	(0.80)
year== 1995.0000	-0.074	-0.084	-0.073	-0.091
	(1.07)	(1.22)	(0.85)	(1.06)
year== 1996.0000	-0.106	-0.098	-0.037	-0.030
	(1.17)	(1.08)	(0.37)	(0.30)
year== 1997.0000	-0.062	-0.057	-0.192	-0.186
	(0.96)	(0.88)	(2.45)*	(2.38)*
year== 1999.0000	-0.087	-0.091	-0.179	-0.186
	(1.56)	(1.64)	(2.42)*	(2.51)*
year== 2000.0000	-0.124	-0.132	-0.280	-0.293
	(1.88)+	(2.02)*	(3.40)**	(3.56)**
year== 2001.0000	-0.264	-0.272	-0.516	-0.531
	(2.78)**	(2.88)**	(4.30)**	(4.42)**
year== 2002.0000	-0.416	-0.432	-0.842	-0.864
	(3.55)**	(3.72)**	(5.65)**	(5.79)**
year== 2003.0000	-0.369	-0.392	-0.834	-0.862
	(3.37)**	(3.60)**	(5.77)**	(5.97)**
Constant	-3.281	-3.179	-3.953	-4.282
	(5.57)**	(5.74)**	(5.06)**	(5.75)**
Observations	6439	6439	7497	7497
Number of group(npdes)	77	77	85	85

Table A 51: Log of relative BOD discharge with past three years average water quality, etc (>=25% monthly observations)

	(1)	(2)	(3)	(4)
	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 quantity to effluent limit	ln ratio of BOD5 quantity to effluent limit
ln past three years mean downstream DO	1.411 (7.66)**		1.390 (6.09)**	
ln past three years median downstream DO		1.411 (8.76)**		1.538 (7.52)**

ln design flow	0.019 (0.90)	0.013 (0.64)	0.018 (0.72)	0.018 (0.70)
predicted inspections	-0.831 (1.25)	-0.651 (0.98)	-3.090 (4.69)**	-3.033 (4.58)**
Constant	-2.576	-2.627	-3.775	-4.316
	(3.89)**	(4.19)**	(4.19)**	(5.00)**
Observations	5600	5600	6590	6590
Number of group(npdes)	60	60	67	67

Table A 52 : Log of relative BOD discharge with past three years average water quality, etc (>=50% monthly observations)

	(1)	(2)	(3)	(4)
	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 quantity to effluent limit	ln ratio of BOD5 quantity to effluent limit
ln past three years mean downstream DO	2.001 (7.97)**		1.395 (4.84)**	
ln past three years median downstream DO		1.903 (9.05)**		1.593 (6.14)**
ln design flow	0.022 (0.90)	0.014 (0.57)	-0.045 (1.83)+	-0.049 (2.01)*
predicted inspections	-0.588 (0.67)	-0.268 (0.31)	-2.681 (2.91)**	-2.408 (2.59)**
Constant	-4.949	-4.887	-5.572	-6.355
	(5.83)**	(6.26)**	(4.66)**	(5.51)**
Observations	4164	4164	4713	4713
Number of group(npdes)	41	41	46	46

Table A 53: Log of relative BOD5 discharge with past year average water quality, etc (full sample)

	(1)	(2)	(3)	(4)
	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 quantity to effluent limit	ln ratio of BOD5 quantity to effluent limit
ln past year mean downstream DO	0.986 (6.70)**		1.214 (6.94)**	
ln past year		0.797		1.037

median downstream DO		(6.64)**		(6.93)**
ln design flow	0.079 (4.57)**	0.079 (4.60)**	0.027 (1.11)	0.024 (1.01)
predicted inspections	-0.858 (1.40)	-0.883 (1.43)	-2.631 (3.15)**	-2.560 (3.06)**
Constant	-2.500 (4.55)**	-2.054 (4.00)**	-2.644 (3.70)**	-2.287 (3.32)**
Observations	7659	7659	8988	8988
Number of group(npdes)	77	77	85	85

Table A 54: Log of relative BOD discharge with past year average water quality, etc (>=25% monthly observations)

	(1)	(2)	(3)	(4)
	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 quantity to effluent limit	ln ratio of BOD5 quantity to effluent limit
ln past year mean downstream DO	1.067 (6.81)**		1.079 (5.68)**	
ln past year median downstream DO		0.897 (6.95)**		0.927 (5.73)**
ln design flow	0.027 (1.36)	0.027 (1.33)	0.016 (0.63)	0.014 (0.56)
predicted inspections	-0.368 (0.57)	-0.332 (0.51)	-3.218 (3.76)**	-3.105 (3.63)**
Constant	-1.879	-1.495	-2.634	-2.334
	(3.05)**	(2.57)*	(3.28)**	(2.99)**
Observations	6760	6760	7951	7951
Number of group(npdes)	60	60	67	67

Table A 55: Log of relative BOD discharge with past year average water quality, etc (>=50% monthly observations)

	(1)	(2)	(3)	(4)
	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 quantity to effluent limit	ln ratio of BOD5 quantity to effluent limit
ln past year	1.219		1.025	

mean downstream DO	(6.24)**		(4.62)**	
ln past year median downstream DO		0.997 (6.24)**		0.985 (5.11)**
ln design flow	0.059 (2.65)**	0.061 (2.71)**	-0.034 (1.49)	-0.035 (1.53)
predicted inspections	-1.352 (1.61)	-1.338 (1.58)	-2.837 (3.43)**	-2.700 (3.22)**
Constant	-2.868	-2.412	-4.516	-4.563
	(3.94)**	(3.52)**	(4.55)**	(4.74)**
Observations	5038	5038	5725	5725
Number of group(npdes)	41	41	46	46

Table A 56: Log of relative BOD discharge with past two years average water quality, etc (full sample)

	(1)	(2)	(3)	(4)
	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 quantity to effluent limit	ln ratio of BOD5 quantity to effluent limit
ln past two years mean downstream DO	1.132 (6.44)**		1.402 (6.91)**	
ln past two years median downstream DO		1.163 (7.86)**		1.498 (8.49)**
ln design flow	0.077 (4.39)**	0.075 (4.30)**	0.026 (1.13)	0.029 (1.23)
predicted inspections	-1.125 (1.80)+	-0.969 (1.55)	-2.522 (3.35)**	-2.360 (3.13)**
Constant	-2.822	-2.968	-3.359	-3.746
	(4.68)**	(5.31)**	(4.33)**	(5.11)**
Observations	7063	7063	8262	8262
Number of group(npdes)	77	77	85	85

Table A 57: Log of relative BOD discharge with past two years average water quality, etc (>=25% monthly observations)

	(1)	(2)	(3)	(4)
	ln ratio of BOD5 concentration	ln ratio of BOD5 concentration	ln ratio of BOD5 quantity to	ln ratio of BOD5 quantity to

	to effluent limit	to effluent limit	effluent limit	effluent limit
ln past two years mean downstream DO	1.221 (6.64)**		1.272 (5.85)**	
ln past two years median downstream DO		1.269 (8.07)**		1.406 (7.36)**
ln design flow	0.020 (0.99)	0.018 (0.89)	0.018 (0.73)	0.021 (0.88)
predicted inspections	-0.608 (0.92)	-0.394 (0.59)	-2.978 (3.88)**	-2.846 (3.70)**
Constant	-2.152	-2.357	-3.354	-3.883
	(3.21)**	(3.75)**	(3.83)**	(4.64)**
Observations	6188	6188	7310	7310
Number of group(npdes)	60	60	67	67

Table A 58: Log of relative BOD discharge with past two years average water quality, etc (>=50% monthly observations)

	(1)	(2)	(3)	(4)
	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 concentration to effluent limit	ln ratio of BOD5 quantity to effluent limit	ln ratio of BOD5 quantity to effluent limit
ln past two years mean downstream DO	1.779 (7.33)**		1.259 (4.69)**	
ln past two years median downstream DO		1.795 (8.97)**		1.629 (6.81)**
ln design flow	0.029 (1.23)	0.021 (0.94)	-0.039 (1.67)+	-0.039 (1.68)+
predicted inspections	-0.861 (1.00)	-0.466 (0.54)	-2.684 (3.02)**	-2.221 (2.50)*
Constant	-4.330	-4.585	-5.075	-6.378
	(5.23)**	(6.05)**	(4.45)**	(5.83)**
Observations	4614	4614	5246	5246
Number of group(npdes)	41	41	46	46

Table A 59: Distance between downstream and upstream stations

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%          1.15           1.15
```

```
         5%          1.82       1.416667
        10%          3.06         1.4875       Obs                  76
        25%        7.8875           1.82       Sum of Wgt.          76

        50%         13.96                      Mean           18.90396
                                 Largest       Std. Dev.      15.10163
        75%         26.63          53.97
        90%          35.9           56.2       Variance       228.0592
        95%         53.97           63.8       Skewness       1.191185
        99%          64.5           64.5       Kurtosis       4.061753
--------------------------------------------------------------------------------------
```

Table A 60: Distance between upstream station and first/only plant

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%          .07            .07
 5%           .2             .1
10%           .5             .2       Obs                  76
25%          1.5             .2       Sum of Wgt.          76

50%        5.375                      Mean             8.1675
                        Largest       Std. Dev.      8.632468
75%       12.505           27.6
90%           20          27.92       Variance       74.51951
95%         27.6             28       Skewness       1.713952
99%           46             46       Kurtosis       6.754434
```

Table A 61: Distance between first/only and second plant/downstream station

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%           .1             .1
 5%           .5            .12
10%           .9             .2       Obs                  76
25%         2.42             .5       Sum of Wgt.          76

50%         4.55                      Mean           8.709211
                        Largest       Std. Dev.       9.52607
75%         10.5             30
90%         23.4             30       Variance       90.74601
95%           30             35       Skewness       1.795313
99%           48             48       Kurtosis        6.26358
```

Table A 62: Distance between second and third plant

```
Variable       Obs    Mean           Std. Dev.      Min     Max


thirddist      17     10.20765       10.0684         0*      32
```

Note *: There is one set of three plants on the same segment of the river, with the second and third plants on the tributary and only the first plant on the main river. In this case, distance between the second and the third plant is considered as zero because their locations are approximated by the point of confluence of the tributary with the main stem.

Table A 63: Distance between third plant and downstream station

```
Variable       Obs    Mean    Std. Dev.      Min    Max

fourthdist     4      7.865   4.951724       3.9    15.1
```

Table A 64: Distance between upstream station and plant

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%          .07            .07
 5%           .2             .1
10%           .7             .2       Obs                  97
25%          2.3             .2       Sum of Wgt.          97

50%            7                      Mean           10.08247
                        Largest       Std. Dev.      9.808414
75%           14           28.6
90%           25           36.2       Variance       96.20499
95%           28           41.5       Skewness       1.329763
99%           46             46       Kurtosis       4.632165
```

Table A 65: Distance between plant and downstream station

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%           .1             .1
 5%           .5            .12
10%         1.06             .4       Obs                  97
25%          3.5             .5       Sum of Wgt.          97

50%          8.2                      Mean           11.35711
                        Largest       Std. Dev.      11.06274
75%         16.4           39.5
90%         26.9           43.4       Variance       122.3841
95%           35             45       Skewness        1.36898
99%           48             48       Kurtosis       4.367741
```

Table A 66: Count of monthly DO data across downstream and upstream stations, by state and year

Year	downstream stations in MD	upstream stations in MD	downstream stations in PA	upstream stations in PA	downstream stations in VA	upstream stations in VA
1990	237	234	155	154	327	244
1991	238	239	147	155	336	286
1992	238	240	149	163	407	349
1993	237	238	163	168	437	373
1994	231	235	159	164	436	385
1995	235	239	157	159	454	386
1996	222	232	154	156	442	377
1997	232	238	139	151	428	361
1998	232	234	121	133	461	395
1999	223	209	72	78	457	412
2000	236	232	79	77	463	409
2001	232	234	67	68	313	318
2002	234	230	70	62	241	258
2003	234	235	98	84	246	241

Table A 67: Dissolved oxygen at downstream and upstream stations, by state

State	Variable	Obs	Mean	Std. Dev.	Min	Max
MD	downstream water quality DO, mg/L	3261	9.48418	2.234304	4.1	18.4
MD	upstream water quality DO, mg/L	3269	9.465483	2.389385	1.9	18.4
PA	downstream water quality DO, mg/L	1730	9.613098	2.543408	3.7	19

PA	upstream water quality DO, mg/L	1772	10.219	2.478831	3.95	18.79
VA	downstream water quality DO, mg/L	5448	9.624584	2.412434	1.41	18.29
VA	upstream water quality DO, mg/L	4794	9.668282	2.366951	1.24	18

Table A 68: Dissolved oxygen at downstream and upstream stations, by year

Year	Variable	Obs	Mean	Std. Dev.	Min	Max
1990	downstream water quality DO, mg/L	719	9.589917	2.382561	2.9	16.8
1990	upstream water quality DO, mg/L	632	9.691535	2.488157	2.7	17.6
1991	downstream water quality DO, mg/L	721	9.387032	2.448768	2.52	19
1991	upstream water quality DO, mg/L	680	9.560132	2.428343	2.33	15.5
1992	downstream water quality DO, mg/L	794	9.640806	2.249488	2.18	16
1992	upstream water quality DO, mg/L	752	9.845705	2.292976	3.4	16.6
1993	downstream water quality DO, mg/L	837	9.515233	2.460607	1.41	17.3
1993	upstream water quality DO, mg/L	779	9.726149	2.469583	2.22	16.8
1994	downstream water quality DO, mg/L	826	9.724467	2.399914	2.01	16.2
1994	upstream water quality DO, mg/L	784	9.873795	2.399377	2.53	16.2
1995	downstream water quality DO, mg/L	846	9.543652	2.453608	1.57	15.32
1995	upstream water quality DO, mg/L	784	9.705153	2.447756	1.24	15.6
1996	downstream water quality DO, mg/L	818	9.774303	2.386188	1.88	18.4
1996	upstream water quality DO, mg/L	765	9.885438	2.367991	1.9	18.4
1997	downstream water quality DO, mg/L	799	9.836984	2.351226	2.18	17.4
1997	upstream water quality DO, mg/L	750	9.922147	2.385516	3.39	16.4
1998	downstream water quality DO, mg/L	814	9.379177	2.147382	1.47	15.4
1998	upstream water quality DO, mg/L	762	9.50689	2.13852	3	15.4
1999	downstream water quality DO, mg/L	752	9.442367	2.315224	2.59	18.29
1999	upstream water quality DO, mg/L	699	9.429027	2.357776	2.48	16.6
2000	downstream water quality DO, mg/L	778	9.555746	2.46614	3.5	17.5
2000	upstream water quality DO, mg/L	718	9.655613	2.449274	2.8	17.5
2001	downstream water quality DO, mg/L	612	9.503627	2.363719	2.83	15.7
2001	upstream water quality DO, mg/L	620	9.533452	2.439077	3.03	16.7
2002	downstream water quality DO, mg/L	545	9.336211	2.443742	3.26	16.6
2002	upstream water quality DO, mg/L	550	9.379736	2.482058	3.38	18
2003	downstream water quality DO, mg/L	578	9.819118	2.433916	3.56	18.79
2003	upstream water quality DO, mg/L	560	10.00459	2.556956	2.9	18.79

Table A 69: Dissolved oxygen at downstream stations, with outliers

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%          4.4           1.41
 5%          5.9           1.47
10%         6.62           1.57       Obs               10527
25%          7.8           1.64       Sum of Wgt.       10527

50%          9.5                      Mean            9.62581
                        Largest       Std. Dev.      2.508877
75%         11.5          37.75
90%         12.8          39.75       Variance       6.294462
95%        13.48           42.1       Skewness       1.030345
99%         14.9           42.1       Kurtosis       13.77442
```

Table A 70: Dissolved oxygen at upstream stations

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%         4.48           1.24
 5%         5.85            1.9
10%          6.6           2.22       Obs                9921
25%          7.9           2.33       Sum of Wgt.        9921

50%          9.7                      Mean           9.723255
                        Largest       Std. Dev.      2.417031
75%         11.6           17.6
90%         12.8             18       Variance       5.842037
95%        13.58           18.4       Skewness      -.0087677
99%         14.9          18.79       Kurtosis       2.543654
-------------------------------------------------------------
```

Table A 71: Dissolved oxygen at downstream stations

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%          4.4           1.41
 5%          5.9           1.47
10%         6.61           1.57       Obs               10519
25%          7.8           1.64       Sum of Wgt.       10519

50%          9.5                      Mean           9.605286
                        Largest       Std. Dev.       2.39415
75%        11.48          18.29
90%         12.8           18.4       Variance       5.731955
95%        13.44          18.79       Skewness       .0461606
99%         14.8             19       Kurtosis       2.594189
```

Table A 72: Dissolved oxygen at downstream and upstream stations, by season

```
-> season = winter

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
downstrea~04 |      2492    12.03269    1.487678          4      18.79
upstreamw~04 |      2389     12.0986    1.489429        4.2      18.79

-------------------------------------------------------------------------------
-> season = spring

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
downstrea~04 |      2743    8.880762     1.68735       2.82         19
upstreamw~04 |      2545    8.910697    1.745741        2.6       17.6

-------------------------------------------------------------------------------
-> season = summer

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
downstrea~04 |      2628    7.442565    1.607551       1.41       16.3
upstreamw~04 |      2491    7.623934    1.687045       1.24       15.4

-------------------------------------------------------------------------------
-> season = fall

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
downstrea~04 |      2576    10.12766    1.968936       3.21      18.29
upstreamw~04 |      2410    10.30207    2.035037        2.9         18
```

Table A 73: BOD5 concentration by 83 plants

```
        -------------------------------------------------------------
              Percentiles      Smallest
         1%        .8914             .1
         5%          1.4             .1
        10%            2             .1       Obs               10015
        25%            3            .17       Sum of Wgt.       10015

        50%          5.7                      Mean           8.501901
                                Largest       Std. Dev.      8.759269
        75%         10.9            107
        90%           19         115.75       Variance       76.72479
        95%           24         126.35       Skewness       3.776078
        99%        40.12          178.5       Kurtosis       34.79842
-----------------------------------------------------------------------------------------
```

Table A 74: Change in downstream from upstream DO

```
-------------------------------------------------------------
      Percentiles      Smallest
 1%        -4.71            -10
 5%         -2.7           -9.1
10%         -1.8           -7.7       Obs                9216
25%    -.7999992           -7.3       Sum of Wgt.        9216

50%    -.0999994                      Mean          -.0827957
                        Largest       Std. Dev.      1.544009
75%     .6000004            7.8
90%          1.6              8       Variance       2.383964
95%          2.4            8.9       Skewness      -.0602331
99%          4.2            9.4       Kurtosis       6.399151
```

Table A 75: Sum BOD5 concentration, mg/L

```
        -------------------------------------------------------------
              Percentiles      Smallest
         1%           .9             .1
         5%          1.5             .1
        10%            2            .17       Obs                8202
        25%          3.6             .2       Sum of Wgt.        8202

        50%          6.9                      Mean           10.07198
                                Largest       Std. Dev.      10.43687
        75%           13            107
        90%           22         115.75       Variance       108.9282
        95%         27.9         126.35       Skewness       3.636794
        99%     51.28333          199.6       Kurtosis       31.15128
-----------------------------------------------------------------------------------------
```

Table A 76: Sum BOD5 quantity, lbs/day

```
        -------------------------------------------------------------
              Percentiles      Smallest
         1%        1.695          .0616
         5%       5.4111           .065
        10%        10.15           .195       Obs                9428
        25%        28.45             .2       Sum of Wgt.        9428

        50%        98.78                      Mean           547.4516
                                Largest       Std. Dev.      2126.447
        75%     335.9333          32491
        90%         1040          34802       Variance        4521777
        95%         1846          39914       Skewness       9.794411
        99%        10680          40522       Kurtosis       119.0947
```

Table A 77: Yearly count of change in downstream water quality, by monitor

```
       Count|      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         27        0.29        0.29
          2 |         52        0.57        0.86
          3 |        108        1.18        2.04
          4 |        224        2.45        4.49
          5 |        220        2.40        6.89
          6 |        270        2.95        9.84
          7 |        147        1.61       11.45
          8 |        160        1.75       13.20
          9 |        603        6.59       19.79
         10 |        620        6.77       26.56
         11 |      1,958       21.39       47.95
         12 |      4,764       52.05      100.00
------------+-----------------------------------
      Total |      9,153      100.00
```

Table A 78: Yearly count of BOD5 concentration, by first/only plant

```
   Count    |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          8        0.10        0.10
          2 |         12        0.15        0.24
          3 |         63        0.76        1.01
          4 |         60        0.73        1.73
          5 |         90        1.09        2.83
          6 |         84        1.02        3.85
          7 |         70        0.85        4.69
          8 |        176        2.14        6.83
          9 |        306        3.71       10.54
         10 |        400        4.85       15.39
         11 |      1,166       14.15       29.54
         12 |      5,808       70.46      100.00
------------+-----------------------------------
      Total |      8,243      100.00
```

Table A 79: Yearly count of BOD5 concentration, by second plant

```
|      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          2        0.14        0.14
          3 |         12        0.81        0.95
          4 |          4        0.27        1.22
          5 |          5        0.34        1.56
          7 |         14        0.95        2.51
          8 |         56        3.79        6.30
          9 |         54        3.66        9.95
         10 |         80        5.42       15.37
         11 |        242       16.38       31.75
         12 |      1,008       68.25      100.00
------------+-----------------------------------
      Total |      1,477      100.00
```

Table A 80: Yearly count of BOD5 concentration, by third plant

```
|      Freq.     Percent        Cum.
------------+-----------------------------------
          3 |          3        0.89        0.89
          4 |          4        1.18        2.07
          6 |          6        1.78        3.85
          7 |         14        4.14        7.99
          8 |          8        2.37       10.36
          9 |          9        2.66       13.02
```

```
         10 |         10        2.96       15.98
         11 |         44       13.02       28.99
         12 |        240       71.01      100.00
------------+-----------------------------------
      Total |        338      100.00
```

Deriving the Streeter-Phelps equation for multiple polluters

If water quality data immediately upstream to the point of discharge of a plant DO_r was available, and a unique pair of upstream and downstream stations could be identified for each plant in the sample then the following modified S-P equation could be "estimated":

$$DO = DO_r e^{-k_R d} - \left(\frac{k_D}{k_R - k_D} \right) L_w (e^{-k_D d} - e^{-k_R d}) \tag{A4.1}$$

d is the distance between the point of outfall of a plant and the downstream monitoring station. Instead, DO data is available at monitoring stations at a certain distance further upstream to the point of outfall. And, multiple polluters between the same two upstream and downstream monitors is a common occurrence.

Consider the situation where there are 3 polluters in between the same set of upstream and downstream stations. The upstream monitoring station is located d_0 miles upstream from the first plant. And, the distance between the first and second plant is d_1 miles, second and third is d_2 miles, and the last plant is d_3 miles upstream to the downstream monitoring station. The concentration of dissolved oxygen in the river, prior to the first plant's discharge, DO_r^1 is DO_r^* times the factor that accounts for the reaeration process for d_0 miles (equation (A4.2)).

$$DO_r^1 = DO_r^* e^{-k_R d_0} \tag{A4.2}$$

Second, the dissolved oxygen concentration d_1 miles downstream to the first plant and prior to the discharge of the second plant, DO^1 is specified in equation (A4.7). DO^1 is the dissolved oxygen concentration in the river prior to the discharge of the second plant. Using equation (A4.1) we get:

$$DO^1 = DO_r^1 e^{-k_R d_1} - \left(\frac{k_D}{k_R - k_D} \right) L_w^1 (e^{-k_D d_1} - e^{-k_R d_1}) \tag{A4.3}$$

Substituting for DO_r^1 from equation (A4.2) we get:

$$DO^1 = DO_r^* e^{-k_R (d_0 + d_1)} - \left(\frac{k_D}{k_R - k_D} \right) L_w^1 (e^{-k_D d_1} - e^{-k_R d_1}) \tag{A4.4}$$

Again, using equation (A4.1), the dissolved oxygen concentration d_2 miles downstream of the second plant and, prior to the effluent discharge point of the third plant DO^2 is expressed in terms of DO^1, and L_w^2 the concentration of BOD in the second plant's effluent discharge.

$$DO^2 = DO^1 e^{-k_R d_2} - \left(\frac{k_D}{k_R - k_D} \right) L_w^2 (e^{-k_D d_2} - e^{-k_R d_2}) \tag{A4.5}$$

Substituting for DO^1 from equation (A4.4) we get:

$$\begin{aligned} DO^2 &= DO_r^* e^{-k_R(d_0+d_1+d_2)} - \left(\frac{k_D}{k_R - k_D} \right) e^{-k_R d_2} L_w^1 (e^{-k_D d_1} - e^{-k_R d_1}) \\ &\quad - \left(\frac{k_D}{k_R - k_D} \right) L_w^2 (e^{-k_D d_2} - e^{-k_R d_2}) \end{aligned} \tag{A4.6}$$

Similarly, DO^3 the ambient water quality d_3 miles downstream from the third plant, is expressed in terms of DO^2, the ambient water quality prior to the discharge of the third plant and L_w^3 the concentration of BOD in the third plant's effluent discharge.
Using equation (A4.1):

$$DO^3 = DO^2 e^{-k_R d_3} - \left(\frac{k_D}{k_R - k_D} \right) L_w^3 (e^{-k_D d_3} - e^{-k_R d_3}) \tag{A4.7}$$

Substituting for each of these ambient water quality expressions we get equation (A4.8), which captures the effect of upstream ambient water quality and effluent discharges on downstream water quality at a certain distance downstream (given stream velocity).

$$\begin{aligned} DO^3 &= DO_r^* e^{-k_R(d_0+d_1+d_2+d_3)} - \left(\frac{k_D}{k_R - k_D} \right) e^{-k_R(d_2+d_3)} L_w^1 (e^{-k_D d_1} - e^{-k_R d_1}) \\ &\quad - \left(\frac{k_D}{k_R - k_D} \right) e^{-k_R d_3} L_w^2 (e^{-k_D d_2} - e^{-k_R d_2}) \\ &\quad - \left(\frac{k_D}{k_R - k_D} \right) L_w^3 (e^{-k_D d_3} - e^{-k_R d_3}) \end{aligned} \tag{A4.8}$$

DO^3 = concentration of DO in the river at the monitoring location downstream to all the 3 point source polluters

DO_r^* = concentration of DO in the river, $d_0 + d_1 + d_2 + d_3$ miles upstream to the downstream monitoring location

L_w^1 = concentration of BOD5 in the effluent discharge reported by plant 1, $d_1 + d_2 + d_3$ miles upstream to the downstream monitoring location

L_w^2 = concentration of BOD5 in the effluent discharge reported by plant 2, $d_2 + d_3$ miles upstream to the downstream monitoring location

L_w^3 = concentration of BOD5 in the effluent discharge reported by plant 3, d_3 miles upstream to the downstream monitoring location

Table A 81: Non-linear estimation of downstream water quality

	Full Sample	Downstream distance of plants >2 and <26 miles
# of monitors	67	43
k_R	-0.0000554	0.0001158
	(-0.26)	(0.49)
k_D	0.0007351	0.0010108
	(3.36)**	(2.51)*
% correct predictions	86.52	85.87
Observations	6322	4209
R-squared	0.7274	0.7093
Prob > F	0.0000	0.0000

References

"Technology-Based Emission and Effluent Standards and the Achievement of Ambient Environmental Objectives", Editorial: *The Yale Law Journal* 1982 vol. 91(4) pp 792-813. 232.

9 VAC 25-720, "Water Quality Management Planning Regulation (Sections 50, 80, 90 and 100)", State Water Control Board, available at: http://www.townhall.state.va.us/L/GetFile.cfm?File=E:/townhall/docroot/103/2208/3866/Text_DEQ_3866_v1.pdf.

Akobundu, Eberechukwu (2004), "Two Essays on Resource Economics: A Study of the Statistical Evidence for Global Warming and An Analysis of Overcompliance with Effluent Standards Among Wastewater Treatment Plants", Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Bandyopadhyay, Sushenjit (2002), "Overcompliance in Point Source Water Pollution in the U.S.", Thesis (Ph. D.)--University of Maryland, College Park, 2002.

Bandyopadhyay, Sushenjit and John K. Horowitz (2006), "Do Plants Overcomply with Water Pollution Regulations? The Role of Discharge Variability", *Topics in Economic Analysis & Policy* vol. 6(1) Article 4, Available at: http://www.bepress.com/bejeap/topics/vol6/iss1/art4.

Bodo, Byron A. (1992), "Statistical Analyses of Regional Surface Water Quality in Southeastern Ontario", *Environmental Monitoring and Assessment* vol. 23 pp 165-187.
Boyd, James. (2000), "The New Face of the Clean Water Act: A Critical Review of the EPA's Proposed TMDL Rules", Resources For the Future Discussion Paper 00-12, March.

Brännlund, R., and K.G. Löfgren (1996), "Emission Standard and Stochastic Waste Load", *Land Economics* vol. 72(2) pp 218-230.

Brooks, K.N., P.F. Ffolliott, H.M. Gregersen & L.F. DeBano (1997), "Hydrology and the Management of Watersheds", 2nd edition, Iowa State University Press: Ames, Iowa.

COMAR, "Water Quality Criteria Specific to Designated Uses", Code of Maryland Regulations, COMAR 26.08.02.03-3 (Subtitle 08 WATER POLLUTION: 02 Water Quality), available at: http://www.dsd.state.md.us/comar/26/26.08.02.03%2D3.htm.

Decker, Christopher S. (2003), "Corporate Environmentalism and Environmental Statutory Permitting", *Journal of Law and Economics* vol. 46(1) pp 103-29.

DeShazo, J.R., and Andres Lerner. (2004), "The Consequences of Devolution for Standard Setting: An Empirical Analysis of the Clean Water Act", Mimeo. UCLA.

DeWald, T., R. Horn, A. Montalbano, L. Manning, P. Taylor, and R. Greenspun. (1985), "STORET Reach Retrieval Documentation", U.S. Environmental Protection Agency, Washington, DC.

Dion, C., Lanoie, P., and B. Laplante (1998), "Monitoring of pollution regulation: Do local conditions matter?", *Journal of Regulatory Economics* vol. 13 pp 5-18.

Doonan, J., Lanoie, P., and B. Laplante (2002), "Environmental Performance of Canadian Pulp and Paper Plants: Why Some do Well and Others do not?", CIRANO. Scientific series, website: http://www.cirano.qc.ca/pdf/publication/2002s-24.pdf.

Earnhart, Dietrich (2004a), "Panel Data Analysis of Regulatory Factors Shaping Environmental Performance", *The Review of Economics and Statistics* vol. 86(1) (Feb. 2004) pp 391-401.

Earnhart, Dietrich (2004b), "Regulatory factors shaping environmental performance at publicly-owned treatment plants", *Journal of Environmental Economics and Management* vol. 48(1) pp 655-681.

Earnhart, Dietrich (2004c), "The Effects of Community Characteristics on Polluter Compliance Levels", *Land Economics* vol. 80(3) (Aug. 2004) pp 408-432.

Earnhart, Dietrich (2007), "Effects of permitted effluent limits on environmental compliance levels", *Ecological Economics* vol. 61(1) pp 178-193.

Ebersole, E., M. Lane, M. Olson, E. Perry, and W. Romano (2002), "Assumptions and Procedures for Calculating Water Quality Status and Trends in Tidal Waters of the Chesapeake Bay and its Tributaries: A cumulative history", Report prepared for Chesapeake Bay Program, Tidal Monitoring and Analysis Workgroup, available at: (http://www.dnr.state.md.us/bay/tribstrat/stat_trend_hist.pdf).

El-Shaarawi, A.H., S.R. Esterby and K.W. Kuntz (1983), "A Statistical Evaluation of Trends on the Water Quality of the Niagara River", *Journal of Great Lakes Research* vol. 2 pp 234-240.

Esterby, Sylvia R. (1996), "Review of Methods for the Detection and Estimation of Trends with Emphasis on Water Quality Applications", *Hydrological Processes* vol. 10 pp 127-149.

Forester, Deborah Lynn (2000), "Water Quality in the Credit River: 1964 to 1998", M.A. Department of Geography/Institute for Environmental Studies, University of Toronto.

Gilliom, R.J. and D. R. Helsel (1986), "Estimation of Distributional Parameters for Censored Trace Level Water Quality Data 1, Estimation Techniques", *Water Resources Research* vol. 22(2) pp 135-146.

Gray, N.F. (1999), *Water Technology: An Introduction for Environmental Scientists and Engineers*, London: Arnold Publishers.

Gray, Wayne B. and Ronald J. Shadbegian (2004), "'Optimal' pollution abatement—whose benefits matter, and how much?", *Journal of Environmental Economics and Management* vol. 47 pp 510-534.

Hammer, M. J. and M. J. Hammer, Jr. (1996), *Water and Wastewater Technology*, New Jersey: Prentice Hall 3rd edition.

Helland, Eric. (1998), "The Enforcement of Pollution Control Laws: Inspections, Violations, and Self-Reporting", *Review of Economics and Statistics* 80 (1) pp 141-53.

Hirsch, Robert M. and James R. Slack (1984), "A Nonparametric Trend Test for Seasonal Data With Serial Dependence", *Water Resources Research* vol. 20(6) pp 727-732.

Hirsch, Robert M., James R. Slack and Richard A. Smith (1982), "Techniques of Trend Analysis for Monthly Water Quality Data", *Water Resources Research* vol. 18(1) pp 107-121.

Houtsma, John (2003), "Municipal Wastewater Treatment: A Comparison of the Performance of Public Sector and Contract Operated Plants", European Regional Science Association (ERSA) conference paper # ersa03p381.

Kagan, R.A., Gunningham, N., and D. Thornton (2003), "Explaining Corporate Environmental Performance: How Does Regulation Matter?", *Law and Society Review* vol. 37(1) pp 51-89.

Keplinger, Keith (2003), "The Economics of Total Maximum Daily Loads", *Natural Resources Journal* vol. 43 (Fall) pp 1057-1091.

Laplante, B. and Paul Rilstone (1996), "Environmental Inspections and Emissions of the Pulp and Paper Industry in Quebec", *Journal of Environmental Economics and Management* vol. 31 pp 19-36.

Liebman, J. C. and Lynn, W. R. (1966), "The optimal allocation of stream dissolved oxygen", *Water Resources Research* vol. 2, pp 581-591.

Luken, R. A., F. R. Johnson, and V. Kibler (1992), "Benefits and Costs of Pulp and Paper Effluent Controls Under the Clean Water Act", *Water Resources Research* vol. 28(3) pp 665–674.

Magat, Wesley A. and W. Kip Viscusi (1990), "Effectiveness of the EPA's Regulatory Enforcement: The Case of Industrial Effluent Standards", *Journal of Law and Economics* vol. 33(2) pp 331-360.

McClelland, J., and J.K. Horowitz (1999), “The Cost of Water Pollution Regulation in the Pulp and Paper Industry”, *Land Economics* vol. 75(2) pp 220-

McConnell, Virginia D. and Gregory E. Schwarz (1992), “The Supply and Demand for Pollution Control: Evidence from Wastewater Treatment”, *Journal of Environmental Economics and Management* vol. 23 pp 54-77.

MDDNR (2008), “Long-Term Health of Chesapeake Bay Tributaries and Maryland's Streams”, Maryland Department of Natural Resources, available at: http://mddnr.chesapeakebay.net/eyesonthebay/tributary_health.cfm.

MDE (1999), “TMDL of Biochemical Oxygen Demand (BOD) for the Western Branch of the Patuxent River, Prince George’s County, MD”, Maryland Department of the Environment, EPA Submittal: December 3, 1999 EPA Approval: June 6, 2000, available at: http://textonly.mde.state.md.us/assets/document/tmdl/westbranch/wb_techmem.pdf.

MDE (2000), “Total Maximum Daily Loads of Biochemical Oxygen Demand (BOD) for Conococheague Creek”, Maryland Department of the Environment, Acceptance Date: February 2, 2001 Completion Date: December 26, 2000, available at: http://www.mde.state.md.us/assets/document/ConocoTech_Memo%20Final.pdf.

MDE (2002), “Total Maximum Daily Loads of Biochemical Oxygen Demand (BOD), Nitrogen and Phosphorus for Town Creek into which the Town of Oxford Wastewater Treatment Plant Discharges Talbot County, Maryland”, Maryland Department of the Environment, Approval: October 16, 2003 Document version: December 26, 2002, available at: http://www.mde.state.md.us/assets/document/Town%20Creek%20Techmemo_final.pdf.

MDEQ (2008), “PERMIT RATIONALE FOR REISSUANCE: Clarksdale POTW Coahoma County Clarksdale, Mississippi Water NPDES No. MS0020311 March 24, 2008”, Mississippi Department of Environmental Quality, available at: http://opc.deq.state.ms.us/get_doc.aspx?dt=eparat&id=467122.

MODNR (1999), “Total Maximum Daily Loads (TMDLs) for Rock Creek, Jefferson County, Missouri”, Missouri Department of Natural Resources Water Pollution Control Program Approved by the EPA December 1, 1999 Completed July 21, 1999, available at: http://www.dnr.mo.gov/env/wpp/tmdl/1714-rock-ck-tmdl.pdf.

MODNR (2005), “Total Maximum Daily Loads (TMDLs) For West Fork Sni-a-Bar Creek Jackson County, Missouri”, Completed: November 21, 2005 Approved: January 6, 2006, available at: http://www.dnr.mo.gov/env/wpp/tmdl/0400-west-fork-sni-a-bar-tmdl.pdf.

Ntelekos, Alexandros A. (2005), “Modeling Dissolved Oxygen Concentrations in the Delaware River Under Drought Conditions.” Processed, Final Project, Princeton.

Prais, S.J. and C.B. Winsten. (1954), “Trend Estimators and Serial Correlation.” Cowles Commission Discussion Paper No 383, Chicago.

SCDHEC (1998), “Index: State of South Carolina Administrative Record; TMDL Submittal for Sanders Branch/Coosawhatchie River”, Bureau of Water Pollution Control, South Carolina Department of Health and Environmental Control, EPA Approval: August 17, 1998, Submittal: January 21, 1998, available at: www.scdhec.net/environment/water/tmdl/docs/tmdlsand.pdf.

Schwarz, Gregory E. and Virginia D. McConnell (1993), “Local Choice and Wastewater Treatment Plant Performance”, *Water Resources Research* vol. 29(6) pp 1589-1600.

Sigman, Hilary (2002), “International Spillovers and Water Quality in Rivers: Do Countries Free Ride?”, *The American Economic Review* vol. 92(4) pp 1152-1159.

Sigman, Hilary (2004), “Does Trade Promote Environmental Coordination?: Pollution in International Rivers”, *Contributions to Economic Analysis & Policy* vol. 3(2) Article 2.

Sigman, Hilary (2005), “Transboundary spillovers and decentralization of environmental policies”, *Journal of Environmental Economics and Management* vol. 50 pp 82-101.

StataCrop (2007), *Stata Longitudinal/Panel-Data Reference Manual: Release 10,* College Station, TX: StataCorp LP.

Stoddard, A., Harcum, J., Simpson, J., Pagenkopf, J. R., and R. K. Bastian (2002), *Municipal Wastewater Treatment: Evaluating Improvements in National Water Quality*, New York: John Wiley & Sons, Inc.

Streeter, H.W. and E.B. Phelps (1925), “A Study of the Pollution and Natural Purification of the Ohio River.” *Public Health Bulletin* 146. Washington D.C.: U.S. Public Health Service.

Summers, J. Kevin, Paul F. Kazyak, and Stephen B. Weisberg (1991), “A Water Quality Model for a River Receiving Paper Mill Effluents and Conventional Sewage”, *Ecological Modelling* vol. 58 pp 25-54.

Thomann, Robert V., and John A. Mueller. 1987. *Principles of Surface Water Quality Modeling and Control*. New York: Harper and Row.

USEPA (1991), “Technical Support Document for Water Quality-based Toxics Control”, EPA 505-290-001. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA (2000a), “A Benefits Assessment of Water Pollution Control Programs Since 1972: Part 1, The Benefits of Point Source Controls for Conventional Pollutants in Rivers and Streams", Final Report U.S. Environmental Protection Agency, Office of Water, Office of Policy, Economics, and Innovation, Washington, DC, January.

USEPA (2000b), "Progress in Water Quality: An Evaluation of the National Investment in Municipal Wastewater Treatment", U.S. Environmental Protection Agency, Office of Wastewater Management, Washington, DC, June.

USEPA (2000c), "A Retrospective Assessment of the Costs of the Clean Water Act: 1972 to 1997", Final Report U.S. Environmental Protection Agency, Office of Water, Office of Policy, Economics, and Innovation, Washington, DC, October.

USEPA (2001a), "Permit Compliance System Data Element Dictionary", PCS-DD01-1.00. USEPA- PCS User Support, Washington, DC, February 23, 2001.

USEPA (2001b), "Better Assessment Science Integrating point and Nonpoint Sources", EPA-823-B-01-001. U.S. Environmental Protection Agency, Office of Water, Washington, DC, June, available at: http://www.epa.gov/waterscience/basins/.

USEPA (2002), "Decision Rationale: Total Maximum Daily Loads of Carbonaceous and Nitrogenous Biochemical Oxygen Demand to Antietam Creek, Washington County, Maryland", Concurrence September 16, 2002 Water Protection Division, U.S. Environmental Protection Agency, Washington, DC, available at: http://www.epa.gov/reg3wapd/tmdl/MD_TMDLs/MercuryTMDLs/AntietamCreek/antietamckDR.pdf

USEPA. (1991), "Technical Support Document for Water Quality-based Toxics Control", EPA 505-290-001. U.S. Environmental Protection Agency, Office of Water, Washington, DC, March.

USEPA. (1996), "NPDES Permit Writers' Manual", EPA 833-B-96-003 U.S. Environmental Protection Agency, Office of Water, Washington, DC, December.

USGS (2005), "USGS surface and water quality models information clearinghouse", USGS, Washington, DC, available at: http://smig.usgs.gov.

VADEQ (2004), "Twenty-year Trend Analyses of Water Quality Variables – 1985-2004", Final Report: 305(b)/303(d) Water Quality Assessment Integrated Report, Virginia Department of Environmental Quality, available at: http://www.deq.Virginia.gov/wqa/ir2004.html.

Vellidis, G., P. Barnes, D.D. Bosch and A.M. Cathey (2006), "Mathematical Simulation Tools for Developing Dissolved Oxygen TMDLs", *Transactions of the ASABE* vol. 49(4) pp 1003-1022.

Zipper, C.E., G. Holtzman, P. Darken, P. Thomas, J. Gildea, L. Shabman, and T. Younos (1998), "Long-Term Water Quality Trends in Virginia's Waterways", Final Report: Virginia Water Resources Research Center, June.