Teevrat Garg, Stuart E. Hamilton, Jacob P. Hochard and Evan Plous Kresch

(Not so) gently down the stream: river pollution and health in Indonesia

Article (Published version) (Refereed)

Original citation:
Garg, Teevrat and Hamilton, Stuart E. and Hochard, Jacob P. and Kresch, Evan Plous (2018) *(Not so) gently down the stream: river pollution and health in Indonesia*. Journal of Environmental Economics and Management. ISSN 0095-0696

Journal of Environmental Economics and Management 92 (2018) 35–53

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

(Not so) gently down the stream: River pollution and health in Indonesia ☆

Teevrat Garg [a], Stuart E. Hamilton [b], Jacob P. Hochard [c,*], Evan Plous Kresch [d], John Talbot [e]

[a] School of Global Policy and Strategy, University of California - San Diego, USA
[b] Department of Geography and Geosciences, Salisbury University, USA
[c] Department of Economics & Institute for Coastal Science and Policy, East Carolina University, USA
[d] Department of Economics, Oberlin College, USA
[e] Eastern Shore Regional GIS Cooperative, USA

ARTICLE INFO

Article history:
Received 18 July 2018
Accepted 16 August 2018
Available online 10 September 2018

JEL codes:
Q56
D62
H23
Q53
I18

Keywords:
Water pollution
Diarrhea
Indonesia
River networks

ABSTRACT

Waterborne diseases, often arising from freshwater pollution, are a leading cause of mortality in developing countries. However, data limitations inhibit our understanding of the extent of damage arising from freshwater pollution. We employ a novel hydrological approach combined with village census data to study the effect of upstream polluting behavior on downstream health in Indonesia. We find that upstream use of rivers for bathing and associated sanitary practices can explain as many as 7.5% of all diarrhea-related deaths annually. We also find suggestive evidence for differential avoidance behavior in response to different pollutants. Our approach relies on publicly available satellite data, open source hydrological models, and coarse village census data allowing us to estimate health externalities from river pollution in particularly vulnerable and data poor environments.

Waterborne diseases are one of the leading causes of mortality in developing countries (Duflo et al., 2015; Gamper-Rabindran et al., 2010; Kremer et al., 2011; Do et al., 2018). These diseases, such as Acute Diarrheal Infections, are devastating because young children in rural areas tend to be the most vulnerable. Access to healthcare among these populations is limited, which is compounded by the absence of formal sanitation facilities and widespread freshwater pollution in rural impoverished areas. In such a context, the social cost of freshwater pollution remains poorly understood, especially at a large national scale.

In this paper, we consider the case of river pollution and resulting waterborne diseases in Indonesia focussing on diarrhea,

☆ This paper has benefited from excellent feedback from Chris Barrett, Goeff Barrows, Prashant Bharadwaj, Josh Graff Zivin, Kelsey Jack, Dan Kaffine, Haiyong Liu, Mushfiq Mobarak, Kevin Roth, Kyle Rozema, Savitar Sundaresan, Jay Shogren, seminar participants at Wageningen University, University of Namur, and conference participants at the CSAE (Oxford), NEUDC (MIT), AERE (Breckenridge, Colorado), Heartland Workshop (University of Illinois - Urbana Champaign), MWIEDC (University of Wisconsin - Madison) and the Southern Economic Association (Atlanta, Georgia). We also thank Carl Flint for testing a random sample of the upstream datasets against alternate algorithms and assisting in the development and processing of the upstream GIS parallel processing operation; Charles Ehlschlaeger, Markus Metz, and the US Army Construction Engineering Research Laboratory for designing open-source algorithms used during the GIS analysis. The FOSS contributors to Linux, GNU Parallel, GDAL, GRASS, POSTGIS, and PostgreSQL were essential to the GIS analysis. Garg acknowledges funding from ESRC Centre for Climate Change Economics and Policy, the Grantham Foundation for the Protection of the Environment and the Cornell Population Center.

* Corresponding author.
E-mail address: hochardj15@ecu.edu (J.P. Hochard).

https://doi.org/10.1016/j.jeem.2018.08.011

which globally accounts for more than 1.5 million deaths each year (WHO, 2014). Freshwater pollution is an important consideration in low- and middle-income countries where untreated river water is routinely consumed, in part due to the lack of access to alternative sources of water and low enforcement of policies intended to prevent contamination.[1] Earlier work in the area of freshwater pollution has focused primarily on industrial waste disposal in rivers (Ebenstein, 2012; Cai et al., 2016; Chen et al., 2016). We emphasize the role of household river pollution, particularly in-river bathing and associated sanitary practices, and find that upstream river-based sanitary practices in Indonesia can explain as many as 7.5 percent of all diarrheal deaths in a given year, which over our four year sample translates to 860 diarrheal fatalities. We find suggestive evidence that households engage in avoidance behavior in response to upstream use of rivers for trash disposal but not to upstream use of rivers for bathing and sanitary practices.

There are two main barriers that researchers face to estimating the effects of freshwater pollution on human populations in poor countries. First, for a number of reasons, data on fresh water pollution levels is incorrect, incomplete or entirely missing especially in locations where this problem is most pressing (Excell and Moses, 2017). In many cases, maps of existing river networks, especially minor rivers that are more likely to be used for water consumption, are unavailable. Second, even when such data are available, finding a relevant instrument or natural experiment to generate plausibly exogenous variation in fresh water pollution remains challenging.

We overcome previously identified data constraints in the area of fresh water pollution in two ways (Currie et al., 2014). First, to identify missing river networks we use publicly-available remote-sensed data on river basins and high-resolution digital elevation models to trace out river pathways. We subsequently "ground-truth" these through village census data that provide information on a village's on-river status. Second, absent comprehensive data on water pollution levels in Indonesia, we use information on polluting behavior of upstream villages as a proxy for downstream pollution levels. This is a reasonable assumption given the natural flow of water, sediment and pollutants downstream. Our approach has the advantage of being generalizable to other settings in low-income countries where on-the-ground pollution monitoring is rare but remotely-sensed data and satellite imagery are publicly available.

We address the identification challenge of finding meaningfully large variation in river pollution that is exogenous to local health outcomes in three ways.[2] First, we use our novel dataset of drainage basins in Indonesia to assign to each of the villages in our sample their respective set of upstream and downstream villages from approximately 5.8 billion possible upstream-downstream hydrological linkages. As a starting point, we rely on the identifying assumption that within-province (the administration of water pollution regulation is at the provincial level) year-to-year changes in upstream polluting behavior are exogenous to downstream health outcomes. This assumption is plausible because household level in-river polluting behavior in Indonesia is *de facto* unregulated, minimizing the selection across villages under the same regulatory jurisdiction. Moreover, we use data on *upstream* polluting behavior rather than data on *local* pollution levels. As a result, we are not relying on correlating local pollution with local health outcomes, which could be spurious for many reasons including but not limited to geographic sorting. Instead, we rely on variation in upstream polluting behavior that, in a *defacto* unregulated environment is plausibly exogenous to downstream health outcomes. As such, our strategy can be employed in other contexts when it is important to understand downstream externalities and the path of the pollutant determines the marginal social cost of pollution.

Second, we test the validity of our identifying assumption and rule out geographic sorting through a battery of falsification tests and find that for a given village, bathing and associated sanitary practices by upstream villages increases diarrheal incidence, while similar polluting behavior by downstream villages has no effect. Furthermore, we show that the effect is specific to diarrhea (consistent with its waterborne nature), with no measurable effect on other non-waterborne diseases. The absence of an effect on other diseases is inconsistent with the existence of a spurious correlation between upstream polluting and downstream health due to, for example, wealth or income, as one would expect to see such a correlation manifest as a significant impact between upstream bathing and at least one of the other diseases as well.[3]

Third, we address the potential of upstream river bathing and associated sanitary practices serving as a proxy for *ease of river access*, and show that there are no meaningful negative health impacts for other upstream river uses, such as transportation and agriculture. Importantly, while alternative explanations could exist for each of our results, it is unlikely that a plausible alternative explanation would rationalize all of our empirical findings.

We contribute to the growing intersection of environmental and development economics. Specifically, we build on the literature emphasizing the estimation of causal impacts of environmental quality on human health, particularly in developing

[1] Greenstone and Hanna (2014) find that while air pollution regulations had a measurable impact on infant mortality in India, water pollution laws had no measurable effect.

[2] While researchers have previously employed randomized designs in subsidies for provision of clean water (Ahuja et al., 2010; Kremer et al., 2011; Guiteras et al., 2016), to the best of our knowledge no one has used experimental or quasi-experimental variation to study the impact of household river pollution on local health outcomes at a national scale. The closest study to our work is Ebenstein (2012) who uses rainfall as an instrument for water quality (and not polluting behavior or biological pollutant concentrations) at certain sites in China to study the impact of poor water quality on digestive cancers. More recently, Lipscomb and Mobarak (2017) have examined the political economy of water pollution in Brazil, and Kahn et al. (2015) have done so in China.

[3] The remaining threat to identification would be if individuals sorted geographically across upstream and downstream villages within a province in our time frame and in response to time-varying omitted variables that were correlated with upstream polluting behavior and diarrheal outbreaks, but not upstream polluting behavior and outbreaks of malaria, measles, respiratory infections or dengue. Given that in Indonesia, downstream (coastal) villages tend to be economically better off than upstream villages, this seems very unlikely.

countries (see Graff Zivin and Neidell (2013) for an exhaustive review).[4] Our contribution is a novel method to generate data on pollution exposure in data-scarce contexts, and subsequently employs an identification strategy relying on successfully being able to match upstream and downstream pairs of villages. We also extend the literature on avoidance, which, in developing countries has focused primarily on governance (Galiani et al., 2005) and accessibility (Jalan and Ravallion, 2003; Ahuja et al., 2010) of clean water provision. While avoidance strategies such as relocation (Currie et al., 2013) and the purchasing of clean bottled water (Graff Zivin et al., 2011) have been documented in the United States, less evidence exists in rural areas. We document differential avoidance behavior across pollutants; households along trash-polluted rivers seek out alternative drinking water sources relative to those downstream of river bathers.[5]

The rest of the paper is organized as follows. In section 1 we provide an overview of water pollution regulation in Indonesia and the epidemiological evidence on the link between river bathing (and associated hygienic practices) and diarrhea. In section 2 we describe the health and demographic data used in this paper, as well as the construction of the upstream-downstream village networks. Section 3 details the econometric strategy that we use and in section 4 we discuss the empirical results. We provide robustness checks, as well as extensions of these results in section 5. Section 6 offers our concluding notes.

1. Background and institutional context

In this section, we provide a brief overview of the state of water pollution and associated regulations in Indonesia to demonstrate that river pollution, particularly originating from households, is *de facto* unregulated. We also review the evidence on the epidemiological foundations of the impact of river bathing and sanitary practices on water quality.

1.1. Water pollution regulation in Indonesia

Indonesia has made recent advances in environmental regulation, including the 2009 Environmental Protection and Management Law that recognizes the "serious problem" of decreasing environmental quality, as well as executive actions designed to reduce emissions and other forms of pollution (Nachmany et al., 2014). Yet, the regulation of water pollution in Indonesia can be characterized as nominally mandated but not regulated for some industries, and fully non-existent for others, especially in the period covered by our study (2000–2008). Ostensibly, any individual or business that purposely pollutes or otherwise damages water sources can face imprisonment for up to 9 years and a maximum fine of 1.5 billion rupiah (USD 105,000), in accordance with Article 94 of Indonesia's Law No. 7/2004 concerning water resources. The provincial governments are responsible for the regulation and supervision of all water resources, including rivers, that fall within their jurisdictions. Despite the steep penalties for polluters, the quality of Indonesia's water sources remains low. Several studies including those by the Ministry of the Environment (MoE) found that all 35 rivers that were tested across Indonesia were unsafe sources for drinking water (AECEN, 2008).[6]

The most unregulated source of water pollution in Indonesia is household and municipal discarding of sewage. Households routinely dispose of waste directly into rivers, while the improper construction of municipal wastewater facilities leads to the disposal of untreated sewage into river waters (Kerstens et al., 2013). For example, nearly two-thirds of the Citarum River's biological oxygen demand (BOD) comes from household pollution, as compared to one-third from all industrial and agricultural activities combined (Kerstens et al., 2013). Regulation of water pollution at the household level is non-existent, with households polluting into lakes and rivers with *de facto* impunity.

Water pollution in Indonesia is also generated from industrial waste and agricultural run-off (GWP, 2013). Industrial polluting causes toxic materials such as heavy metals and mercury to enter and poison drinking water sources. The 1989 Clean Water Program (CWP), a government initiative to curb water pollution, achieved spotty reductions in industrial pollutants with disproportionate success in East Java (Lucas and Djati, 2000). The mixed success of the CWP may be attributable to its enforceability, as the program was designed to be voluntary (Bedner, 2010) and water pollution regulations across Indonesia generally do not apply to small firms and home industries (Braadbaart, 1995).[7]

1.2. Use of rivers for bathing and sanitation

The use of rivers for bathing and sanitation purposes poses two major risks to human health, both of which are symptomatic of diarrhea. First, river bathing increases the amount of free carbon dioxide (CO_2) and decreases the amount of dissolved oxygen (DO) in rivers (Bhatnagar and Sangwan, 2009; Sharma et al., 2012). Organic and biodegradable waste from the bathers is decomposed by microbes that use oxygen and release carbon dioxide back into river water. This effect is amplified by the use

[4] Our work is positioned in the area water pollution (Carson et al., 2011; Ebenstein, 2012) with specific focus on a cause of diarrheal diseases (Dasgupta, 2004) in a country with a documented history of under-regulated and poor quality drinking water (Resosudarmo, 2003).

[5] We remain agnostic as to the reason behind differential avoidance behavior as it is beyond the scope of the exercise in this paper.

[6] In fact, the Citraum River has been found to be one of the most polluted places on Earth (Bernhardt and Gysi, 2013).

[7] This is true for small scale industrial enterprises, as well as small and family-run farms, which account for nearly a fifth of national GDP. In general, agriculture has been relatively less regulated than industrial activity. For example, the Clean Water Program only applies to industrial firms and no counterpart has been created for the agricultural sector.

of soaps and detergents that are absorbed by aquatic flora. Higher CO_2 levels drive phosphate and alkalinity concentrations, which lead to river eutrophication. Consumption of water from eutrophic rivers has been linked to gastroenteritis (WHO, 2002) and cyanobacterial toxins (Scott et al., 1985; Wu, 1999), both of which can cause symptoms of intestinal pain, nausea, and diarrhea.

Second, river use for bathing and sanitation can lead to an increased presence of coliform bacteria, which are associated with harmful pathogens known to cause nausea, vomiting, and bloody diarrhea, especially among infants and those with compromised immune systems (Joshi and Sati, 2011; Tyagi et al., 2013). While fecal coliforms do not necessarily cause diarrhea, their presence is correlated with diarrhea pathogens that may arise from removing trace amounts of fecal matter from the body during bathing. This is especially true for the cleaning of infants directly in the river after defecation, which greatly increases the amount of fecal matter present in the water.

2. Data

In this section, we describe the health and demographic data, the geospatial methods used to conduct the upstream and downstream assignment of villages along Indonesia's hydrological network, and the classification of each village into an identifiable drainage basin.

2.1. Health and demographic data

The Indonesian statistical agency, *Badan Pusat Statistik* (BPS), conducts a biennial census of all Indonesian villages known as *Podes*. Our sample consists of an unbalanced panel of 32,107 villages over four years (2000, 2003, 2006 and 2008) spread across all major Indonesian islands with the exception of Java.[8,9] The census is conducted in a short span of 4–6 weeks in October and November and consists of an exhaustive questionnaire to which village heads respond.[10] The census contains village-level information on health, population, river location and other demographic variables.

2.2. Disease data

The *Podes* provides the most exhaustive dataset on outbreaks and resulting deaths from five major diseases: dengue, diarrhea, malaria, measles and respiratory infections. *Podes* data is collected every 2–3 years in over 66,000 Indonesian villages. In each year of the census, the village head is asked to report if there was an outbreak of each of the different infections in that year. However, the village head is not provided any instructions - such as a cut-off or a point of comparison with respect to deaths or infections - in determining whether an outbreak has taken place. If the village-heads deems there is an outbreak, he/she provides an estimate of the total number of diarrhea-related fatalities. As such it is feasible that there were disease-driven deaths in a given year that were unreported in *Podes* because the degree of spread was not determined "high" enough to be considered an outbreak. As a result in section A.1, we demonstrate that any such bunching occurs around the critical threshold of zero, allowing us to minimize such concerns.

Table 1 provides summary statistics on the probability of village-level outbreaks for different diseases. Diarrhea, following only malaria, is the second most prevalent disease in Indonesia, followed closely by respiratory infections. Diarrhea is slightly more prevalent in hilly areas relative to flatter ones, and in rural areas relative to urban settings. A spatial mapping of diarrheal outbreaks across the Indonesian villages in our sample is presented in Fig. 1. This geographic pattern of disease prevalence is consistent across all diseases with the exception of dengue, which is endemic to urban areas. The differences between these groups are statistically significant but not meaningfully large (columns 2 and 3).

2.3. Population and demographic data

Our primary variable of interest is information on river bathing practices of villages in Indonesia. Importantly, river bathing is associated with river use for other sanitary activities, such as defecation (Hutton et al., 2008; Wasonga et al., 2014). Households that use the river for bathing also typically lack proper sanitation facilities, and use the river as an open sewerage system through the use of "hanging latrines" that deposit fecal matter directly into the river. Individuals - especially infants and young children - may defecate while bathing as well, which would contribute to an increase in the overall level of dangerous water-borne pathogens.

[8] We have 25,894 villages in 2000, 22,952 villages in 2003, 28,041 villages in 2006 and 32,107 villages in 2008. We construct our key independent variable - upstream bathing - separately for each year, reducing concerns over an unbalanced panel. In section 4 we show that our results are robust when we consider only a balanced panel.

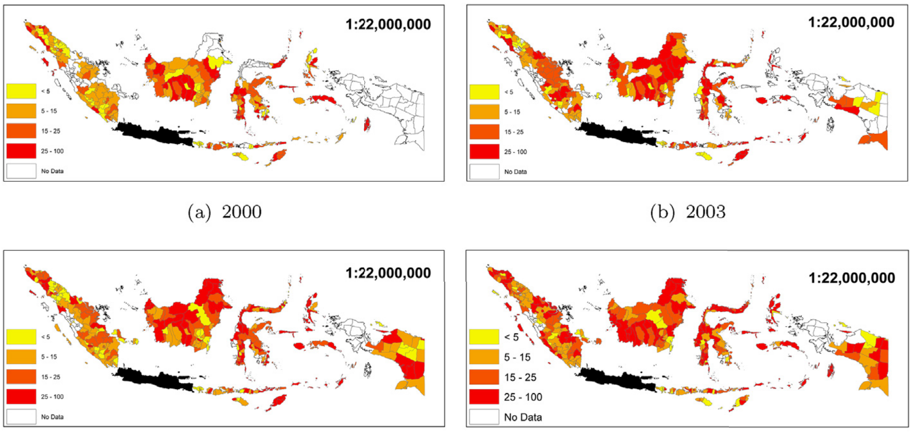
[9] Java, where the capital city of Jakarta is located, is the most densely populated island with an unusually high level of urbanization relative to the other islands. Additionally, Java has differential level of enforcement of water pollution that we do not have the data to capture. We leave it as a future follow-up exercise, and in this paper we focus on the other major Indonesian islands.

[10] We are aware that some components of this information are verified at the sub-district or district offices but we do not have information of the precise sections of the survey that are verified.

Table 1
Probability of village-level disease outbreaks.

VARIABLES	(1) Full	(2) Flat Villages	(3) Hilly	(4) Urban	(5) Rural
Was there an outbreak of diarrhea	0.183	0.158	0.188	0.158	0.185
	(0.387)	(0.365)	(0.391)	(0.365)	(0.389)
Was there an outbreak of respiratory	0.111	0.0978	0.118	0.101	0.112
	(0.315)	(0.297)	(0.323)	(0.301)	(0.316)
Was there an outbreak of measles	0.0807	0.0748	0.0866	0.0682	0.0820
	(0.272)	(0.263)	(0.281)	(0.252)	(0.274)
Was there an outbreak of malaria	0.207	0.165	0.231	0.127	0.216
	(0.405)	(0.371)	(0.421)	(0.333)	(0.411)
Was there an outbreak of dengue	0.0613	0.0698	0.0420	0.213	0.0454
	(0.240)	(0.255)	(0.201)	(0.409)	(0.208)
Observations	108,991	54,122	23,124	10,298	98,693

The table reports the proportion of villages (with standard deviations in parenthesis) in a given year reporting the incidence of diarrhea (row 1), respiratory diseases (row 2), measles (row 3), malaria (row 4) and dengue (row 5) in the full sample (column 1). We break down the full sample into flat villages (column 2) and hilly villages (column 3) as well as urban (column 4) and rural villages (column 5).



(a) 2000 (b) 2003 (c) 2006 (d) 2008

Fig. 1. Percentage of villages declaring a diarrheal outbreak in a year.

Podes reports the dominant location of bathing activity (e.g. in-river, other) for each village. This is a binary variable and *Podes* does not contain information on the number of individuals engaging in bathing. As an alternative, we use village populations to construct our key explanatory variables: the number of individuals polluting in the river through in-river bathing. To create this variable, we replicate the following exercise on river bathing for both upstream and downstream aggregate measures. For a given year t, village v, with n_{vt} villages upstream along the river that passes through village v, we define our key independent variable - the number of individuals engaging in upstream river bathing as,

$$Upstream_{vt} = \sum_{n_{vt}} \left[population_{nt} * bathing_{nt}\right] \tag{1}$$

where $population_{nt}$ is the population of the n_t^{th} upstream village and $bathing_{nt}$ is a binary variable, which is equal to 1 if the majority of households in the village bathe in the river in year t, and 0 otherwise. We repeat this exercise for all downstream villages.

There are two important considerations when using the above measure as our key independent variable. First, it is likely that in-river bathing is associated with other sanitary practices that result in increased concentrations of fecal contaminants in the river. Unfortunately, *Podes* does not directly ask about the nature and location of village defecating behavior.[11] However, it is likely the villages that use the river as a primary bathing source use it for sanitation purposes as well. Thus, we cautiously interpret our bathing variable as a proxy for the total of *all associated sanitation practices* that take place in the river.

[11] *Podes* does contain variables that ask about the presence of public and private latrines, but it does not specify the type of latrine (e.g. hanging latrine vs pit latrine) nor the location of the public latrines.

Table 2
Dumping variables summary statistics.

	Mean (Std Dev)
Do most people bathe in the river?	0.791
	(0.407)
Do most people drink from the river?	0.343
	(0.475)
Do most people primarily dump trash in the river?	0.0915
	(0.288)
Observations	108,991

The table reports the proportion of villages (with standard deviations in parenthesis) where most people use river for bathing and related sanitary practices (row 1), drink water from the river (row 2) and dispose off trash in the river (row 3).

Table 3
Upstream-downstream pairs summary statistics.

	Mean	Std Dev	Min	Max
Distance (Km)	60.125	60.779	0.09	1352.352
Population difference	2292.549	3660.449	0	163,161
Same river bathing practice?	0.645	0.478	0	1
Same river trash disposal practice?	0.885	0.319	0	1
In the same province?	0.860	0.347	0	1
Observations				13,123,273

The table reports the relationship between village pairs along the river network. Column 1 presents the calculated average across all village pairs, with columns 2–4 displaying the standard deviation, minimum, and maximum values, respectively. Relationships of interest between village pairs include the distance between pairs (row 1), the difference in population (row 2), correlated river bathing and trash dumping behaviors (rows 3 and 4), and shared province (row 5).

Second, it is important to note that our independent variables could be either over- or understated. Including all individuals as river bathers, where only the majority of individuals engage in river bathing, likely overstates the number of individuals bathing in rivers. However, excluding any individuals as river bathers in villages where less than a majority of individuals bathe in rivers understates the true number of individuals bathing in rivers. The concern of unpredictable measurement error in our key independent variables is addressed in our placebo tests in Table 6. Since we construct our upstream and downstream variables in the same way, any bias should be present in both, and given that the downstream effects are negligible, this measurement error is an unlikely source of bias in our results.[12] Cautiously, we may interpret our results as the differential impact of upstream bathing health effects net of downstream bathing health effects. We find equivalence in the two interpretations due to the approximate null effects of downstream river bathing.[13]

We also obtain population and demographic data from *Podes*. In addition to village population information, the census also contains information on whether a river passes through a village, which we use to ground truth the hydrological river network data. *Podes* also contains information on a range of socio-economic variables such as dominant source of income in village, geography of village, quality of governance (e.g. education of village head), access to medical facilities in the village and political status of the village.

2.4. Construction of drainage basins data and assignment of upstream and downstream villages

We link villages along a hydrological network enabling us to track the impact of upstream river bathers on downstream river users. Indonesia contains seven major islands with relatively mountainous and high-elevation interiors that create a complex hydrological network of streams and rivers. The official river network for the country fails to identify minor waterways that are being used by villages for bathing, drinking and trash disposal. As such, conducting a classic hydrological network analysis poses the risk of failing to assign villages located along minor rivers to the river network, which may understate upstream pollutant runoff (Fig. 2).

[12] Another concern may be turnover in village heads significantly altering the responses in *Podes*. *Podes* does not report turnover in village heads so it is not possible to include village head fixed effects. However, in order for this concern to undermine our identification strategy, village head turnover would have to be correlated with the changes in upstream behaviors, which seems unlikely.

[13] Another way to characterize our key independent variable described in equation (A.5) is as a scalar multiple of the population weighted average of the number of villages where a plurality of households engage in polluting behavior, with zero weight being assigned to villages where the dominant location of bathing is not in the river.

Table 4
Impact of upstream bathing on downstream diarrheal incidence.

Was there an outbreak of diarrhea?	(1) Full Sample	(2) Full Sample	(3) Drink from River	(4) Not Drink from River
Panel A: Full Panel				
Upstream Bathing	0.0222**	0.0250**	0.0530***	−0.000609
(100,000 individuals)	(0.0105)	(0.0107)	(0.0180)	(0.0132)
Downstream Bathing	0.00354	0.00294	0.0346	0.00121
(100,000 individuals)	(0.00248)	(0.00244)	(0.0297)	(0.00269)
Observations	108,991	106,797	36,819	69,978
R-squared	0.012	0.013	0.023	0.012
Village FE	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes
Panel B: Balanced Panel				
Upstream Bathing	0.0327**	0.0334**	0.0545**	0.00106
(100,000 individuals)	(0.0134)	(0.0139)	(0.0228)	(0.0165)
Downstream Bathing	0.00445	0.00360	0.111*	0.000286
(100,000 individuals)	(0.00393)	(0.00409)	(0.0566)	(0.00420)
Observations	48,128	47,054	18,326	28,728
R-squared	0.011	0.012	0.026	0.015
Village FE	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes

Cluster robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$ All specifications include village and province-year fixed effects. Standard errors are clustered at the river basin levels. All specifications include additional controls for total village population, total upstream population and total downstream population. Columns (2)–(4) include additional controls: dominant source of income in village, geography of village, quality of governance (education of village head), access to medical facilities in the village and political status of the village.

Table 5
Impact of upstream bathing weighted by distance.

Was there an outbreak of diarrhea?	(1) Full Sample	(2) Full Sample	(3) Drink from River	(4) Not Drink from River
Upstream Bathing	0.205*	0.229**	0.580***	−0.0551
(100,000 individuals)	(0.111)	(0.114)	(0.198)	(0.141)
Downstream Bathing	0.0318	0.0533	0.681	−0.282
(100,000 individuals)	(0.185)	(0.189)	(0.651)	(0.194)
Observations	108,991	106,797	36,819	69,978
R-squared	0.012	0.013	0.023	0.012
Village FE	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes

Cluster robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$ All specifications include village and province-year fixed effects. Standard errors are clustered at the river basin levels. All specifications include additional controls for total village population, total upstream population and total downstream population. Columns (2)–(4) include additional controls: dominant source of income in village, geography of village, quality of governance (education of village head), access to medical facilities in the village and political status of the village.

Instead of tracing the hydrological network directly, we adopt a watershed approach that identifies all upstream-downstream relationships within each basin using a high resolution digital elevation model (DEM), and then determine on-river status using survey data. Proprietary approaches to processing such a DEM are less adept to managing canopy interference - where the presence of tree canopy is mistaken for terrain - which could also render an underestimation of the number of upstream villages connected topographically to a given downstream village. The problem of canopy interference is compounded by the approximately 5.8 billion possible village relationships across Indonesia.[14]

We employ a high-resolution 30 m void-filled DEM alongside village administrative boundaries from *Podes*. Next, we constructed a pour point for each individual village that self-reported location on a river.[15] The use of self-reported locations in this way provides a means to "ground truth" our data to avoid misidentifying on-river (type 1 error) and off-river (type 2 error) villages.

[14] We manage canopy interference and computational processing constraints by developing a clustered implantation of the r.watershed and r.water.outlet algorithms in GRASS GIS v7.

[15] Each pour point identified the village's maximum upstream catchment, which is bound by its drainage basin.

Table 6
Placebo tests.

Was there an outbreak of	(1) Diarrhea	(2) Respiratory	(3) Measles	(4) Malaria	(5) Dengue
Upstream Bathing	0.0222**	0.00564	−7.24e-05	0.0108	−0.00687
(100,000 individuals)	(0.0105)	(0.00773)	(0.00848)	(0.0116)	(0.00747)
Downstream Bathing	0.00354	0.000322	0.00318	0.00637***	−0.000341
(100,000 individuals)	(0.00248)	(0.00214)	(0.00229)	(0.00243)	(0.00266)
Observations	108,991	108,991	108,991	108,991	108,991
R-squared	0.012	0.011	0.010	0.015	0.025

Cluster robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ All specifications include village and province-year fixed effects. Standard errors are clustered at the river basin levels. The specifications include additional controls for total village population, total upstream population and total downstream population. The sample is limited to villages that self-report proximity to a river.

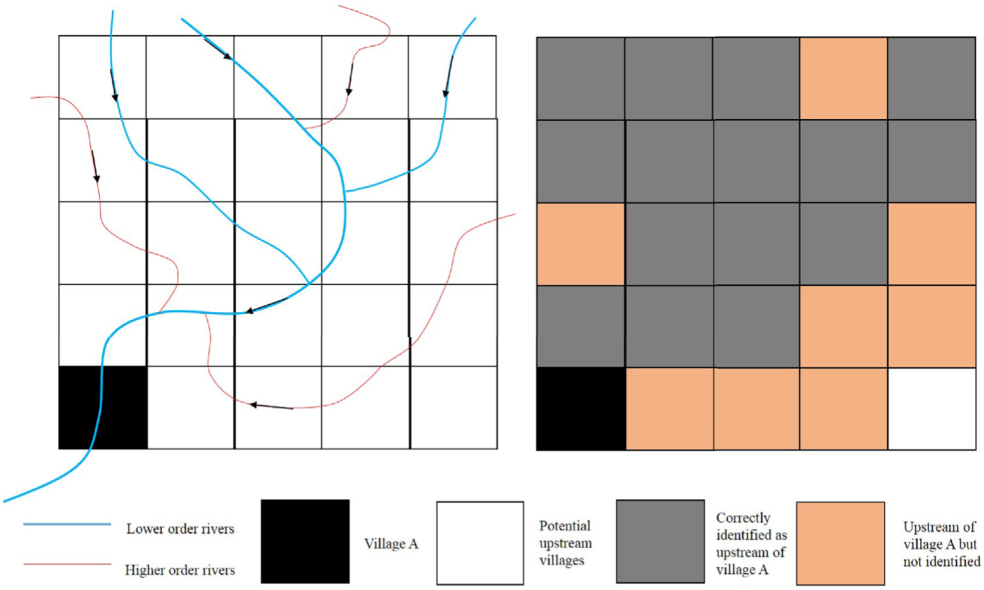


Fig. 2. Construction of river network data. (Left) low and high order rivers (right) improved assignment.

We map upstream and downstream village relationships independently for each sample year (2000, 2003, 2006 and 2008) to accommodate the redistricting of administrative units.[16] The product of the hydrological analysis is a list of every Indonesian village and its ordered upstream counterparts across the four sample years (approximately 13.7 million upstream observations). Summary statistics of these upstream-downstream pairs are provided in Table 3.

3. Estimation and identification strategy

The challenge in identifying the effects of water polluting behavior on human health is finding exogenous variation in water pollution that is large enough to capture an economically measurable effect. There are many plausible reasons why exposure to, and consumption of, impure water may be determined endogenously. For instance, poorer individuals who have a lower stock of health may be financially or behaviorally constrained from consuming clean water. Instead of correlating local water pollution with local health outcomes, we focus on the diarrheal incidence in a given village due to individuals who are geographically separated but whose (unregulated or unenforced) polluting behavior may affect downstream villages through river networks. Relying on the identifying assumption that within-province year-to-year changes in upstream polluting behavior are exogenous to downstream diarrheal incidence, we estimate a linear probability model:

$$I(DiarrhealOutbreak = 1)_{vpt} = \beta_1 Upstream_{vpt} + \beta_2 Downstream_{vpt} + \mathbf{X}'_{vpt}\delta + \gamma_v + \eta_{pt} + \epsilon_{vpt} \quad (2)$$

[16] Three sets of verifications were conducted. First, the GIS open source algorithms were compared against the ESRI algorithms. Second, the flow accumulations were consistent with the official Indonesian River Network. Third, the construction of drainage basins was verified by ensuring the official rivers network flowed properly through each basin ensuring that the constructed basins were logically sound.

$I(\cdot)$ is an indicator variable equal to 1 if the village v in province p at time t had an outbreak of diarrheal disease, and equal to 0 otherwise. $Upstream_{vpt}$ is the number of people upstream of village v that are engaging in river bathing in province p at time t. Similarly, $Downstream_{vpt}$ is the number of people downstream of village v that are engaging in bathing and other hygienic activities in the river. Appendix A.2 provides further discussion of the year-to-year variation in upstream populations and river bathing practices used to construct $Upstream_{vpt}$.

We also include a vector of controls $\mathbf{X}_{vpt}$ to account for the total population upstream, downstream, and within a given village. γ_v denotes the village fixed effects that control for time-invariant heterogeneity across villages and η_{pt} are the province-year fixed effects that control for provincial-level shocks across all years. ϵ_{vrt} is the error term, with standard errors clustered at the drainage basin to allow for arbitrary serial correlation across villages along the same drainage basin over time.

While focussing on geographically-separated polluting behavior can avert some endogeneity concerns, to the extent that individuals could geographically sort over time with wealthier individuals ending up in villages with cleaner water, the coefficient on upstream polluting, β_1 would remain biased. Since we use province-year fixed effects and thereby control for all province-specific changes over time, we will focus on such geographic sorting within a province over time. To overcome these concerns, we test the validity of our identifying assumption with a battery of placebo tests. First, upstream pollution could have an effect on downstream individuals' polluting behavior, but downstream polluting should not have a direct effect on upstream individuals' health. Second, we estimate the effect of polluting behavior on diseases that are not transmitted through ingestion of contaminated water, such as measles, malaria, respiratory infections, and dengue. If we are estimating a spurious geography-health correlation instead of the causal effect of upstream bathing on downstream diarrheal incidence, then we should also see association with these other diseases that are not waterborne. The absence of such effects would support the identifying assumption. Third, we add a range of time-varying control variables for changing demographic and poverty characteristics. As detailed in the results section, each of these tests supports our identifying assumption and thereby increases our confidence that we are estimating the causal effect of polluting behavior on diarrheal incidence.

Following equation (2), we also estimate avoidance behavior. In particular we test whether individuals reduce consumption of drinking water from the river in response to upstream polluting.

$$H(DrinkFromRiver)_{vpt} = \beta_1 Upstream_{vpt} + \beta_2 Downstream_{vpt} + \mathbf{X}'_{vpt}\delta + \gamma_v + \eta_{pt} + \epsilon_{vpt} \tag{3}$$

where $H(\cdot)$ is an indicator function equal to 1 if most people in that village drink water from the river.

Two additional econometric issues bear noting. First, there could be potential concerns over the choice of our estimator. We find that our estimates are robust to different choices of estimator. In particular, we use the fixed effects logit estimator and find that our results are qualitatively similar, as reported in Appendix Table A.1. We also find that a mere 0.3 percent of our observations (433 out of 108,991) had predicted values outside the [0,1] range, suggesting that fit is not a concern with the use of a linear probability model. Second, we cluster our standard errors at the drainage basin to allow for errors to be correlated across villages along the same river segment and over time. Given that the pollutants accumulate along a river segment, clustering at the drainage basin allows for conservative inference on the effects of upstream polluting behavior.

4. Results

4.1. Health effects

We are primarily interested in the effects of in-river bathing and other household sanitary practices on downstream health outcomes. We estimate the effects of polluting behavior on both upstream and downstream locations simultaneously using equation (2) and find that upstream bathing causes increased diarrheal incidence whereas downstream polluting behavior has no effect on upstream health outcomes. Specifically, we find that an additional 100,000 individuals upstream using the river for bathing and related sanitary activities increases diarrhea incidence by 2.52 percentage points or 13.7 percent (Table 4, panel A, column 1–2). To contextualize, a one sample standard deviation increase in the number of people upstream from a village (182,940 individuals) using the river for bathing and sanitary activities increases the probability of diarrheal outbreak in that village by 4.59 percentage points. Using the within-sample average of 18.29 percent diarrheal incidence, this corresponds to a 25.10 percent effect. The result is stable to the choice of specification (Appendix Table A.1). The magnitude of the effect of upstream bathing on downstream diarrheal incidence is considerably larger when limiting our sample to only those villages where the primary source of drinking water is the river (column 3) with a negligible effect when limiting our sample to villages where the primary source of drinking water is from sources other than the river (column 4). Cumulatively yet conservatively, we estimate that upstream hygienic activities can explain 860 deaths, which is 7.5 percent of all diarrheal deaths in our sample, suggesting that there is a large human cost to household polluting behavior.[17]

We find no evidence on the impact of diarrheal incidence from downstream river bathing (Table 4, panel A). As any pollutants from downstream villages cannot travel up the river naturally, the economically small and statistically insignificant coefficient associated with downstream bathing behavior lends further credibility to our identification strategy. As noted in section 2, to overcome challenges of potential measurement error in the construction of our main independent variable (upstream household

[17] It is important to note that these are equilibrium effects, net of any possible avoidance behavior. The lack of significant effect for those villages that do not drink from the river may in part be an endogenous effect of those villages avoiding particularly polluted sections of the river.

polluting behavior), we can think of the net effect as the difference between the coefficient on upstream and downstream polluting behavior, since the measurement error should be similar given the identical method of construction of both variables. The null effect of downstream polluting behavior suggests that measurement error does not appreciably bias our main result.

Additionally, as noted earlier, our panel data set is unbalanced. To address concerns over selection of different villages over different time periods, we limit our sample to the balanced panel (Table 4, panel B) and find that our results are qualitatively and quantitatively similar to those from the full sample.

Finally, we address the concern that pollution dissipates over long distances along the river network. In the main specification, the $Upstream_{vt}$ indicator provides equal weight to each individual in a village that partakes in river bathing. However, it is possible that bathing activity from more distant villages may have little to no impact on disease outcomes if the contaminants dissipate over the course of the river. Given the method of river network construction discussed in section 2, we can address this mechanism by weighting individuals in closer villages more than those who are farther away. To do this, we calculate an alternative upstream (and downstream) measure for bathing behavior as:

$$Upstream_{vt} = \sum_{n_{vt}} \left[population_{nt} * bathing_{nt} * f(distance_{nt})\right] \quad (4)$$

where $distance_{nt}$ is the distance (in meters) of the centroid of village n from village v. The function $f(\cdot)$ weights the relative distance of the villages with the property that $f' \leq 0$.

Table 5 presents main regression results using this distance-weighting indicator.[18] The results are similar to those in Table 4 in both sign and statistical significance. Notably, the distance-weighted results are larger by an order of magnitude, which is consistent with the fact that bathing populations farther upstream have a dissipated impact on downstream human health. This magnitude increase is offset by the lower weighting of far upstream populations, and using the distance-weighted estimates predicts approximately 700 deaths over 4 years, which is a similar prediction to the mortality estimate using the main specification.

5. Robustness checks and extensions

In this section, we present additional validation for our identification strategy. We report results from (1) falsification tests to test the possibility of a correlation between location on a river stream, subsequent river use and health in general (not specific to diarrhea), (2) heterogeneity by geographic and topographic factors, (3) including rainfall and flooding as possible confounding variables along with causal bounds following Oster (2017), and (4) considering forms of upstream polluting behavior other than bathing and associated sanitary practices. We conclude this section by providing suggestive evidence of differential avoidance behavior occurring as a result of visible river pollution (trash dumping) as compared to inconspicuous forms of pollution (river bathing and sanitation).

5.1. Falsification tests

In this section, we provide the results from a battery of falsification tests to further validate our identification strategy. Table 6 shows that the effect we find is (1) specific to diarrheal incidence and (2) specific to *upstream* polluting behavior. As shown in the table, we find no evidence of an effect on diarrheal (or any other diseases) incidence in a given village in response to downstream bathing. Not only is the effect of downstream bathing statistically indistinguishable from zero, it is also an order of magnitude smaller than the effect of upstream polluting. Therefore, we can reasonably rule out concerns over drainage basin specific, time-varying factors that are correlated with both polluting behavior and diarrheal outcomes.

Additionally, we follow Garg (2017) and provide falsification tests on other diseases (Table 6, columns 2–5). Spurious correlation between downstream health (driven by, for example, demographic or political economy factors) and upstream polluting behavior should be shared when predicting the impact of upstream pollutants on other disease outbreaks. For instance, if individuals were geographically sorting over time in response to poor health conditions (but not to upstream bathing), then we should expect to see a correlation with incidence of at least one of the other diseases. The absence of any noticeable or meaningful effect on any of the other diseases further supports the validity of our identification strategy.[19]

5.2. Topographic and geographic factors

Next we isolate effects based on different geographies to deduce whether (1) the results are consistent with the topography of villages (flat versus hilly), which would affect the rate of water flow, and (2) there are systematic differences in the effects of upstream bathing across urban and rural villages in our sample. In Table 7, we provide results breaking down the effect of upstream bathing by topography (columns 2–3) and by urbanization (columns 4–5). Consistent with our intuition, we find that the bulk of the human health effect is concentrated in flat rather than hilly villages due to the propensity for eutrophication

[18] In this specification, we use the weighting function $f(distance) = 1/\log(distance)$.

[19] We also test for non-linearities in Appendix Table A.2 and find no evidence to support a non-linear relationship between exposure to upstream river bathing and downstream diarrheal incidence.

Table 7
Results disaggregated by geography.

Was there an outbreak of diarrhea?	(1) Full Sample	(2) Flat	(3) Hilly	(4) Urban	(5) Rural
Upstream Bathing	0.0222**	0.0248*	−0.0431	0.0178	0.0203*
(100,000 individuals)	(0.0105)	(0.0142)	(0.0531)	(0.0211)	(0.0117)
Downstream Bathing	0.00354	0.00281	−0.000537	0.00257	0.00615*
(100,000 individuals)	(0.00248)	(0.00294)	(0.00921)	(0.00430)	(0.00356)
Observations	108,991	54,122	23,124	10,298	98,693
R-squared	0.012	0.012	0.029	0.046	0.013

Cluster robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ All specifications include village and province-year fixed effects. Standard errors are clustered at the river basin levels. The specifications include additional controls for total village population, total upstream population and total downstream population. The sample is limited to villages that self-report proximity to a river.

Table 8
Impact of upstream bathing controlling for rainfall and flooding.

Was there an outbreak of diarrhea?	(1) Full Sample	(2) Full Sample	(3) Full Sample
Upstream Bathing	0.0220**	0.0226**	0.0212**
(100,000 individuals)	(0.0105)	(0.0104)	(0.0103)
Downstream Bathing	0.00366	0.00348	0.00341
(100,000 individuals)	(0.00248)	(0.00247)	(0.00247)
Observations	108,991	108,991	108,991
R-squared	0.012	0.014	0.014
Annual Rainfall	Desa Only	Desa Only	Desa & Upstream
Deviation from Average	Desa Only	Desa Only	Desa & Upstream
Flooding	No	Yes	Yes
Village FE	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Delta	1.003	1.273	1.340

Cluster robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ All specifications include village and province-year fixed effects. Standard errors are clustered at the river basin levels. All specifications include additional controls for total village population, total upstream population and total downstream population. Column (1) includes controls for the annual rainfall in a given village, as well as the deviation in daily rainfall for the 5-year average. Column (2) uses the same controls as Column (1), in addition to a variable that measures the number of floods that the village experienced in the previous 3 years. Column (3) uses the same controls as Column (2), and includes the average annual rainfall and deviation from 5-year average for all of the upstream villages. The last row presents estimates of the coefficient of proportionality following (Oster, 2017), assuming values of $\beta = 0$ and $\overline{R}_{\max} = 0.3$.

in stagnant water (Jiménez, 2006). By contrast, we find no evidence of differential impacts in rural versus urban locations, suggesting that avoidance behavior may not be region-specific.

5.3. Rainfall and flooding

Another channel that may affect the estimation results is the presence of heavy rainfall events. Our results are robust to controlling for the frequency and level of rainfall in villages in a given year to control for both the direct physical channel, as well as the indirect income channel that may affect health outcomes in a village. We address these concerns by constructing village-level rainfall measures using data from NOAA's Climate Prediction Center (CPC).[20] The CPC's *Merged Analysis pentad* dataset provides globally gridded, 5-day precipitation estimates (in mm/day) from Jan 1979 to May 2013. We calculate each village's spatially-weighted average of rainfall in 5-day intervals between 1995 and 2012. With these estimates, we create various rainfall measures that are used as additional control variables for the main estimating equation. The results of these regressions are presented in Table 8 and are similar in both magnitude and significance to the coefficient estimate of Table 4 - indicating that neither rainfall nor flooding appears to play as significant a role in village health outcomes as the presence of upstream hygienic activities. Furthermore, we follow (Oster, 2017) and estimate that, conditional on the included fixed effects, the selection on unobservable variables would have to be 1.34 times the selection on observable variables for our estimate to be

[20] Data was downloaded from the IRI Data library of Columbia University's International Research Institute for Climate and Society, Earth Institute: http://iridl.ldeo.columbia.edu/.

Table 9
Measures of river access.

Was there an outbreak of diarrhea?	(1) Full Sample	(2) Full Sample	(3) Full Sample	(4) Full Sample	(5) Full Sample	(6) Full Sample	(7) Full Sample
Upstream Trash	0.00449						−0.000271
(100,000 individuals)	(0.0123)						(0.0121)
Upstream Irrigation		−0.0218				−0.0193	−0.0288*
(100,000 individuals)		(0.0140)				(0.0146)	(0.0149)
Upstream Industry			−0.000799			0.0108	0.00638
(100,000 individuals)			(0.00102)			(0.0141)	(0.0141)
Upstream Transportation				−0.000522		−0.00784*	−0.00833*
(100,000 individuals)				(0.000564)		(0.00454)	(0.00451)
Upstream Other Activities					−0.000214	0.000836	0.00238
(100,000 individuals)					(0.000339)	(0.00423)	(0.00418)
Upstream Bathing							0.0307***
(100,000 individuals)							(0.00786)
Observations	106,797	106,797	106,797	106,797	106,797	106,797	106,797
R-squared	0.012	0.013	0.013	0.013	0.013	0.013	0.014
Village FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Cluster robust standard errors in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$ All specifications include village and province-year fixed effects. Standard errors are clustered at the river basin levels. All specifications include additional controls for total village population, total upstream population, total downstream population, dominant source of income in village, geography of village, quality of governance (education of village head), access to medical facilities in the village and political status of the village.

biased away from zero.

5.4. Proxy for river access

While our identification strategy (complemented with the falsification tests) allows us to be reasonably confident we are likely estimating the causal effect of upstream polluting activity on downstream health, it is possible that our river bathing and sanitation variable is merely a proxy for upstream river *access*. Villages in which the river is more accessible (e.g. not flowing through a deep gorge) may promote increased river use for bathing and sanitation as well as for other purposes. While the falsification tests presented in Table 6 allow us to reasonably exclude a non-waterborne transmission mechanism, they do not rule out other pollution sources besides bathing, such as agricultural runoff that may be driving the main result.

We address this by using a question from *Podes* regarding the village's river use. The question provides a binary indicator for whether the river is used for a host of economic activities: trash disposal, agriculture, industry, transportation, and "other" miscellaneous activities. We replicate the calculation from equation (2) using these respective indicators in place of the river bathing indicator.

Table 9 presents the results of using these accessibility-to-river population indicators as a predictor of diarrheal outbreak. None of these activities alone (columns 1–5) have economically meaningful and statistically significant predictive power on the probability of a downstream diarrheal outbreak. When considering the set of upstream polluting activities on which we have information (column 6), we see a strong effect of upstream bathing and sanitary activities whereas the absence of any measurable relationship between upstream river access and downstream human health outcomes remains. As column 6 should display any of the predictive power related to river access besides river bathing, we can be cautiously optimistic that the effect we find in Table 4 is being driven by upstream river pollution related to bathing activities and not generically river access.

5.5. Do individuals protect themselves from river pollution?

Since upstream use of rivers for bathing and sanitation results in large downstream health externalities, to what extent do individuals seek to protect themselves from such river pollution? We use a question in *Podes* on the primary source of drinking water to understand the extent to which the use of the river for drinking water changes in response to upstream polluting behavior. In Table 10 we consider avoidance behavior in response to upstream use of rivers for bathing and sanitation versus trash disposal. In Panel A, we show that upstream bathing and sanitation results in no measurable change in drinking raw water from the river (column 1), whereas upstream trash disposal significantly (p-value = 0.06 for a two-sided test) reduces the number of villages consuming raw water from the river (column 2).[21] These results hold when we include the two variables together. In contrast, in Panel B, we show that while upstream bathing and sanitation has a strong effect on diarrheal incidence, no such effect is observed from upstream trash disposal (column 1), even when limiting our sample to villages where the primary source of drinking water is the river (column 2).

[21] We estimate avoidance behavior on the full set of potential upstream polluting variables considered in Table 9 and report these in Table A.3.

Table 10
Avoidance behavior.

Panel A: Avoidance Behavior			
Is the river used for raw drinking?	(1) Full Sample	(2) Full Sample	(3) Full Sample
Upstream Bathing	0.0155		0.0165
(100,000 individuals)	(0.0115)		(0.0114)
Upstream Trash		−0.0190*	−0.0195*
(100,000 individuals)		(0.0105)	(0.0104)
Observations	106,797	106,797	106,797
R-squared	0.017	0.017	0.018
Village FE	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Panel B: Main Result			
Was there an outbreak of diarrhea?	(1) Full Sample	(2) Drink from River	(3) Not Drink from River
Upstream Bathing	0.0235**	0.0528***	−0.00155
(100,000 individuals)	(0.0106)	(0.0177)	(0.0133)
Upstream Trash	0.00352	−0.0126	−0.00347
(100,000 individuals)	(0.0122)	(0.0168)	(0.0141)
Observations	106,797	36,819	69,978
R-squared	0.013	0.023	0.012
Village FE	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Cluster robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ All specifications include village and province-year fixed effects. Standard errors are clustered at the river basin levels. All specifications include additional controls for total village population, total upstream population and total downstream population. Panel A uses a binary variable “do most people drink raw water from the river” as the explanatory variable. Columns (1) and (2) include upstream bathing and trash separately whereas Column (3) includes both variables together. In Panel B, we report results on the effect of upstream bathing and trash on diarrheal incidence. Column (1) estimates this relationship over the full sample, whereas we limit the sample to those villages where the river is the primary source of drinking water (Column 2) and where the river is not the primary source of drinking water (Column 3). An extended version of Panel A containing other potential forms of upstream pollution is reported as Appendix Table A.3.

While we cannot explicitly test the motivations behind this gap in avoidance behavior, we hypothesize two plausible explanations. First, trash disposal in the river may pose a higher health risk than other river uses, and as such individuals are more active in avoiding the risk arising from the former. However, the primary effect of trash (particularly plastics) on human health is through endocrine disruptions, which may not manifest as noticeably as diarrheal episodes caused by organic water contamination (Thompson et al., 2009). Second, trash disposal could result in pollutants that are visible to the naked eye, in contrast to impurities generated from bathing that are less or not visible to individuals. This explanation is consistent with our results in Table 10 that individuals stop drinking water from the river in response to upstream trash polluting but not in response to upstream bathing. However, data limitations prevent us from identifying the exact mechanism behind the differences in avoidance behavior and we leave that as an open question for future research.[22]

6. Conclusion

In this paper, we construct and employ a novel data set on Indonesia's drainage basins to provide the first causal evidence that household-level polluting behavior and in particular upstream in-river bathing and associated sanitary activities generates large downstream health externalities. Our results are particularly relevant to policymakers for several reasons. First, we find that upstream river hygienic activities can explain as many as 860 deaths over four years representing 7.5 percent of all diarrheal deaths in our sample. This represents a large human cost from a source of river pollution that remains under-explored in the

[22] There is also suggestive evidence that villages are discretionary in their bathing activity as it relates to local agricultural activities - i.e., areas using the river for agriculture are less likely to bathe in the same waters, - which is confirmed by (i) a negative in-sample correlation coefficient of −0.13 between the *within village* river bathing and irrigation variables from *Podes* and (ii) an increase in the marginal effect of upstream river bathing on downstream diarrheal incidence in Table 9 when these upstream indicators are included together. Moreover, the Table 9 coefficient on “Upstream irrigation”, which serves as our single best proxy for upstream agricultural activity, is negative - further suggesting that villages are differentially avoiding use of the river for bathing and sanitation when it is simultaneously used for irrigation.

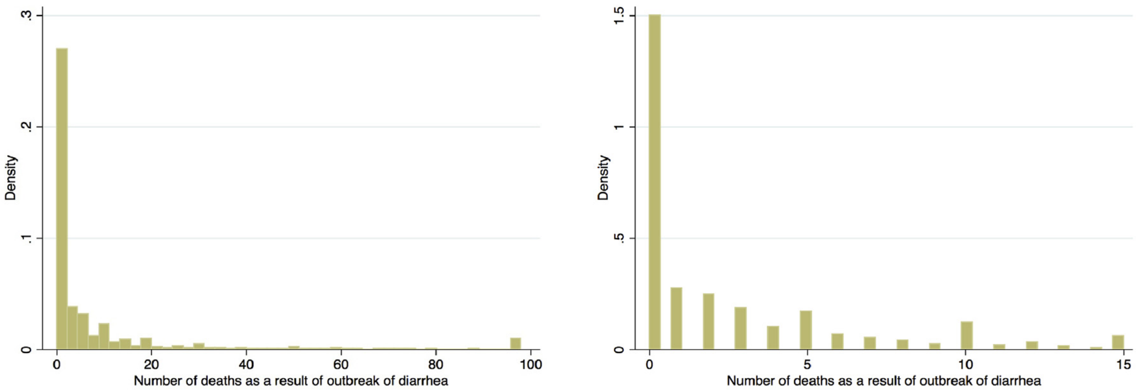
literature.[23]

We find suggestive evidence for differential avoidance behavior across different forms of upstream polluting behaviors. One of many candidate explanations is that individuals avoid visible but not invisible pollutants by seeking alternative sources of fresh water. If this is the case, considerable health savings may be obtained by investing in the prevention of “silent killers” against which households are less prone to taking defensive actions. Instead of large scale government programs aimed at river basin cleanup, which may be financially or technologically infeasible, policymakers may focus on preventing those polluting activities that are difficult for downstream households to detect. Future research should examine behavioral explanations for differences in avoidance behavior across pollutants.

A. Appendix

A.1. Validation of outbreak variable

Given that our outbreak variable is binary we provide below a histogram of the distribution of the number of deaths conditional on an outbreak. The bunching around zero suggests that the outbreak variable is sensitive enough to pick up even cases when disease occurs, but no deaths do.



(a) Full Sample (b) Less than 15 deaths

Fig. A.1. Distribution of number of deaths conditional on a diarrhea outbreak.

A.2. Variation in upstream bathing behavior

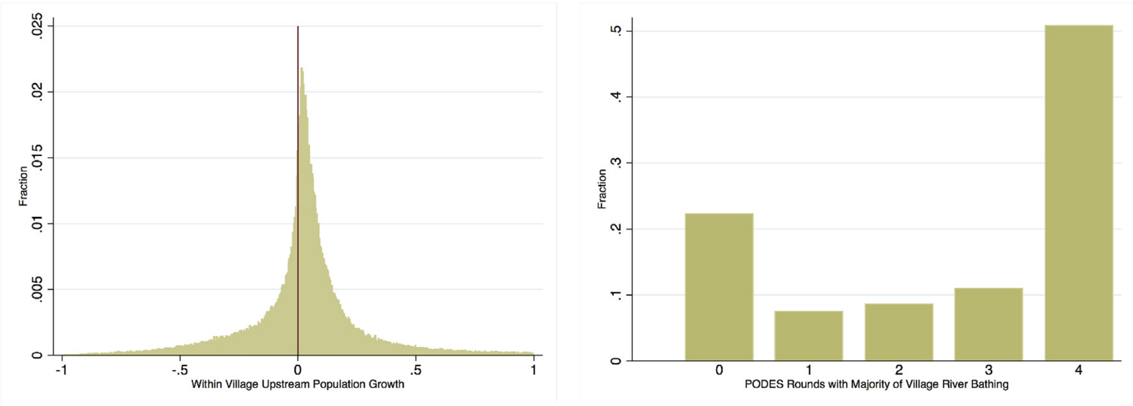
As the identification strategy relies on the *changes* in upstream river bathing populations, in this section we investigate the source of this variation across rounds of *Podes*. The key independent variable - the number of individuals engaging in upstream river bathing - is defined for a given year t, village v, with n_{vt} villages upstream along the river that passes through village v as,

$$Upstream_{vt} = \sum_{n_{vt}} \left[population_{nt} * bathing_{nt}\right] \tag{A.5}$$

where $population_{nt}$ is the population of the n_t^{th} upstream village and $bathing_{nt}$ is a binary variable, which is equal to 1 if the majority of households in the village bathe in the river in year t, and 0 otherwise.

Given that year-to-year variations in $Upstream_{vt}$ can come from changes in upstream populations, the changes in river bathing behavior of these upstream villages, or both, we provide below a histogram of the variation in these variables across the four rounds of *Podes*.

[23] Back of the envelope calculations suggest that policy targeting based on geographic location of the source of pollution can result in substantial health savings (see Appendix A.6). In particular, a 1 percent decrease in in-river bathers in the most upstream decile reduces fatalities by 2.54 percent. By contrast, a 1 percent decrease in in-river bathers in the most downstream decile reduces fatalities by only 1.62 percent. Policy options tailored to local geographic considerations are particularly relevant in Indonesia and other developing countries where limited resources for enforcement require precision targeting of point-source pollution.



(a) Upstream Population (b) River Bathing Behavior

Fig. A.2. Within village variation in river bathing and upstream population across years.

The within-village median change in total upstream population (i.e., regardless of river bathing practices) across the four rounds of *Podes* is 3.26%, with the histogram closely resembling the normal distribution. As population growth is not monotonically increasing or decreasing - 35% of population changes are *losses* - it is unlikely that these population changes are the sole driver of the observed effects in Table 2.

Panel (b) shows the year-to-year changes in river bathing practices for a given village. While nearly three quarters of villages either *always* or *never* use the river for bathing, there is an appreciable amount of “switching” within villages. Of the 25% of instances where villages switch river bathing practices, nearly half were villages that transitioned to non-river bathing behavior across the study period.

As both upstream population and river bathing practices display significant shares of positive and negative variation across *Podes* rounds, our identifying variation is coming from both changes in population and changes in use of rivers for bathing and sanitation.

A.3. Robustness to choice of estimator

In this section, we validate the choice of estimator. Column 1 presents the coefficients of a linear probability model (LPM) in predicting the diarrheal epidemic in a village in a given year. Using an LPM generates only 0.3% of predicted values that are outside the [0,1] range. Column 2 performs a similar estimation using a panel logit model instead of an LPM. The results of the estimation are qualitatively similar to the LPM regression, and maintains both the sign and level of significance for the bathing estimator.

Table A.1
Robustness to choice of estimator.

	(1) LPM	(2) Logit
Upstream Bathing	0.0222**	0.172**
(100,000 individuals)	(0.0105)	(0.0710)
Downstream Bathing	0.00354	0.0348
(100,000 individuals)	(0.00248)	(0.0299)
Observations	108,991	41,332
R-squared	0.012	0.026

Cluster robust standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$ All specifications include village and province-year fixed effects. Standard errors are clustered at the river basin levels. Both specifications include additional controls for total village population, total upstream population and total downstream population. Column (1) presents estimation using an OLS linear probability model, with approximately 0.3% of predicted values lying outside the [0,1] range. Column (2) presents the results of the fixed effects logit model. R-squared presented in the table is the psuedo-R2 calculated by the logit estimation. The number of observations using the FE Logit model is smaller than Column (1) due to all-positive or all-negative outcomes for village across all 4 years of the panel being dropped. The fixed effect logit model does not allow for the computation of the marginal effect at the mean value for the variables of interest - for a more detailed discussion, see Kitazawa (2012).

A.4. Testing for non-linearities

In this section, we test for whether there exists a non-linear relationship between upstream bathing behavior and diarrheal incidence in downstream villages. Column 1 presents the main result of the paper using the OLS estimator, and is identical to the corresponding column in Table 4. Column 2 applies a quadratic fit to the bathing estimator, and finds no significant relationship between diarrheal outbreak and the square of upstream bathing populations. The third column runs an OLS regression using the log of upstream bathing values. This monotonic transform of the explanatory variable is qualitatively similar to the main specification, and although the estimator is an order of magnitude smaller, it still estimates a positive and significant relationship. Thus we find no evidence to support a non-linear relationship between exposure to upstream river bathing and downstream diarrheal incidence.

Table A.2
Testing for non-linearities.

	(1) Linear	(2) Quadratic	(3) Logarithmic
Upstream Bathing	0.0222**	0.0227	
(100,000 individuals)	(0.0105)	(0.0139)	
Square [Upstream Bathing]		−8.20e-06	
(100,000 individuals)		(0.000365)	
Log[Upstream Bathing]			0.00492***
(100,000 individuals)			(0.00114)
Observations	108,991	108,991	108,991
R-squared	0.012	0.012	0.013
Village FE	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes
Controls	No	No	No

Cluster robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ All specifications include village and province-year fixed effects. Standard errors are clustered at the river basin levels. All specifications include additional controls for total village population, total upstream population and total downstream population. The sample is limited to villages that self-report proximity to a river. The mean number of upstream bathing households is 54,840.

A.5. Avoidance behavior and river access

In this section, we expand on Table 10 to include other forms of potential upstream polluting behaviors.

Table A.3
Avoidance behavior and river access.

Is the River Used for Raw Drinking?	(1) Full Sample	(2) Full Sample	(3) Full Sample	(4) Full Sample	(5) Full Sample	(6) Full Sample
Upstream Bathing	0.0165	0.0137	0.0136	0.0129	0.0136	0.0123
(100,000 individuals)	(0.0114)	(0.0121)	(0.0111)	(0.0111)	(0.0112)	(0.0121)
Upstream Trash	−0.0195*					−0.0206*
(100,000 individuals)	(0.0104)					(0.0107)
Upstream Irrigation		0.00628				−1.17e-05
(100,000 individuals)		(0.0119)				(0.0118)
Upstream Industry			7.95e-06			−0.00741
(100,000 individuals)			(0.00110)			(0.0132)
Upstream Transportation				9.78e-05		0.00161
(100,000 individuals)				(0.000622)		(0.00406)
Upstream Other Activities					1.65e-05	0.00187
(100,000 individuals)					(0.000367)	(0.00373)
Observations	106,797	106,797	106,797	106,797	106,797	106,797
R-squared	0.018	0.017	0.018	0.018	0.018	0.019
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Cluster robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ All specifications include village and province-year fixed effects. Standard errors are clustered at the river basin levels. All specifications include additional controls for total village population, total upstream population, total downstream population, dominant source of income in village, geography of village, quality of governance (education of village head), access to medical facilities in the village and political status of the village. Columns (1)–(5) report the results with individual sources of upstream polluting behavior whereas Column (6) reports results from the joint estimation.

A.6. Policy simulation

The impact of upstream bathing on downstream health suggests that policy responses to river pollution should be targeted with consideration to geography. Therefore, we conduct a set of policy simulations by imposing increasingly stringent moratoriums on river bathing (Table A.4).[24]

We show that the geography of targeting is essential to cost-effective policy. We categorize all sample villages into deciles based on total downriver population. Villages located near a river's headwaters with a large downstream population are grouped into the first decile while most downstream villages are grouped into the tenth decile (Table A.4). Targeting upstream villages generates the largest benefit - targeting the most upstream villages is an order of magnitude more effective than targeting downstream villages. Specifically, avoiding a single diarrheal death requires preventing 971,000 individuals in the most downstream decile from bathing but only 82,000 individuals in the most upstream decile. Our findings are therefore consistent with recent work on the political economy of water pollution (Lipscomb and Mobarak, 2017).

The baseline case, which most closely resembles the current state of affairs in Indonesia, has no regulation on river bathing activity. Population deciles with the largest downstream populations are then targeted incrementally until a complete moratorium on river bathing is achieved. Column 2 in Table A.4 shows the number of individuals bathing in the river in each decile.[25] In our two extreme cases, the absence of regulation on river bathing allows the 860 deaths attributable to river bathing to persist while a strict moratorium on river bathing prevents all of these deaths (Table A.4, column 3).

However, avoided deaths on decile-level moratoriums is not a comparable measure across the different deciles that have varying number of bathers. We generate two measures that allow us to compare moratoriums on different deciles - average and marginal number of individuals who must stop bathing to avoid a single instance of diarrheal mortality (Table A.4, columns 3 and 4 respectively). These cost calculations are akin to average and marginal costs of the policies per unit of benefit (Fig. A.3). Columns (3) and (4) show that a policymaker interested in reducing diarrheal deaths would have to inconvenience (or compensate) the fewest number of individuals per avoided death in the most upstream decile - 82,000 individuals who bathe in a river, versus 971,000 individuals in the most downstream decile. Conversely, in columns 5 and 6 (Table A.4), we develop an elasticity measure that shows reductions in mortality from a 1 percent reduction in decile-specific river bathers (Fig. A.4). Reducing top-decile bathers by 1 percent reduces marginal downstream diarrhea-related mortalities by 2.54 percent but reducing the lowest-decile bathers by 1 percent reduces marginal mortalities by only 1.61 percent.

Table A.4
Policy simulations.

Targeting Rule	(1) Bathing individuals impacted (millions)	(2) Predicted mortality reduction	(3) Average: Bathers removed per avoided death (thousands)	(4) Marginal: Bathers removed per avoided death (thousands)	(5) Average: % Change in mortality per 1% decrease in bathers	(6) Marginal: % Change in mortality per 1% decrease in bathers
No regulation of river bathing	0	0	–	–	–	–
Top decile only	2.4	29	82	82	2.54	2.54
Deciles (1–2)	5.5	62	88	93	2.37	2.28
Deciles (1–3)	11.5	123	93	99	2.23	2.21
Deciles (1–4)	21.8	226	97	101	2.16	2.27
Deciles (1–5)	37.8	356	106	122	1.97	2.03
Deciles (1–6)	55.8	493	113	132	1.84	2.13
Deciles (1–7)	77.6	620	125	171	1.67	1.97
Deciles (1–8)	103.7	731	142	236	1.47	1.80
Deciles (1–9)	133.6	813	164	364	1.27	1.61
Complete Moratorium on river bathing	179.5	860	209	975	1.00	1.00

Simulations include 108,990 village-year observations (2000, 2003, 2006 and 2008). Each 100,000 upstream bathing individuals increase epidemic rates by 0.025 while each epidemic yields 0.58 deaths, which is the sample average for all on-river villages across all years (excluding Java). The targeting rule deciles are based on total downstream population. For example, the top decile includes only those top 10% of villages with the largest downstream populations. Column one shows that increasingly stringent moratoriums on river bathing increase the number of bathing individuals affected, which corresponds directly with a reduction in predicted mortality displayed in column 2. Columns 3 and 4 capture the number of bathers that must be removed to prevent a death where column 3 is the average effect across all deciles under the moratorium and column 4 is the marginal effect of the most downstream moratorium decile. Columns 5 and 6 present estimates of the elasticity of placing a moratorium on river bathing. Column 5 captures the change in mortality per 1% decrease in bathers, averaged over the deciles under the moratorium. Column 6 presents the effect of the change in mortality per 1% decrease in bathers for the marginal decile placed under moratorium. Here, elasticities are generally decreasing resembling the impact of targeting the most upstream and detrimental individuals. However, because the policy rule is based on cumulative downstream population along the river, not the downstream bathing population, it is possible for the marginal elasticity measure to increase with the addition of a new decile.

[24] Simulations are cumulative across all four sample years (2000, 2003, 2006 and 2008) and the exposure variable (upstream bathing population) is updated by recalculating the sample mean with a moratorium imposed on villages that fall under the targeting rule. Epidemics are predicted using the rate of 0.0252 per 100,000 upstream individuals and deaths are predicted using the sample average of 0.58 diarrhea-related fatalities occurring within a village per diarrhea outbreak. Elasticity measures are calculated as the percentage change in diarrhea-related deaths divided by the percentage change in bathing individuals for each policy increment. Note that the marginal elasticities are not forced to be decreasing as they are calculated for incremental adjustments to the policy rather than as percentage changes from the no-regulation baseline (which would be necessarily decreasing in magnitude).

[25] Deciles are constructed based on total population, not bathing-in-river populations.

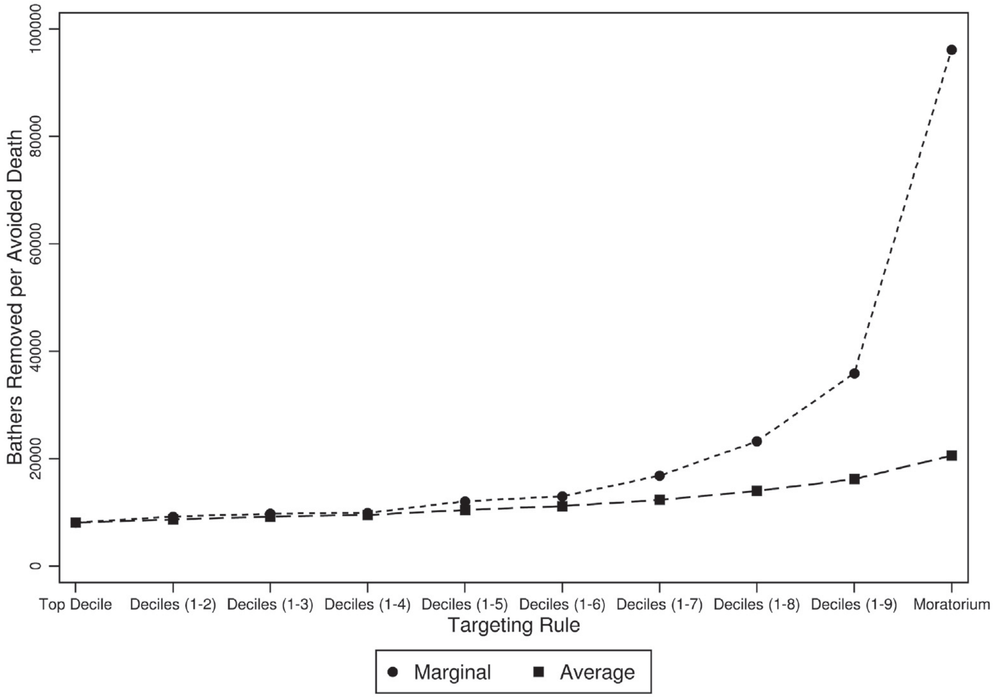


Fig. A.3. Marginal and average cost by geographic deciles.

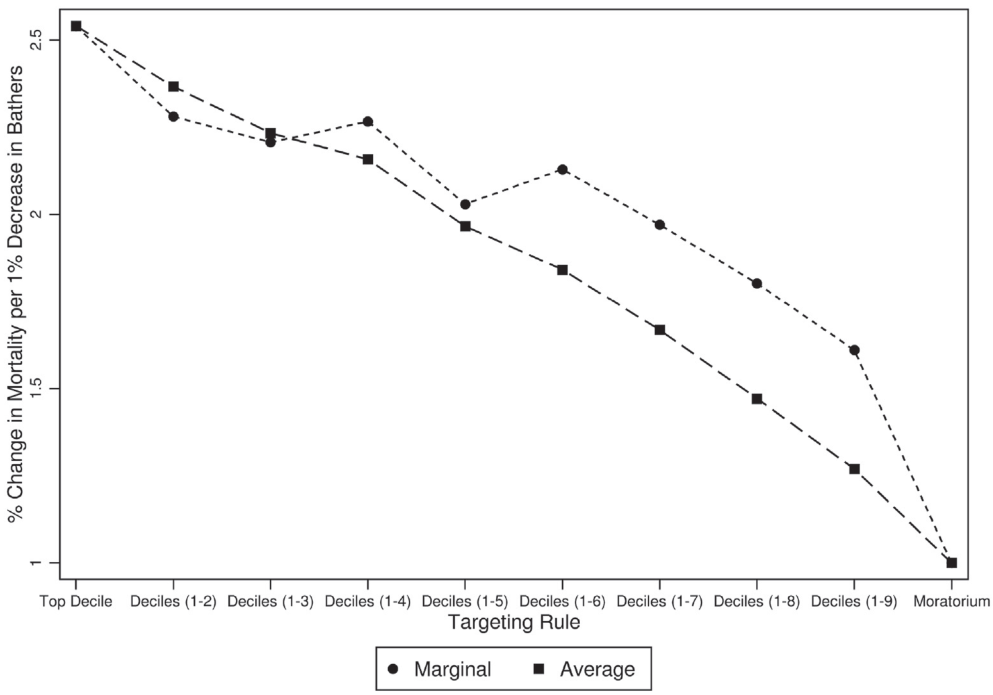


Fig. A.4. Percentage reduction in mortality per 1% decrease in bathers by targeting rule.

References

AECEN, 2008. Environmental Compliance and Enforcement in Indonesia: Rapid Assessment.
Ahuja, Amrita, Kremer, Michael, Peterson Zwane, Alix, 2010. Providing safe water: evidence from randomized evaluations. Ann. Rev. Resour. Econ. 2 (1), 237–256.
Bedner, Adriaan, 2010. Consequences of decentralization: environmental Impact Assessment and water pollution control in Indonesia. Law Pol. 32 (1), 38–60.

Bernhardt, A., Gysi, N., 2013. Blacksmith Institute and Green Cross Switzerland.
Bhatnagar, Anita, Sangwan, P., 2009. Impact of mass bathing on water quality. Int. J. Environ. Res. 3 (2), 247–252.
Braadbaart, Okke, 1995. Regulatory strategies and rational polluters: industrial waste-water control in Indonesia, 1982-1992. Third World Plann. Rev. 17 (4), 439.
Cai, Hongbin, Chen, Yuyu, Gong, Qing, 2016. Polluting thy neighbor: unintended consequences of China s pollution reduction mandates. J. Environ. Econ. Manag. 76, 86–104.
Carson, Richard T., Koundouri, Phoebe, Nauges, Céline, 2011. Arsenic mitigation in Bangladesh: a household labor market approach. Am. J. Agric. Econ. 93 (2), 407–414.
Chen, Zhao, Kahn, Matthew E., Liu, Yu, Wang, Zhi, 2016. The Consequences of Spatially Differentiated Water Pollution Regulation in China. National Bureau of Economic Research.
Currie, Janet, Zivin, Joshua Graff, Mullins, Jamie, Neidell, Matthew, 2014. What do we know about short-and long-term effects of early-life exposure to pollution? Ann. Rev. Resour. Econ. 6, 1.
Currie, Janet, Zivin, Joshua Graff, Meckel, Katherine, Neidell, Matthew, Schlenker, Wolfram, 2013. Something in the water: contaminated drinking water and infant health. Can. J. Econ. 46 (3), 791–810.
Dasgupta, Purnamita, 2004. Valuing health damages from water pollution in urban Delhi, India: a health production function approach. Environ. Dev. Econ. 9 (1), 83–106.
Do, Quy-Toan, Joshi, Shareen, Stolper, Samuel, 2018. Can environmental policy reduce infant mortality? Evidence from the Ganga Pollution Cases. J. Dev. Econ. 133, 306–325.
Duflo, Esther, Greenstone, Michael, Guiteras, Raymond, Clasen, Thomas, 2015. Toilets Can Work: Short and Medium Run Health Impacts of Addressing Complementarities and Externalities in Water and Sanitation. National Bureau of Economic Research.
Ebenstein, Avraham, 2012. The consequences of industrialization: evidence from water pollution and digestive cancers in China. Rev. Econ. Stat. 94 (1), 186–201.
Excell, Carole, Moses, Elizabeth, 2017. Thirsting for Justice. Transparency and Poor People's Struggle for Clean Water in Indonesia, Mongolia, and Thailand. World Resources Institute.
Galiani, Sebastian, Gertler, Paul, Schargrodsky, Ernesto, 2005. Water for life: the impact of the privatization of water services on child mortality. J. Polit. Econ. 113 (1), 83–120.
Gamper-Rabindran, Shanti, Khan, Shakeeb, Timmins, Christopher, 2010. The impact of piped water provision on infant mortality in Brazil: a quantile panel data approach. J. Dev. Econ. 92 (2), 188–200.
Garg, Teevrat, 2017. Ecosystems and Human Health: the Local Benefits of Forest Cover. UC San Diego Working Paper.
Graff Zivin, Joshua, Neidell, Matthew, 2013. Environment, health, and human capital. J. Econ. Lit. 51 (3), 689–730.
Graff Zivin, Joshua, Neidell, Matthew, Schlenker, Wolfram, 2011. Water quality violations and avoidance behavior: evidence from bottled water consumption. Am. Econ. Rev. 101 (3), 448–453.
Greenstone, Michael, Hanna, Rema, 2014. Environmental regulations, air and water pollution, and infant mortality in India. Am. Econ. Rev. 104 (10), 3038–3072.
Guiteras, Raymond P., Levine, David I., Luby, Stephen P., Polley, Thomas H., Jannat, Kaniz Khatun-e, Unicomb, Leanne, 2016. Disgust, shame, and soapy water: tests of novel interventions to promote safe water and hygiene. J. Assoc. Environ. Resour. Econ. 3 (2), 321–359.
GWP, Global Water Partnership, 2013. National Stakeholder Consultations on Water: Supporting the Post-2015 Development Agenda.
Hutton, A., Rodriguez, U.E., Napitupulu, L., Thang, P., Kov, P., 2008. Economic Impacts of Sanitation in Southeast Asia.
Jalan, Jyotsna, Ravallion, Martin, 2003. Does piped water reduce diarrhea for children in rural India? J. Econom. 112 (1), 153–173.
Jiménez, Blanca, 2006. Irrigation in developing countries using wastewater. Int. Rev. Environ. Strat. 6 (2), 229–250.
Joshi, Namita, Sati, Vandana, 2011. Assessment of water quality of river Ganges at Haridwar during Kumbh Mela-2010. Rep. Opin. 3 (7), 30–36.
Kahn, Matthew E., Li, Pei, Zhao, Daxuan, 2015. "Water pollution progress at borders: the role of changes in China's political promotion incentives. Am. Econ. J. Econ. Pol. 7 (4), 223–242.
Kerstens, S.M., Hutton, G., Van Nes, A., Firmansyah, I., 2013. Downstream Impacts of Water Pollution in the Upper Citarum River, West Java, Indonesia Economic Assessment of Interventions to Improve Water Quality, p. 32 Jakarta, Indonesia.
Kitazawa, Yoshitsugu, 2012. Hyperbolic transformation and average elasticity in the framework of the fixed effects logit model. Theor. Econ. Lett. 2 (2), 192.
Kremer, Michael, Leino, Jessica, Miguel, Edward, Zwane, Alix-Peterson, 2011. Spring cleaning: rural water impacts, valuation and property rights institutions. Q. J. Econ. 126, 145–205.
Lipscomb, Molly, Mobarak, Ahmed Mushfiq, 2017. Decentralization and pollution spillovers: evidence from the Re-drawing of county borders in Brazil. Rev. Econ. Stud. 84 (1), 464–502.
Lucas, Anton, Djati, Arief, 2000. The Dog Is Dead So Throw it in the River: Environmental Politics and Water Pollution in Indonesia: an East Java Case Study. Monash Asia Inst.
Nachmany, Michal, Fankhauser, Samuel, Townshend, Terry, Collins, Murray, Landesman, Tucker, Matthews, Adam, Pavese, Carolina, Rietig, Katharina, Schleifer, Philip, Setzer, Joana, 2014. The GLOBE Climate Legislation Study: a Review of Climate Change Legislation in 66 Countries.
Oster, Emily, 2017. Unobservable selection and coefficient stability: theory and evidence. J. Bus. Econ. Stat., https://doi.org/10.1080/07350015.2016.1227711.
Resosudarmo, Budy P., 2003. River water pollution in Indonesia: an input-output analysis. Int. J. Environ. Sustain Dev. 2 (1), 62–77.
Scott, W.E., Van Steenderen, R.A., Welch, D.I., Water Research Commission, et al., 1985. Health Aspects of Eutrophication. Available from the National Technical Information Service, Springfield VA. 22161, as PB 87-225645.
Sharma, Vijay, Bhadula, Sushil, Joshi, B.D., 2012. Impact of mass bathing on water quality of Ganga river during Maha Kumbh-2010. Nat. Sci. 10 (6), 1–5.
Thompson, Richard C., Moore, Charles J., Vom Saal, Frederick S., Swan, Shanna H., 2009. Plastics, the environment and human health: current consensus and future trends. Phil. Trans. R Sco. B: Biol. Sci. 364 (1526), 2153–2166.
Tyagi, Vinay Kumar, Bhatia, Akanksha, Gaur, Rubia Zahid, Khan, Abid Ali, Ali, Muntajir, Khursheed, Anwar, Kazmi, Absar Ahmad, Lo, Shang-Lien, 2013. Impairment in water quality of Ganges River and consequential health risks on account of mass ritualistic bathing. Desalin. Water Treat. 51 (10–12), 2121–2129.
Wasonga, Job, Olang'o, Charles Omondi, Kioli, Felix, 2014. Improving households knowledge and attitude on water, sanitation, and hygiene practices through school health programme in Nyakach, Kisumu county in Western Kenya. J. Anthropol. 2014, 1–6.
WHO, 2002. Eutrophication and Health. Office for Official Publications of the European Communities, Luxembourg. Geneva: WHO.
WHO, 2014. Preventing Diarrhea through Better Water, Sanitation and Hygiene: Exposures and Impacts in Low- and Middle-income Countries.
Wu, R.S.S., 1999. Eutrophication, water borne pathogens and xenobiotic compounds: environmental risks and challenges. Mar. Pollut. Bull. 39 (1), 11–22.