Número 621

Impact of Upstream Plant Level Pollution on Downstream Water Quality

EVIDENCE FROM THE CLEAN WATER ACT

LOPAMUDRA CHAKRABORTI

AGOSTO 2019

CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONONÓMICAS

CIDE

Advertencia

Los Documentos de Trabajo del CIDE son una herramienta para fomentar la discusión entre las comunidades académicas. A partir de la difusión, en este formato, de los avances de investigación se busca que los autores puedan recibir comentarios y retroalimentación de sus pares nacionales e internacionales en un estado aún temprano de la investigación.

De acuerdo con esta práctica internacional congruente con el trabajo académico contemporáneo, muchos de estos documentos buscan convertirse posteriormente en una publicación formal, como libro, capítulo de libro o artículo en revista especializada.

Acknowledgements

I am grateful for the generous guidance provided by Ted McConnell, Jay Shimshack, Erik Lichtenberg, John Horowitz and Maureen Cropper at the inception of this research.

Abstract

This is the first study to assess whether pollutant inputs of major point sources have a negative impact on downstream water quality, net of upstream pollution levels and controlling for location specific factors. We utilize a panel data on monthly biochemical oxygen demand (BOD) concentration for a sample of 87 municipal and industrial plants located in the states of Maryland, Pennsylvania and Virginia, for the period 1990-2003. Water quality is measured by monthly dissolved oxygen (DO) from 67 locations within 25 miles downstream. We find that upon increase in aggregate BOD (by one or more plant) downstream DO net of ambient levels before their effluent outfalls declines by 0.001 mg/L. Despite the small magnitude (due to natural attenuation), the results are robust to distance traveled by pollutant and seasonal considerations of high temperature or low stream flow. From our results we infer that self-reported pollution does not exhibit underreporting biases.

Keywords: U.S. Clean Water Act, Ambient Water Quality Model, Self-Reported Pollution, Total Maximum Daily Loads, Over-compliance

Resumen

Este es el primer estudio que evalúa si los aportes de contaminantes de las principales fuentes puntuales tienen un impacto negativo en la calidad del agua río abajo, tomando en cuenta los niveles de contaminación río arriba y discriminando los factores específicos de la ubicación. Utilizamos un panel de datos sobre la concentración mensual de demanda bioquímica de oxígeno (DBO) para una muestra de 87 plantas municipales e industriales ubicadas en los estados de Maryland, Pennsylvania y Virginia, para el período 1990-2003. La calidad del agua se mide mediante datos de oxígeno disuelto (OD) mensual obtenido de 67 ubicaciones a lo largo de 25 millas río abajo. Encontramos que al aumentar la DBO agregada en 1% (por una o más plantas) río abajo el OD disminuye en 0.001 mg / L. A pesar de la pequeña magnitud (debido a la atenuación natural), los resultados son robustos en relación a la distancia recorrida por contaminantes y las consideraciones estacionales de alta temperatura o bajo flujo de agua. Nuestros resultados infieren que la contaminación auto-reportada no exhibe sesgos de subregistro.

Palabras claves: EEUU Clean Water Act, modelo de calidad del agua ambiental, contaminación auto-reportada, cargas diarias máximas totales, sobre cumplimiento

Introducción

Ambient water quality has not been studied extensively in the water pollution regulation literature primarily on account of unavailability of good time series data for water bodies across the U.S. McConnell and Schwartz (1992) 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 pollutants namely ambient biochemical oxygen demand (Sigman, 2002, 2004), dissolved oxygen, fecal coliform, total suspended solids, phosphorus and nitrogen (Sigman, 2005) have been modeled as measures of water quality. However, none of the studies that model ambient water quality estimate the impact of upstream pollutant inputs on in-stream water quality. Instead, proxies intended to capture the impact of pollutant inputs on water quality such as stream flow and temperature have been used (Sigman, 2005).

More recently, water quality has taken the spotlight based on concerns of non-point source pollution resulting in no significant improvements in ambient water quality trends after the initial improvements. Concurrently, various policy studies and academic reports have documented the high costs of the Clean Water Act (USEPA, 2000). Studies on lakes and estuaries have reported dismal trends (Smith and Wolloh,

2012). Until recently, research on ambient water quality has been limited mostly due to concerns about statistical representativeness and comparability across states of long-term water quality measurements (GAO, 2000). Grant and Langpap (2019) consider the measure of Dissolved Oxygen Deficit (DOD) over the period 1996–2008. Keiser and Shapiro's (2019), comprehensive ambient water quality database updated till 2001 finds positive trends in ambient water quality in rivers and streams and highlight unaccounted for benefits components (Keiser, Kling and Shapiro, 2019). In this paper we study the impact of upstream pollutant discharges on downstream water quality after controlling for ambient pollution, i.e. before the points of effluent outfall, and location specific factors. Our results have implications for lack of evidence on underreporting biases for self-reported pollution data which is the backbone of all CWA monitoring and enforcement.[1] Our result is contrary to preliminary evidence on under-reporting by sewage treatment plants across the U.S. found in Chakraborti and Shimshack (2012).

We utilize a sample of 87 major polluters in the three neighboring states of Maryland, Pennsylvania and Virginia, for the period 1990 to 2003. These plants are regulated by the NPDES permits (National Pollutant Discharge Elimination System) and are mandated by law to report monthly average pollutant discharges. We gather a comprehensive water quality database from the EPA's central STORET (Storage and Data Retrieval System) databases and state FOIA (Freedom of Information Act) requests. Our data might be considered slightly dated. However, both EPA's Permit Compliance System (PCS) the source of self-reported pollution data and ambient water quality coverages declined during the post 2000 period. Decentralizing of EPA's various databases and states transitioning to new reporting protocols and data platforms meant that historical databases failed to remain centralized. As a researcher looking into ambient water quality, long time period is a requirement due to the slow evolution of ambient water quality. Hence, we consider the most complete coverage of pollutant inputs and ambient water quality. Besides, the implications for self-reporting remain

[1] Self-reported pollution forms the basis of monitoring and enforcement for toxic as well as conventional pollutants like biochemical oxygen demand (BOD) and total suspended solids (TSS).

valid and can be applied to more recent decades, given no significant change in policy since the 2000s.

We estimate panel data models that control for all time-invariant location specific effects such as biophysical aspects of the stream segment. We consider a more general fixed effects model that controls for time varying changes at the state level such as economic or climate shocks such as drought or flood or extreme weather events, land use changes, changes in non-point source trends, changes in environmental preferences of government or citizens. We rely on distributional considerations of previous water quality and plant level pollution studies to estimate linear relationships.

We find that pollutant discharges from major point sources have a significantly negative impact on downstream water quality after controlling for upstream ambient water quality. Our results show that a one percent increase in upstream BOD pollutant inputs leads to a decline in downstream net of upstream dissolved oxygen levels (DO) by 0.001 mg/L. As expected, the magnitude of the coefficients is small (due to assimilative capacity of water bodies) but they are consistently negative and statistically significant in our various robustness checks based on distance traveled by pollutants or seasonal considerations.

BACKGROUND AND LITERATURE

The Ambient Water Quality Model

Evolution of ambient dissolved oxygen levels is determined by upstream pollution inputs of BOD, upstream ambient concentrations, flow, temperature and assimilative capacity of the stream. In this section, we present the Streeter-Phelps (Dissolved) Oxygen Sag Curve to gain an understanding of the physical relationship between observed ambient dissolved oxygen and the BOD effluent discharges of "major" point source polluters.

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 and seasonal cycles.[2] Seasonally, DO concentrations are higher 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 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 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. 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:

$$\frac{dD}{dt} = k_D L_0 e^{-k_D t} - k_R D \qquad \text{.....(1)}$$

[2] 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.

Where $D = DO$ 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. Equation (1) is essentially a balance between DO consumption due to BOD inputs and stream reaeration. The first term captures the rate of deoxygenation i.e. consumption of DO, while the second term 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. BOD of the river/wastewater mixture (L_0), also known as the ultimate BOD is given by:

$$L_0 = \frac{Q_r L_r + Q_w L_w}{Q_r + Q_w} \qquad \text{.....(2)}$$

Where:

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

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

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

$Q_w =$ Flow of the wastewater discharge

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

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.[3] 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 (2)).[4] Consequently, regulators (federal and state EPAs) focus their permitting, monitoring and compliance activities towards "major" polluters where one of the criteria used is design flow greater than 1 million gallons per day. 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.

The solution to the differential equation (1) gives the ambient water quality at distance$d$ (downstream to wastewater discharges) where u is average water velocity:

$$D = \frac{k_D L_0}{k_R - k_D}(e^{-k_D d/u} - e^{-k_R d/u}) + D_0 e^{-k_R d/u} \quad \text{.....(3)}$$

This is the well-known Streeter-Phelps oxygen-sag curve formula (Streeter and Phelps, 1925), which was originally developed for use on the Ohio River, back in 1914. Substituting for initial dissolved oxygen deficit $D_0 = DO_S - DO_0$, and dissolved oxygen deficit $D = DO_S - DO$, the relationship in equation (3) can be expressed in terms of ambient dissolved oxygen levels instead of DO deficits:

$$DO = DO_S(1 - e^{-k_R t}) - \frac{k_D L_0}{k_R - k_D}(e^{-k_D t} - e^{-k_R t}) + DO_0 e^{-k_R t} \quad \text{.....(4)}$$

The initial dissolved oxygen (DO_0) is calculated using the same formula that was used for calculating BOD of the river/wastewater mixture at the point of outfall, L_0(equation (2)). Thus, ambient water qualityDOat a certain distance downstream (given stream velocity) depends on initial dissolved oxygen DO_0, concentration of BOD of the river/wastewater mixture at the outfall L_0, reaeration (k_R) and deoxygenation (k_D)

[3] 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 paper mill discharging its effluents to the Pigeon River in NC. The authors find that simulated BOD in the river is reduced by 20 ppm in the vicinity of the discharge.

[4] On the other hand, as stream flow approaches an infinitesimally big number ultimate BOD is determined solely by the upstream concentration of BOD,Lr (in equation (2)). Hence, it is suspected that non-point sources of pollution are contributing the most to ambient pollution during high stream flow seasons.

coefficients, and the saturation level of DO in the river waterDO_S. This is a steady-state model relating dissolved oxygen concentration in a free-flowing stream to BOD. Equation (4) predicts a negative relation between $L_0$and downstream ambient water quality measured by the concentration of dissolved oxygen,DO.

USES OF STREETER-PHELPS IN THE CWA

Regulators undertaking water quality evaluations every permit cycle have used the Streeter-Phelps model of a free-flowing stream, and hence to determine whether water quality-based effluent limits for BOD are needed. For example, in Maryland and Virginia water quality-based limits for BOD were derived from non-TMDL waste load allocations in the 1990s. 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 is simulated.[5] If 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 implemented.

As the water pollution regulation moved towards the TMDL regime (in order to incorporate non-point source pollution) the Streeter-Phelps equation continue to be used as the underlying model.[6] 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.[7] The Streeter-Phelps 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

[5] For Maryland, 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.

[6] TMDLs utilize a steady-state model that is 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.

[7] Low dissolved oxygen can arise 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 waterbody. 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.

sources and waste-load allocations from all point sources could be permitted so that the ambient water quality standard is met in future time periods.[8] Alternatively, it is used to determine how much pollutant inputs from all relevant sources must be reduced in order to maintain the water quality standard for a stream segment.

Using the basic concept of Streeter-Phelps many increasingly complex mathematical models have cropped up to accurately simulate DO dynamics in streams. "Most were developed to simulate parameters associated with [the NPDES] permits" (Vellidis *et al*. 2006, 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) models for developing DO TMDLs (USGS, 2005), while HSPF (Hydrological Simulation Program- Fortran) is a dynamic model. 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.

Several previous studies such as El-Shaarawi, Esterby and Kuntz (1983), Bodo (1992) and Esterby (1996) report that robust and multivariate regression analysis has been extensively used 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.[9] Therefore, a non-linear relationship could be estimated using least squares method for each station; given that the constants of Streeter-Phelps are location specific.

Drawing from the above discussion, a simplified linear relationship (across distinct monitoring locations) between ambient DO at a certain distance downstream and aggregate BOD pollution from all the relevant point sources is estimated. Water quality data immediately upstream or downstream to the point of discharge of a plant

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

[9] 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).

is not available. Instead, DO data is available at monitoring stations located at a certain distance upstream or downstream to the point of outfall. Hence, a unique pair of upstream and downstream stations could not be identified for each plant in the sample. Incorporating multiple polluters for the same pair of upstream and downstream monitors, either due to lack of water quality data or monitoring stations, we present our version of the empirical water quality model in section 4.

DATA

Matching Water Quality and Pollution Inputs

EPA's STORET databases (Legacy and Warehouse), Chesapeake Bay Program (CBP) and Virginia Department for Environmental Quality (VADEQ) are the primary sources for ambient water quality data. We consider the period from 1990 to 2003. Water Quality coverage and data varied a lot by state. More recent data beyond 2000 is scarcer for Pennsylvania and Virginia. But, at least one of the three states considered were among the most monitored for water quality based on its importance relative to the Chesapeake Bay (Maryland). Dissolved oxygen (DO) is the measure of water quality data that was collected. It is also the most monitored measure of water quality across the U.S. (Keiser and Shapiro 2019). Higher DO levels in mg/L means better water quality as it reflects the oxygen that is available for fish and aquatic life to survive.

The choice and location of downstream as well as upstream monitoring stations was primarily driven by availability of water quality data. We map the NPDES majors reporting monthly BOD5 during 1990 to 2003 to identify the relevant upstream and downstream monitoring stations for each polluter.[10] There are 97 major industrial and municipal plants that reported monthly BOD5 (either concentration or quantity loads). We assign monitors 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. For example, there are three pairs of upstream and downstream monitoring locations with one plant on a tributary whereas the

[10] 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 degree centigrade) over a five-day period. It has a detection limit of 1 mg/L.

others are 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. There is no monitoring data available from the tributaries.[11]

There are 76 water quality monitoring locations with its corresponding upstream water quality data. These monitoring stations were downstream to at least one major point source (either industrial or municipal plant). They also have good coverage of dissolved oxygen data from 1990 to 2003. Actual water quality records are averaged on a monthly basis upon availability of multiple records. 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 have one other major facility discharging into the same stream segment and hence they have the same pair of upstream and downstream stations. The other 12 plants have two other plants polluting into the same stream segment i.e. there are three plants discharging "in between" the same pair of upstream and downstream monitoring stations.

Summary Statistics

In the empirical model we describe why concentration measure of BOD has an unambiguously negative impact on dissolved oxygen that in turn is determined by the ultimate BOD in the river/wastewater mixture i.e. at the point of effluent outfall (equation (2)). On the other hand, increase in BOD load might lead to higher dissolved oxygen when effluent BOD concentration is higher than ambient or stream BOD concentration. Hence, our final sample of NPDES majors drop to 87 plants reporting BOD concentration in the EPA's Permit Compliance System (PCS) database. These 87 plants are matched to 67 distinct downstream (and upstream) monitoring locations. By states, the distribution of stations is 37 for Virginia, 18 for Maryland, and 12 for Pennsylvania. The distribution of the 87 major facilities is 47 plants in VA, 24 plants in

[11] Another plant could not be included since we could not identify its relevant upstream station. The plant discharged its effluents near the point of confluence of two tributaries.

Maryland, and 16 plants in Pennsylvania. Across states, more than 62% of the 87 plants are municipal plants (54) and the remaining 33 plants are industrial facilities.

Table 1. Summary Statistics of Water Quality and Pollution Inputs

	Obs.	Mean	Std. Dev.	Max.
FIXED FACTORS				
Distance to downstream station	87	11.11	10.77	48.00
Distance to upstream station	87	9.83	10.02	46.00
Total distance between stations	67	19.31	15.51	64.50
WATER QUALITY				
Downstream, DO	6,514	9.47	2.38	18.40
Upstream, DO	6,514	9.52	2.40	17.50
POLLUTION INPUTS				
Monthly BOD concentration	7,946	8.63	8.98	178.50
Aggregate BOD concentration	6,514	1 0.53	10.90	199.60
REGRESSION VARS.				
Downstream-Upstream, DO	6,514	-0.05	1.47	8.90
log Aggregate BOD	6,514	1.95	0.91	5.30

Notes: Distance in miles to downstream/upstream stations are for 87 majors with BOD5 concentration data. Total distance between stations are for 67 upstream-downstream monitoring segments. Water Quality is monthly average Dissolved Oxygen (DO) in mg/L for 67 unique downstream and upstream stations, for the period 1990-2003. Pollution Inputs are monthly average concentration of BOD5 in mg/L reported by 87 plants, over 1990-2003. Aggregate BOD is the sum of discharges by plants located in the same upstream-downstream segment. Regression Variables are dependent variable: downstream net of upstream monthly DO at 67 stream segments, and primary regressor is log aggregate BOD concentration of 87 major polluters.

Table 1 first presents summary statistics on fixed factors such as the average downstream, upstream distance of each plant to water quality monitoring stations for the 87 plants. And the total distance between upstream and downstream monitoring stations. All distance is measured in miles. We adhere to past studies e.g. Sigman (2005) that use 50 miles criterion based on the physical rates of attenuation for oxygen depletion. In robustness checks we apply more stringent criteria of total distance between upstream and downstream monitoring stations less than 40 miles, distance to downstream/upstream monitor less than 25 miles (Keiser and Shapiro 2019). Table 1 also presents summary statistics on the continuous variables of monthly average, downstream and upstream dissolved oxygen, measured in concentration units of mg/L, for the regression sample. It also presents the monthly average aggregate BOD concentration in mg/L and finally the dependent variable and the primary regressor for the empirical models.

Table 1 shows that average upstream and downstream distance for the current sample of plants are 10 and 11 miles respectively with 75% of the observations less than 14.0 and 16.5 miles. Total distance between upstream and downstream monitoring stations is on average 19.3 miles with 75% of the observations less than 28 miles. On average, both downstream and upstream dissolved oxygen levels at 9.5 mg/L are significantly higher than the ambient standards of 4-5 mg/L required to maintain aquatic life. In the current sample, correlation between contemporaneous upstream and downstream water quality is about 75%. Monthly BOD concentration for 87 plants is 8.6 mg/L. Compared to average concentration limits closer to 30 mg/L, these plants were significantly overcomplying with their permits (Bandyopadhyay and Horowitz, 2006). On average, monthly aggregate BOD concentration of 10.5 mg/L means that the significant overcompliance pattern is maintained even when pollution discharges are summed across multiple polluters. Last, we present the dependent variable, downstream net of upstream dissolved oxygen levels. On average, the mean difference in water quality is close to zero with standard deviation close to 1.5 meaning that like previous studies our measure of net of upstream, downstream dissolved oxygen roughly follows normal distribution. On the other hand, studies have cited BOD

following a log-normal distribution which can be seen upon log transforming the aggregate BOD measure with mean 1.95 and standard deviation close to 1.

EMPIRICAL STRATEGY

The central question that is the subject of this paper is whether pollution discharges from major point sources have a significant impact on downstream water quality. As mentioned in the introduction, the question though may appear trivial given that the point sources in the sample are major sources of pollution. However, we believe our study provides important evidence on assessing validity of self-reported pollution data. If self-reported pollution is widely underreported (as found in some Latin American countries, Caffera and Lagomarsino, 2014) then upstream pollution from major polluters might not have any impact on downstream water quality net of ambient pollution levels and other location or stream specific factors as captured by ambient water quality upstream to these polluters.

Model

Our main model is based on the theoretical foundations of the Streeter-Phelps dissolved oxygen sag curve as presented in section 2.2. Current water quality models have evolved considerably in complexity with the aid of advanced computing abilities. But the simple equation that describes downstream water quality as determined by upstream water quality, pollution inputs and temperature and stream flow conditions, remains valid. Equation (5) below presents the main model based on the Streeter-Phelps. The dependent variable is downstream dissolved oxygen net of upstream dissolved oxygen in month*t*. For each stream segment, we identify the relevant upstream and downstream monitoring locations with dissolved oxygen data. A stream segment is considered if there is a major point source discharging its effluents in that waterbody. Upstream water quality in the same month is endogenous as weather and biophysical conditions that determine water quality downstream to a point source's point of effluent outfall also determine water quality at a nearby upstream location. Hence, we include downstream net of upstream water quality as our dependent

variable. This strategy also controls for unobserved or hard to measure ambient conditions such as terrain, land use, elevation etc. that jointly determine water quality at nearby locations. [12]

Our primary regressor is pollution inputs from major point sources. The pollutant input that is relevant for dissolved oxygen is biochemical oxygen demand (BOD). Given that we could not match each point source to a distinct pair of downstream-upstream monitoring stations, our measure of pollution inputs is aggregate BOD discharged by one or more major point source in each stream segment. Later we check for sensitivity by considering only single polluter stream segments in our estimations. In this case, the measure of pollutant input is BOD discharges by a single plant for each stream segment.

In equation (5) below, the impact of pollutant inputs from point source dischargers on ambient water quality is examined. 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}). As mentioned above, we consider this dependent variable as upstream water quality is endogenously determined in same month $t$based on factors that also explain downstream quality (except of course, the pollutant discharges in between).

$$DO_{jt} - DO_{ut} = \alpha + \beta \sum_{i=1}^{3} BOD_{ijt} + \delta_j + \gamma_T + \theta_s + \varepsilon_{jt} \qquad \text{.....(5)}$$

Where:

$DO_{jt} =$ 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 and month t, upstream to the plant location i

[12] 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 close to each other.

$BOD_{ijt} =$ concentration 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 $T = 1,2.\ldots13$

$\theta_s =$ quarterly seasonal indicator for each month with s = 1,2,3,4

ε_{jt} = error term for monitor j in month t

The primary explanatory variable of interest is the sum of concentration 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}). Linear distance along the stream flow are recorded for the distance of each plant i to its downstream monitoring location j (d_{ij}). We expect pollutants from the plant that is located further upstream from station j has traveled longer, and hence are likely to have undergone more attenuation. In later robustness, we focus on plants within 25 miles downstream to verify if the estimated coefficients are larger in magnitude than the entire sample of 87 plants.

Among the controls, time invariant location specific effects that have not been captured by water quality at the upstream location are captured by station level dummy variables. [13] Specifically, it controls for long-term variations in non-point source pollution for the segment of the stream in between the upstream and downstream station. In the absence of data on non-point source pollution, regulators while conducting 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 annual changes in downstream water quality net of upstream condition across the three states from 1990 to 2003. Lastly, quarterly seasonal indicators are included mainly to facilitate robustness check on the estimates. Seasonal variations in temperature and rainfall are expected to affect downstream and upstream water quality in an identical manner, and hence season is

[13] For example, differences in land use across downstream locations since this is unlikely to change rapidly in a short period of time.

not anticipated to influence the change in downstream from upstream DO. However, monthly BOD discharged by the plants also exhibits seasonal variability. Hence, we include seasonal indicators in order to capture the effect of BOD net of seasonal effects.[14] Lastly, annual dummy variables are included to control for any possible annual trends over the 14 years of data.

For our empirical model, we consider the concentration measure of BOD5 in order to capture the impact of pollutant discharges on downstream water quality net of upstream ambient pollution, i.e. before point sources' effluent outfalls. Our discussion is based on equation (2), which calculates the ultimate BOD that can be observed in the wastewater- river mixture after a point source's effluent outfall and controlling for upstream ambient pollution. Consider a plant, which faces only concentration limits, and hence discharges high loads into the river during high flow seasons. However, the impact of high loads on ultimate BOD is not clear i.e. it might not lead to higher pollution in the river/wastewater-mixing zone. In general, it is expected that the effluent flow of a plant will be lower than the stream flow, especially during high flow conditions.[15] Consequently, Q_w in equation (2) increases but Q_r in equation (2) rises more during high flow seasons, meaning that the weight assigned to the effluent concentration (L_w in equation (2)) actually goes down. On the other hand, the weight assigned to upstream concentration (L_r in equation (2)) is higher, since Q_w/Q_r falls. Therefore, ceteris paribus, ultimate BOD falls under the condition that effluent concentration exceeds upstream concentration.

On the contrary, consider a plant that faces only quantity limits and hence can discharge a high concentration of BOD during low flow seasons. In terms of ultimate BOD, low flow conditions imply that both Q_w and Q_r are lower. However, the decline in the effluent flow rate Q_w is expected to be less than the reduction in the stream flow Q_r. In equation (2), the weight assigned to effluent concentration L_w is higher, while the

[14] 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. Plants can reduce discharges during higher temperatures that support better efficiency in the biological processes involving wastewater treatment technology.

[15] We illustrate using a simple numerical example: receiving stream 7Q10 is 30 cubic feet per second (cfs), while the design flow of the POTW in concern has a design flow of 5 mgd or 7.7 cfs. Actual stream flow during low flow conditions are higher than the 7Q10, while the actual recorded flows of wastewater will never exceed the volume that a plant is designed to accommodate.

weight assigned to upstream concentrationL_r is lower. Therefore, the impact of higher effluent concentration (under low flow conditions) results in unambiguously higher ultimate BOD. Hence, theory guides us in our choice of BOD5 concentration rather than quantity loads as we have clear expectations on the sign of the estimated coefficient on the log of aggregate BOD5 concentration.

Our argument for using concentration rather than load as the measure of pollutant inputs is supported by how policies are implemented. For instance, regulators assign TMDLs for stream segments that are "impaired" (i.e. not meeting minimum ambient standards for designated uses) in terms of limits on the concentration of BOD5 under critical 7Q10 drought-like stream flow conditions. Given that background pollutionL_r is zero under 7Q10 conditions, BOD in the river/wastewater mixture are solely captured by effluent concentration. The corresponding maximum allowable load is then 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).

Results

From the discussion in the previous section, one expects that the direction of impact of BOD concentration on ambient dissolved oxygen will be negative. BOD is from organic pollutants discharged by plants that are significant sources of oxygen demand and hence their impact on ambient dissolved oxygen levels is expected to be negative. Higher BOD concentration mean lower dissolved oxygen in the downstream locations; it is empirical evidence of this impact controlling for other location specific factors that is absent in the literature.

Table 2. Fixed Effects Estimations, Main Panel Results

Panel A. Monitoring location by month data				
DEP.VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**
log aggregate BOD	-0.078**	-0.100***	-0.071**	-0.091***
	(0.034)	(0.035)	(0.033)	(0.034)
Year FE	X	X		
Year X State FE			X	X
Station FE	X	X	X	X
Season FE	X	X	X	X
# stations	67	53	67	53
# observations	6,514	6,047	6,514	6,047
Panel B. Single polluter monitoring segments				
DEP.VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**
log aggregate BOD	-0.103**	-0.120**	-0.090*	-0.107**
	(0.047)	(0.048)	(0.047)	(0.047)
Year FE	X	X		
Year X State FE			X	X
Station FE	X	X	X	X
Season FE	X	X	X	X
# stations	51	39	51	39
# observations	4,531	4,173	4,531	4,173

Notes: Model 1 includes year fixed effects and Model 2 includes year interacted with state fixed effects. In Panel A, Full sample refers to monthly downstream net of upstream DO data from 67 stream segments with unique downstream-upstream stations. Balanced sample refers to 53 stream segments with at least 50% coverage i.e. 84 months of the total possible 168 months for each downstream-upstream pair of stations. In Panel B, Full sample refers to downstream net of upstream DO from 51 stream segments with single point source and Balanced sample refers to 39 locations with unique upstream-downstream monitors with single polluter and 50% water quality coverage. Clustered standard errors within stations in parenthesis: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Overall, we find that BOD concentration discharges from major polluters exert a negative impact on downstream net of upstream ambient DO levels. Table 2 presents the fixed effects estimation results for regression sample of 87 plants (Panel A) and sample of 51 single polluters' stream segments (Panel B). The balanced criterion in columns (2) and (4) is for ambient water quality measurements with at least 50% coverage i.e. 84 months out of the total possible 168 months. As seen in the number of stations reported, the water quality coverage is good in general as sample size falls from 67 to 53. The coefficients can be interpreted as a one percent increase in aggregate BOD concentration discharged by polluters results in 0.001 mg/L decline in net downstream DO (columns (1) – (4) of Table 2). As expected, the estimated impact is small but statistically significant and negative in sign. In Panel B, for the single polluter stream segments, the estimated coefficients are similar in magnitude. We infer that "free-riding" is not apparent as the coefficients for the 87 plants including (multiple polluter stream segments) are not much different from the single polluter stream segments. The coefficients in model (2) that control for time varying state level changes can be interpreted as a one percent increase in aggregate BOD concentration discharged by polluters results in 0.001 mg/L decline in net downstream DO (columns (3) and (4) of Table 2).

Robustness

Our first set of robustness tests is for alternative distance criterion. Our expectation is that since ambient water quality dilutes pollutant concentration the further downstream it travels, the coefficients will be larger upon stricter distance criteria i.e. smaller stream segments with upstream and downstream monitoring stations. Table 3 presents three panels of estimations: Panel A applies total stream segments i.e. distance between upstream and downstream monitoring locations less than 40 miles, Panel B with distance to downstream monitor for each plants less than 25 miles, and Panel C with distance to upstream monitor for each plant less than 25 miles.

Overall, results are robust to all three distance criteria with larger magnitudes for downstream monitoring stations less than 25 miles. This is expected as distance

pollutant inputs from major polluters travel downstream is more significant than either total distance or distance to upstream monitoring location as upstream water quality serves more as a control for the impact of pollutant inputs net of ambient concentrations. Number of stations in Panel A shows that most of our stream segments are less than 40 miles (less than 39 miles in the data). Sample size drops from 67 to 62 only. The magnitudes are close to Table 2 results specifically for the balanced sample criterion of 50% coverage of ambient water quality (columns (2) and (4) of Table 3). The coefficient can be interpreted as one percent increase in aggregate BOD concentration discharged by polluters results in 0.001 mg/L decline in net downstream DO (rounding to the nearest decimal). Panel B are larger than Panel A magnitudes with sample size dropping from 67 to 56. The coefficients in columns (2) and (4) can be interpreted as one percent increase in aggregate BOD concentration discharged by polluters results in 0.0011 mg/L decline in net downstream DO. Panel C applies the criterion of distance to upstream monitoring station less than 25 miles for each plant (mostly for consistency checks). This criterion is not as restrictive as Panel B as sample size falls from 67 to 60 for "full" sample and from 53 to 47 for "balanced" sample. The coefficients in columns (2) and (4) can be interpreted as one percent increase in aggregate BOD concentration discharged by polluters results in 0.001 mg/L decline in net downstream DO (rounding up to the nearest decimal).

Table 3. Robustness: Sensitivity to distance

Panel A. Total distance between upstream and downstream stations less than 40 miles				
DEP.VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**
log aggregate BOD	-0.067**	-0.081**	-0.064**	-0.075**
	(0.032)	(0.033)	(0.032)	(0.033)
Year FE	X	X		
Year X State FE			X	X
# stations	62	49	62	49
# observations	6,029	5,616	6,029	5,616

Panel B. Distance to downstream monitors less than 25 miles				
DEP.VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**
log aggregate BOD	-0.106***	-0.112***	-0.107***	-0.110***
	(0.032)	(0.034)	(0.030)	(0.032)
Year FE	X	X		
Year X State FE			X	X
# stations	56	45	56	45
# observations	5,478	5,145	5,478	5,145
Panel C. Distance to upstream monitors less than 25 miles				
DEP. VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**
log aggregate BOD	-0.070**	-0.083**	-0.065*	-0.076**
	(0.034)	(0.034)	(0.033)	(0.033)
Year FE	X	X		
Year X State FE			X	X
# stations	60	47	60	47
# observations	5,857	5,443	5,857	5,443

Notes: Clustered standard errors within stations in parenthesis: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Our second set of robustness tests applies sensitivity to different seasons. As seen in discussion on applications of Streeter-Phelps equations to TMDLs and water quality-based effluent limits, the critical conditions of high temperature and low stream flows are important due to minimal assimilative capacity of water bodies during this period or season. Some of these plants in the regressions, face seasonal limits during a year with higher permitted levels during low temperature high stream flow seasons and lower permitted levels during high temperature and low stream flow seasons. We expect larger negative impact of pollutant inputs during high temperature and low stream flow seasons in contrast to low temperature and high stream flow seasons.

Table 4 presents results by quarter. The coefficients in Table 4 for summer months are larger than winter months. Column (4) of Panel C (for summer months) can

be interpreted as one percent increase in aggregate BOD concentration discharged by polluters results in 0.0020 mg/L decline in net downstream DO (rounding up to the nearest decimal). In contrast, column (4) of Panel A (for winter months) can be interpreted as one percent increase in aggregate BOD concentration discharged by polluters results in 0.001 mg/L decline in net downstream DO (rounding up to the nearest decimal). The coefficients in spring (Panel B) versus fall (Panel D) can be compared to infer the importance of stream flow conditions. For Maryland, Pennsylvania and Virginia, stream flow is higher during fall than spring (other than snow melt) so coefficients for spring months exert a statistically significant negative impact on net downstream DO (Panel B) as opposed to coefficients for fall months that are not of the expected sign (or significance).

Table 4. Robustness: Panel Estimations by Season (Downstream monitors <25 miles)

Panel A. Winter Season (months=1,2,3)				
DEP.VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**
log aggregate BOD	-0.076	-0.100*	-0.068	-0.093*
	(0.051)	(0.050)	(0.048)	(0.049)
# stations	55	44	55	44
# observations	1,262	1,177	1,262	1,177
Panel B. Spring Season (months=4,5,6)				
DEP.VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**
log aggregate BOD	-0.112**	-0.123***	-0.106**	-0.117**
	(0.047)	(0.046)	(0.051)	(0.048)
# stations	55	44	55	44
# observations	1,382	1,308	1,382	1,308
Panel C. Summer Season (months=7,8,9)				
DEP. VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**

log aggregate BOD	-0.144**	-0.136*	-0.174**	-0.162**
	(0.070)	(0.072)	(0.069)	(0.071)
# stations	55	44	55	44
# observations	1,347	1,261	1,347	1,261
Panel D. Fall Season (months=10,11,12)				
DEP. VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**
log aggregate BOD	0.048	0.064	0.036	0.065
	(0.046)	(0.050)	(0.046)	(0.048)
# stations	55	44	55	44
# observations	1,348	1,260	1,348	1,260

Notes: Model 1 includes year fixed effects and Model 2 includes year interacted with state fixed effects. Clustered standard errors within stations in parenthesis: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Our last robustness test consists of a investigating heterogeneity based on a different seasonal classification. We divide each year into two 6-mont periods: November to April as low temperature and high stream flow conditions and May to October as the high temperature and low stream flow conditions. Table 5 presents results for each 6-month period of a year. Coefficients in Panel A are for low temperature and high stream flow conditions and Panel B are for high temperature and low stream flow conditions. We expect higher temperatures to have a larger magnitude of negative impact on net ambient water quality in contrast to lower temperature conditions. Column (4) of Model 2 in Panel B for higher temperature conditions can be interpreted as one percent increase in aggregate BOD concentration discharged by polluters results in 0.0013 mg/L decline in net downstream DO. Column (4) of Model 2 in Panel A for lower temperature conditions can be interpreted as one percent increase in aggregate BOD concentration discharged by polluters results in 0.0005 mg/L decline in net downstream DO (rounding up to the nearest decimal). The magnitude is about half the size for higher temperature months.

Table 5. Robustness: Panel Estimations by temperature
(Downstream monitors <25 miles**)**

Panel A. Low temperature, high stream flow (November-April)				
DEP.VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**
log aggregate BOD	-0.036	-0.050	-0.035	-0.048
	(0.034)	(0.035)	(0.032)	(0.033)
# stations	55	44	55	44
# observations	2,584	2,415	2,584	2,415
Panel B. High temperature, low stream flow (May-October)				
DEP.VAR: Net Downstream DO:	**Model 1**		**Model 2**	
Sample:	**Full**	**Balanced**	**Full**	**Balanced**
log aggregate BOD	-0.135**	-0.138**	-0.135**	-0.132**
	(0.053)	(0.055)	(0.051)	(0.052)
# stations	55	39	51	39
# observations	2,755	2,591	2,755	2,591

Notes: Model 1 includes year fixed effects and Model 2 includes year interacted with state fixed effects. Clustered standard errors within stations in parenthesis: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Conclusions

This is the first study to our knowledge to estimate an empirical ambient water quality model with the objective of assessing the impact of pollutant inputs from major point sources regulated under the CWA. Our question may appear obvious upfront as major polluters are classified by the CWA based on their design effluent flow i.e. volume of effluents discharged into waterbodies. However, the implications of these results for validating the self-reported pollution database can be useful for policy purposes. Despite a slightly dated period of 1990-2003, our conclusions for self-reported pollution data remains valid until today and might be applied to other regions or states.

We find that indeed upstream pollutant inputs of BOD lead to a decline in downstream net of upstream ambient DO levels. The magnitude of the estimated coefficients is small as expected from the natural attenuation of water bodies but negative and statistically significant in our main results as well as various robustness checks. We find that a one percent increase in upstream BOD leads to a decline in downstream DO net of upstream DO (prior to pollutant discharges) by 0.001 mg/L. Our results indicate that BOD concentration reported by polluters might not suffer from any reporting biases

References

Bandyopadhyay, S. & Horowitz, J. K. (2006). Do Plants Overcomply with Water Pollution Regulations? The Role of Discharge Variability. *Topics in Economic Analysis & Policy*, 6(1) Article 4, Available at: http://www.bepress.com/bejeap/topics/vol6/iss1/art4.

Bodo, B. A. (1992). Statistical Analyses of Regional Surface Water Quality in Southeastern Ontario. *Environmental Monitoring and Assessment*, 23(1-3): 165-187.

Caffera, M. & Lagomarsino, A. (2014). Under reporting of water pollution by industrial plants: empirical evidence. Latin American and Caribbean Environmental Economics Program *LACEEP Working Paper Series*, #WP66.

Chakraborti, L. & Shimshack, J.P. (2012). Assessing the accuracy of self-reported pollution Data. *Tulane University Working Paper*.

El-Shaarawi, A. H., Esterby, S. R. & Kuntz, K. W. (1983). A Statistical Evaluation of Trends on the Water Quality of the Niagara River. *Journal of Great Lakes Research,* 9(2): 234-240.

Esterby, S. R. (1996). Review of Methods for the Detection and Estimation of Trends with Emphasis on Water Quality Applications. *Hydrological Processes*, 10(2): 127-149.

GAO (General Accounting Office). (2000). Water quality: key EPA and state decisions limited by inconsistent and incomplete data. Washington, DC: United States General Accounting Office. GAO/RCED-00-54.

Gilliom, R.J. & Helsel, D. R. (1986). Estimation of Distributional Parameters for Censored Trace Level Water Quality Data: 1. Estimation Techniques. *Water Resources Research* ,22(2): 135-146.

Grant, L. E. & Langpap, C. (2019). Private provision of public goods by environmental groups.

PNAS, 116 (12) 5334-5340.

Hirsch, R. M. & Slack, J. R. (1984). A Nonparametric Trend Test for Seasonal Data with Serial Dependence. *Water Resources Research*, 20(6): 727-732.

Keiser D.A., & Shapiro, J. S. (2019). Consequences of the Clean Water Act and the Demand for Water Quality. *The Quarterly Journal of Economics*, 134(1): 349–396, doi:10.1093/qje/qjy019.

Keiser D.A., Kling, C. L., & Shapiro, J. S. (2019). The low but uncertain measured benefits of US water quality policy. *Proc Natl Acad Sci* USA PNAS October 8, 2018 https://doi.org/10.1073/pnas.1802870115.

McConnell, V. D. & Schwarz, G.E. (1992). The Supply and Demand for Pollution Control: Evidence from Wastewater Treatment. *Journal of Environmental Economics and Management*, 23(1): 54-77.

MDE (Maryland Department of the Environment). (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. EPA Approval: October 16, 2003 Document version: December 26, 2002: https://mde.state.md.us/programs/Water/TMDL/ApprovedFinalTMDLs/Documents/www.mde.state.md.us/assets/document/Town%20Creek%20Main%20Document_final.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: https://mde.state.md.us/programs/Water/TMDL/ApprovedFinalTMDLs/Documents/www.mde.state.md.us/assets/document/conoco_TMDL_ReportFinal(1).pdf.

MDE. (1999). TMDL of Biochemical Oxygen Demand (BOD) for the Western Branch of the Patuxent River, Prince George's County, MD. EPA Submittal: December 3, 1999 EPA Approval: June 6, 2000: https://mde.state.md.us/programs/Water/TMDL/ApprovedFinalTMDLs/Documents/www.mde.state.md.us/assets/document/tmdl/westbranch/wb_tmdl_fin2.PDF.

MODNR (Missouri Department of Natural Resources) (2005). Total Maximum Daily Loads (TMDLS) For West Fork Sni-a-Bar Creek Jackson County, Missouri. Completed: November 21, 2005 Approved: January 6, 2006, https://dnr.mo.gov/env/wpp/tmdl/docs/0400-west-fork-sni-a-bar-tmdl.pdf.

MODNR. (1999). Total Maximum Daily Loads (TMDLs) for Rock Creek, Jefferson County, Missouri. MODNR Water Pollution Control Program Approved by the EPA December 1, 1999 Completed July 21, 1999, https://dnr.mo.gov/env/wpp/tmdl/docs/1714-rock-ck-tmdl.pdf.

SCDHEC (South Carolina Department of Health and Environmental Control). (1998), TMDL Submittal for Sanders Branch/Coosawhatchie River. Bureau of Water Pollution Control, State of South Carolina Administrative Record; EPA Approval: August 17, 1998, Submittal: January 21, 1998: https://scdhec.gov/sites/default/files/docs/HomeAndEnvironment/Docs/tmdlsand.pdf.

Sigman, H. (2005). Transboundary Spillovers and Decentralization of Environmental Policies. *Journal of Environmental Economics and Management,* 50(1): 82-101.

Sigman, H. (2004). Does Trade Promote Environmental Coordination?: Pollution in International Rivers. *Contributions to Economic Analysis & Policy*, 3(2) Article 2, https://doi.org/10.2202/1538-0645.1251.

Sigman, H. (2002). International Spillovers and Water Quality in Rivers: Do Countries Free Ride? *The American Economic Review*, 92(4): 1152-1159.

Smith, V. K., & Wolloh, C.V. (2012). Has surface water quality improved since the Clean Water Act? *NBER Working Paper* No. 18192, June 2012. Available at www.nber.org/papers/w18192.

Streeter, H.W. & Phelps, E.B. (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. K., Kazyak, P. F. & Weisberg, S. B. (1991). A Water Quality Model for a River Receiving Paper Mill Effluents and Conventional Sewage. *Ecological Modelling*, 58: 25-54.

USEPA. (2000). 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, October 2000.

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.

USGS. (2005). USGS surface and water quality models information clearinghouse. United States Geological Survey, Washington, DC, available at: http://smig.usgs.gov.

Vellidis, G., Barnes, P., Bosch, D.D. & Cathey, A.M. (2006). Mathematical Simulation Tools for Developing Dissolved Oxygen TMDLs. *Transactions of the ASABE*, 49(4): 1003-1022.

.