Waterborne diseases and children's learning

Steve Berggreen and Linn Mattisson*

February 4, 2023

Abstract

Unsafe water kills 1.2 million people every year, yet there is little research on the environmental and social risk factors of waterborne disease incidence and its consequences on children's learning. Building on recent advances in hydrological engineering, we construct a hydrological model for Tanzania that simulates the appearance of stagnant water pools – essential to the growth and spread of waterborne pathogens – which we use as a measure of waterborne disease potential. Using a difference-in-differences approach, we find that children exposed to one standard deviation larger waterborne disease potential have 0.03 standard deviations lower test scores, and the main symptom of waterborne diseases, diarrhoea, increases by 11%. These results mask important heterogeneities: We find that the most vulnerable children are those who live in urban areas with poor sanitation. Access to safe sanitation attenuates the negative effect of waterborne disease potential on both children's health and learning, which suggests that policy-makers should incorporate local environmental risk factors when implementing sanitation policies in regions vulnerable to waterborne disease.

*Berggreen: Department of Economics, Lund University, steve.berggreen-clausen@nek.lu.se. Mattisson: Department of Economics, Lund University, linn.mattisson@nek.lu.se. **Acknowledgements:** We are grateful to Jan Bietenbeck, Gunes Gokmen, Therese Nilsson, Mohammad Sepahvand, Anna Tompsett, Gianmarco León-Ciliotta, Matthew Collins and Martin Fisher for valuable comments and suggestions. We are also grateful for comments and discussions with seminar participants at the Lund lunch seminar, ASWEDE conference, and the CBU Climate change and inequality workshop.

1 Introduction

"For ten days or so in 2015 the world was watching the images from Nepal, where 9,000 people had died [in an earthquake]. During the same ten days, diarrhoea from contaminated drinking water also killed 9,000 children across the world. There were no camera teams around as these children fainted in the arms of their crying parents. No cool helicopters swooped in. Helicopters, anyway, dont work against this child killer (one of the worlds worst). All that is needed to stop a child from accidentally drinking her neighbor's still lukewarm poo is a few plastic pipes, a water pump, some soap, and a basic sewage system. Much cheaper than a helicopter."
– Rosling et al. (2018)

Waterborne diseases are a leading cause of death and disability in the world, contributing to 5.3% of DALYs (disability-adjusted life years) from all diseases and leading to 1.2 million deaths annually, of which 485,000 are from diarrhea alone.[1] This means the global disease burden of waterborne diseases is higher than that of AIDS or malaria (UNICEF, 2022b). In contrast to other leading global diseases, 90% of the disease burden specifically affects children under 5 years (Prüss et al., 2002). Waterborne diseases are spread by drinking or having physical contact with water, caused by microbes, bacteria and parasites harmful to humans. For example, cholera causes acute diarrhoea and is caused by the intake of cholera bacteria, and the virus E.Coli causes diarrhoea and abdominal pain.[2] In severe cases, in particular for young children and infants, these diseases are fatal.

The large documented costs to human life that waterborne diseases pose are made all the more tragic by the fact that waterborne diseases are almost completely preventable, as seen by the very low prevalence in high-income countries. While waterborne diseases only represent 0.4% of the total disease burden of Europe, the disease burden in Africa is upwards of 14%, making waterborne diseases one of the continent's most debilitating conditions. This makes policy to tackle waterborne diseases even more relevant, as large reductions in waterborne diseases are possible: For instance, the WHO estimates that 94% of diarrhoeal cases are preventable through changes in water, sanitation and hygiene (WASH) practices or other changes in the environment (WHO, 2010).

Despite the high disease burden of waterborne diseases, existing evidence is scarce on how waterborne diseases spread and debilitate children's health and capacity for learning if they are exposed to waterborne diseases during their schooling. We address this gap by simulating a plausibly exogenous risk factor of waterborne diseases: The formation of stagnant water pools in a hydrological model. We analyse this risk factor on key measures of children's learning: Proficiency tests in Mathematics, English and Swahili for school-aged children in Tanzania. We find that children exposed to one standard deviation larger waterborne disease potential (WBD Potential) have 0.03 standard deviations lower test scores. We also estimate an increase in the probability of contracting diarrhoea specifically, which is indicative of a higher prevalence of waterborne disease: One stan-

[1] According to estimates by the World Health Organization (WHO, 2022a) The WHO also states that 829,000 people die from diarrhea annually if combining deaths from unsafe drinking water and sanitation practice

[2] This is in contrast to e.g. mosquito-born malaria which is contracted in the bloodstream from a parasite.

dard deviation increase in WBD Potential increases the prevalence of diarrhoea by 10.6% relative to the mean diarrhoea prevalence. Interestingly, we find that WASH quality at the community level matters more than the household level for attenuating the negative effects of WBD potential. This suggests the existence of positive spill-over effects from households investing in improved WASH infrastructure and speaks against a "piecemeal approach" to improving WASH infrastructure, in line with recent findings. Exploratory analysis suggests that the children most affected by waterborne diseases are those who have been exposed to more waterborne disease risk factors in the past, suggesting the diseases act cumulatively on children.

Our outcomes of interest are related to children's health and education, as children are the most affected demographic by waterborne diseases. Our main analysis uses the Uwezo surveys for the period 2011-2017, a subnationally representative dataset for children in Tanzania. Key to our understanding of the effect of WBD Potential, Uwezo surveys employ standardised testing on all children aged 6-16 in the surveyed household. Hence, an important contribution of this paper is that we can more accurately assess to what extent disease affects children's current school performance and learning. Additionally, in contrast to many other education surveys we can assess the score of children not enrolled in school, allowing us to exclude the channel that our results are driven by selective school dropouts. Lastly, we complement the test scores with detailed survey information on children's health from three waves of the Demographic and Health Survey of Tanzania covering the years 1999, 2010 and 2015.

Our findings have several important policy implications: First, we show that where stagnant water pools form and the risk of waterborne diseases is higher, there are meaningful effects on the incidence of waterborne diseases and effects on children's performance in learning. Thus, the implementation of flood forecast models, which are by hydrological standards relatively easily maintained, to predict areas vulnerable to disease outbreaks can mitigate the burden of waterborne diseases in the future. Second, programmes seeking to reduce the incidence of waterborne diseases should aim to address sanitation practices at a village or community level as opposed to units of households, in particular in urban areas with higher population density.

To the best of our knowledge, this paper is the first to causally assess the costs of waterborne diseases to children's health and learning from climatic risk factors of waterborne disease. Most existing literature is epidemiological, in the form of observational studies, focusing on specific outbreaks or methodologies to trace and quantify disease outbreaks. Recent studies find an association between single intense rainfall events and disease outbreaks, and highlight the challenges in predicting these events (Levy et al., 2016; Lo Iacono et al., 2017). It is well-known that floods caused by heavy rainfall account for the majority of outbreaks of waterborne diseases (Cann et al., 2013). However, rainfall may affect socio-economic outcomes through many channels, especially in developing countries (Mellon, 2021). Our key contribution is that we can separate the impacts of rainfall from stagnant water, which is arguably a better measure of waterborne disease potential, by using hydrological modelling to simulate surface water flow. Methodologically, Deryugina et al. (2019) similarly model air pollution using local wind direction to causally infer the effect of air pollution on life expectancy among the elderly population in the US. Rinaldo (2020) use

machine-learning to model how the prevalence of the snail-host of schistosomiasis, a waterborne disease, affects economic development in Burkina Faso. In this work, we also make use of exogenous natural variation and develop a hydrological model customised to simulate the emergence and disappearance of areas of stagnant water over time. This time variation allows us to disentangle the effects of recently emerging stagnant water from inherent flood risk.

While we find no literature that studies the effect of waterborne disease on health and learning in terms of climatic and environmental risk factors, some studies exist that explore the effect of improving WASH infrastructure on health and mortality. Kremer et al. (2022) provide a meta-analysis of water-quality random control trials and find that water-quality interventions improve both reported diarrhoea incidence and mortality. Duflo et al. (2015) provide evidence that household-level programs to expand good sanitation practices reduces diarrhoea episodes by 30-50%. From a historical perspective, Alsan and Goldin (2019) find that improved water sewerage in Boston reduced child mortality by one-third of log points, Cutler and Miller (2005) find that expansion of high-quality water treatment in the US halved total mortality, and Knutsson (2020) find that cholera mortality almost disappeared with the introduction of water-filtration in 19th century Stockholm. Our results complement these findings by analysing an environmental risk source of waterborne disease driven by local climate, and by further analysing to what extent water and sanitation practices mitigate these effects.

The rest of this paper is organised as follows: Section 2 provides additional information about waterborne disease and education in Tanzania. Section 3 summarises our data sources and key variable, while Section 4 explains our empirical strategy. Section 5 presents our main results and Section 6 provide complementary robustness checks. Section 7 provides an exploration of the channels of the effects on learning. Finally, Section 8 concludes the discussion and presents our policy recommendations.

2 Institutional background

2.1 Prevalence and causes of waterborne disease in Tanzania

Waterborne diseases are illnesses caused by pathogens that are transmitted through contaminated water sources[3]. Tanzania lies in the 'belt' of the world's highest waterborne disease incidence, which stretches through the central part of Sub-Saharan Africa, a region characterised both by a favourable climate for waterborne pathogens as well as lack of access to safe water and sanitation. The number of diarrhoeal episodes per person per year is 1.1 episodes in Eastern Sub-Sahara where Tanzania is located, which is only second in the world to Central Sub-Saharan Africa (1.21) (Troeger et al., 2018). Tanzania has had several outbreaks of waterborne diseases such as cholera in the past few decades. A recent example of a cholera outbreak occurred around Tanzania's capital, Dar es Salaam, in 2015, with 16,521 reported cholera cases in total (Chae et al., 2022), but other waterborne disease outbreaks have been documented as recently as in 2022 (Masunga et al., 2022;

[3] See Appendix B for more details on the definition and characteristics of waterborne diseases

WHO, 2022b). Waterborne diseases have historically been classified as 'neglected' diseases by Tanzanian authorities and have received much less attention than officially prioritised diseases such as HIV and malaria (Tanzania Ministry of Health and Social Welfare, 2008), despite the fact that waterborne diseases carry a higher disease burden than HIV and Malaria for children in Tanzania (Vos et al., 2020).

Environmental factors play a critical role in enabling survival, growth and transmission of waterborne pathogens[4]. El Niño events which occur every 3-5 years and lead to an increase in rainfall and flood events across East Africa are associated with a 3-fold increase in cholera incidence (Moore et al., 2017). In a review of the epidemiological literature, Levy et al. (2016) found that the most important predictors of diarrhoeal disease outbreaks were floods, followed by heavy rainfall and high temperature. Heavy rainfall is thought to affect diarrhoea incidence primarily through its effect on floods and surface water contamination (Levy et al., 2016). Detailed mapping in a district in Tanzania showed that low-lying areas with high water-tables are more susceptible to cholera as a result of rainfall, as this leads to the accumulation of surface water which contaminates pit latrines and shallow drinking water wells (Mayala et al., 2003).

2.2 Prevention and treatment of waterborne disease in Tanzania

Tanzania has made great strides in reducing childhood diarrhoea mortality, with the mortality rate declining by 89% from 1980 to 2015 (Masanja et al., 2019), but incidence remains high. Most of the lives saved are attributed to treatment with oral rehydration salts (ORS), which replenishes fluids lost by diarrhoea, in combination with improved nutritional status, making children more resilient to the effects of diarrhoea. These are relatively affordable on the global market, costing close to 0.56 USD per treatment (UNICEF, 2022a). In contrast, less than 3% of the reduction in diarrhoea-related mortality is attributed to improved water and sanitation, which is evident by the fact that coverage increased only from 8 to 15% over the same time period (Masanja et al., 2019). Hence, disease management has been focused on treatment rather than prevention, leaving room for large variation in local disease environment and incidence. In high-income countries, large-scale historical investment in water and sanitation were crucial in reducing childhood mortality and diarrhoeal incidence (Alsan and Goldin, 2019). However, this infrastructure is costly, and lower-cost piecemeal approaches in Tanzania have been unable to curb diarrhoea-related morbidity (Briceño et al., 2017).

Treatment with ORS or antibiotics is usually administered at health clinics, which implies that access to high-quality healthcare is necessary to relieve consequences of waterborne disease infection. In 1984 Tanzania set up the National Control of Diarrheal Disease to combat childhood mortality, which focused on creating local clinics that could administer ORS (Masanja et al., 2019). This may explain why care-seeking for diarrhoea in children in Tanzania is among the highest in Sub-Saharan Africa, with more than half of children sick with diarrhoea taken to a health clinic (Schellenberg et al., 2003), and more than 90% live closer than 5 km from a primary health facility

[4]See Appendix B.2 for more details on transmission mechanisms

(Tanzania Ministry of Health and Social Welfare, 2008). However, this masks large inequalities in the quality of received healthcare. A study in rural Tanzania found that only 1 of 6 children with diarrhoea received ORS, citing low diagnostic capabilities and a lack of medical supplies (Schellenberg et al., 2003), while other parts of Tanzania see ORS administration rates over 50% (Masanja et al., 2019). It is not uncommon for rural families to bypass the local health clinic, especially if living close to a hospital (Kahabuka et al., 2011), which suggests that while there is high access to health clinics, quality is often poor and unevenly distributed. Thus, we expect waterborne diseases to spread more easily in urban, densely populated areas, where contamination of water sources is more common, but the health consequences once afflicted by a waterborne disease may be worse in rural areas with lower access to high-quality healthcare.

2.3 Education and learning in Tanzania

In Tanzania, schooling starts with pre-primary schooling at ages 5-6, although it is common that children also attend pre-school for 2-3 years, which by global standards is relatively formal schooling (Bietenbeck et al., 2019). This is followed by seven years of primary schooling at ages 7-13, four years of ordinary secondary school (ages 14-17), and two years of advanced level secondary school (ages 18-19)[5]. The adult literacy rate was 77.5% for men and 62.2% for women in 2012, but literacy among current pupils is higher, with the literacy rate being 86% for 15-24 year-olds. In 2012, 83% of the population reported having attained primary schooling, and 12.9% secondary schooling. Schooling expansion has been swift in Tanzania, and for example ordinary secondary school enrolment has increased from 6% in 2002 to 34% in 2013.

The effect of waterborne diseases on learning and education as a whole is understudied, but there are key links between health and the state of learning in Tanzania. For one, many students drop out and do not finish primary schooling; 65.1% (72.8%) of boys (girls) finish primary schooling. Most dropouts are due to unknown reasons (truancy), but health reasons are common. For one, early-life stunting, from e.g. repeated exposure to diarrhoeal diseases, is an important cause of both absence and lower cognitive ability. Second, current health issues also lead to greater dropouts. For example, in 2014, a survey conducted in three Tanzanian regions stated that health problems caused 6.2% of dropouts, and in a 2013 school census, 3% of survey respondents reported dropping out due to illness or having to take care of someone ill (UNESCO, 2014). Even school absence as short as 10 days has been shown to have long-term consequences (Cattan et al., 2023) indicating that even if health is only temporarily affected, losses from learning may be permanent. Lastly, even when attending school, learning can be limited: In a survey performed by the World Bank, only 40% of students in year 4 (mainly ten-year-olds) could perform a year 3 mathematics task such as $6 \div 3$ or 7×8. An important cause of this is low quality of educational resources (Mgema, 2022; Ilomo and Mlavi, 2016), fatigue or frequent absences due to waterborne diseases are likely to also negatively affect the capacity for learning.

[5]This information and subsequent on Tanzania school system from UNESCO (2014) and World Bank (2016)

3 Data

3.1 Geographic data

To generate our treatment measure, waterborne disease potential, we construct a high-resolution hydrological model to simulate the flow of surface water across all of Tanzania. We aim to generate a treatment measure with the highest possible geographic resolution that is feasible for simulation at this scale, using common input data components that are necessary for hydrological (flow) and hydraulic (distribution) simulations. We therefore make use of the latest available global gridded datasets at the highest possible resolution, shown in Figure 1. For topographic data, which underlies the hydraulic flow component of the model, we use data from the Shuttle Radar Topography Mission (Farr et al., 2007). Rainfall and evaporation data, which allows water to enter and leave the model, are extracted from ERA5 reanalysis data (Hersbach et al., 2018). Finally, to model soil infiltration processes, we use the ISRIC 2.0 global soil database (Poggio et al., 2021), which contains high-resolution data on soil composition across the globe. Appendix A.1 contains more detailed descriptions of each data source and how they are used to generate our treatment measure.

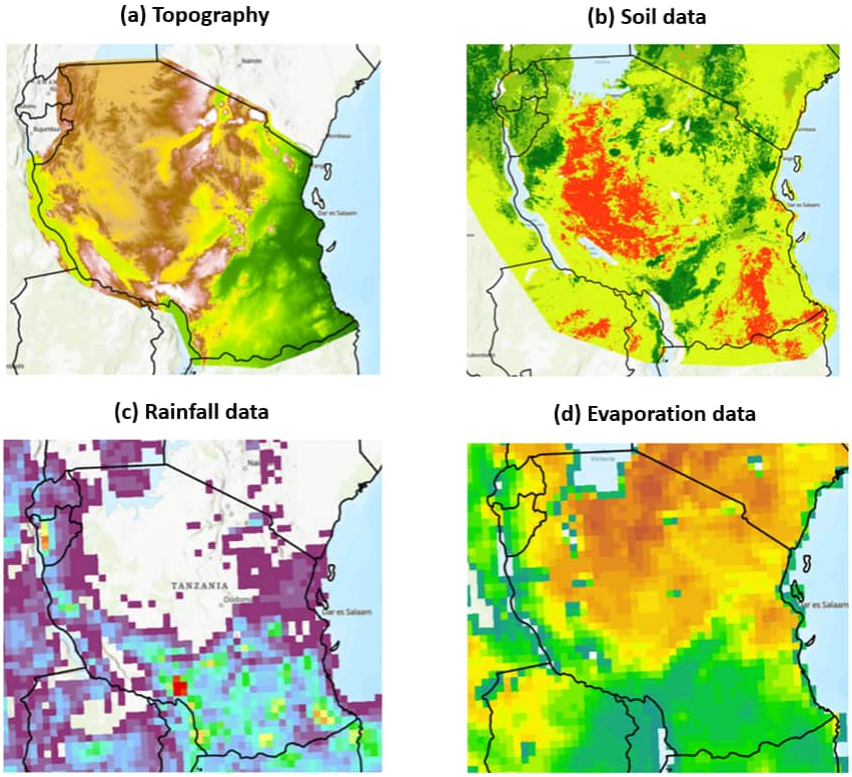

Figure 1
Data sources used to construct the WBD Potential measure

Note: The algorithm for WBD Potential uses four high-resolution gridded geographic datasets: (a) topographic data from SRTM, (b) soil infiltration data from the ISRIC 2.0 soil database, and (c) hourly rainfall and (d) potential evapotranspiration from ERA-5 reanalysis data.

In order to investigate treatment heterogeneity in terms of local climate, we use historical annual mean precipitation and temperature from the WorldClim v2 database (Fick and Hijmans, 2017),

covering the period 1970-2000.

3.2 Uwezo surveys

The Uwezo surveys are large-scale assessments of school-age children in Kenya, Uganda and Tanzania. The surveys are administered to schools, communities and households, and are repeated cross-sections. We utilise surveys from the survey waves in the years 2011, 2013, 2014, 2015 and 2017, and all survey waves are representative at the district level. In our paper, we employ the household surveys.[6] For each household, there is collected information on household wealth and assets, including whether the household has a toilet, as well as basic socio-economic information about the child's age, year of schooling, whether they are currently enrolled, mother's age and mother's schooling.

There are three features of the Uwezo surveys that are crucial for our purpose of studying the effect of waterborne diseases on learning. First, all school-aged children (aged 6-16) in the surveyed household are tested in basic Mathematics, English and Kiswahili by being administered a random sample of question cards in each subject. This allows us to measure learning and cognitive ability, which our main income of interest. Another advantage is that test scores are collected for all children of school-age in the household which makes us able to capture the effects also on those children who drop out of school due to e.g. ill health. Lastly, the surveys are administered by TWAWEZA, an NGO which works largely independent of Tanzanian government/authority, and the tests have no bearing for the school grades of the children or the evaluation of teacher performance. Hence, there is little reason to believe that parents or schools would influence performance during the tests.

Depending on how well the child answers the question sets, the children get allocated a score: The scores are discrete and are 1-5 for reading in English and Kiswahili, and 1-6 for Mathematics, except in wave 2015 where scores are given between 1 and 9. For example, scores in reading are based on whether the child can recognise letters, recognise words, read a paragraph or read a short story[7]. We have no reason to think that the effects would be different for verbal and quantitative abilities, and hence we construct our main outcome as the mean age-standardised test score across the three subjects. We standardise each child's test score by wave to have mean zero and standard deviation one, and then take the average and standardise again such that

$$\text{Mean standardised test scores} = \frac{\textit{Avg Score} - \text{Mean}(\textit{Avg Score})}{\text{SD}(\textit{Avg Score})} \tag{1}$$

where *Avg Score* is the average for each child across the three subjects.[8] We thus end up with

[6] For more information, visit https://uwezotanzania.or.tz/

[7] In Mathematics, scores are based on whether the children can count, recognize numbers, order numbers, add, subtract or multiply.

[8] That is, for each child and $\textit{Subject} \in \{\textit{English, Kiswahili, Maths}\}$ we standardise each score by wave:

$$\text{Subject}_{\text{std}} = \frac{\text{Score} - \text{Mean}(\text{Score})}{\text{SD}(\text{Score})}$$

a continuous measure of children's current performance and learning which is comparable across waves. In our results, unless otherwise stated we refer to the mean standardised test score as "test scores". Test scores are missing for 8% of our sample which we exclude from our analysis. If these differ systematically from the non-missing values this could bias our result. For example, if the test scores are missing because children did not get tested as they were ill, we may underestimate the effect of WBD Potential on test scores.

The second important feature of the Uwezo surveys is that the exact date of the survey, and thus testing of children is recorded. This allows us to link recent changes in stagnant water change to the test scores of the children. The third important feature of the Uwezo feature is that it contains accurate geographic information which is time-invariant. Since we model WBD Potential based on climactic and geographic factors, we require geographical information on households. While data are collected at the household level, the smallest geographical unit which is geocoded is *wards*, the smallest unit of geographical administrative area in Tanzania.

3.3 Demographic and Health Surveys (DHS)

To explore health and nutrition-related outcomes linked to waterborne diseases we use three surveys of the Demographic and Health Surveys (DHS). We use data from three available waves of DHS; 1999, 2010 and 2015[9].

Similar to the Uwezo surveys DHS are repeated cross-sectional surveys representative at the subnational level, but the DHS surveys differ from the Uwezo surveys in two regards that are important for our analysis. First, children surveyed are between 0-5 years and so more sensitive to waterborne diseases than the children in the Uwezo surveys (6-16 years). Assuming that the risk of "catching" waterborne disease reduces by age, we can interpret these results as upper-bound estimates for children in school-going age. Additionally, since the DHS surveys mostly precede the Uwezo surveys by a few years, there is potentially large overlap in terms of the age cohorts covered by the surveys.

The second key difference to Uwezo is that villages (or neighbourhoods in urban areas) are geocoded with a displacement. Thus, we slightly have to alter how we define the allocation of treatment. Each village or neighborhood, defined as a 'cluster', has its coordinate randomly shifted by 0-5 km from the sampling location and may not be visited more than once, which means that locations will necessarily vary between the survey waves, unlike the wards used in our baseline specification. To account for unobserved geographic variation, we allocate each household to a 50 km grid cell which is fixed over time. The cells are large enough that they capture the variation in

where mean and standard deviation, SD, are wave-specific. We then take the mean of the standardised subject-specific variables and standardise, again by wave, to construct *Avg Score*:

$$\text{Avg Score} = \frac{English_{std} + Kiswahili_{std} + Maths_{std}}{3}$$

[9]The DHS Program is funded by USAID and has collected survey data with a particular focus on children and women's health since 1984. For more information, visit https://dhsprogram.com/

the share of stagnant water over time, but arguably small enough to control for most unobserved geographical variation. Figure C.1 in the Appendix shows the geographical distribution of DHS clusters and the 50 km grid cells.

4 Empirical strategy

4.1 Algorithm to model waterborne disease potential

We are interested in estimating the effect waterborne diseases on children's learning. Since disease incidence likely correlates with several local conditions, in order to extract a causal estimate of waterborne diseases we require exogenous variation in the probability of contracting waterborne diseases. A naïve approach would be to simply use variation in rainfall to proxy for this risk, as extreme rainfall events are associated with disease outbreak (Levy et al., 2016; Lo Iacono et al., 2017). However, local rainfall in itself is an insufficient predictor of the risk of waterborne diseases: Rainfall contributes to flooding and the formation of pools of stagnant water, but also non-local rainfall from upstream can form significant stagnant water pools which equally enables a waterborne transmission pathway, usually found to be the most significant source of transmission (Eisenberg et al., 2013; Leclerc et al., 2002). To improve upon the rainfall measure, we use well-established hydrological modelling methods and novel high-resolution geospatial datasets to simulate the overland flow of water and, importantly, the formation of stagnant water pools. Since growth of waterborne pathogens occur mostly in stagnant water, and we want to use variation in this measure as our treatment, we subtract permanent water bodies and flowing water from the simulation output.

The modelled waterborne disease potential (WBD Potential) is thus measured at the ward level, following the geographic unit given in the Uwezo surveys. We define WBD Potential for each ward and survey year, $S_{w,y} \in [0,1]$, as the time average of the share of the ward's area A_w covered by stagnant water $A_{S,t}$ at day t over the past n days from the date of the visit:

$$S_{w,y} = \frac{\sum_{t=1}^{n} \frac{A_{S,t}}{A_w}}{n} \tag{2}$$

In our main specification we average over the last **two weeks** prior to the date of the survey. We choose two weeks mainly for two reasons. First, we prefer a short time span since waterborne pathogens can grow and infect humans within days and even hours in their ideal conditions[10]. Defining treatment across two weeks also allows us to better account for the additional transmission time of the disease via the faecal-oral channel. Second, there is up to a week of variation in the date of the survey. Thus, extending the definition of treatment to two weeks allow us to also capture households surveyed later within the ward.[11]

Figure 2 shows the output of the hydrological simulation for specific point in time for northern Tanzania, and how this is translated into our waterborne disease potential measure. A more intense

[10] See e.g. Appendix B.2.

[11] In robustness checks in Section 6 we vary the number of weeks included in $S_{w,y}$.

purple colour implies a higher value of $S_{w,t}$. In addition to a more accurate measurement of our variable of interest, stagnant water, our methodology is likely more robust to the potential issues that arise when using rainfall as an instrument (Mellon, 2021). Rainfall may affect our outcome variables through many different channels, as a large and growing literature using exogenous variation in rainfall to estimate effects on children's health and education outcomes has shown (Maccini and Yang, 2009; Shah and Steinberg, 2017; Bandyopadhyay et al., 2012). By instead using rainfall to simulate stagnant water over time, we reduce both measurement error as well as the possible number of ways in which the exclusion restriction can be violated.

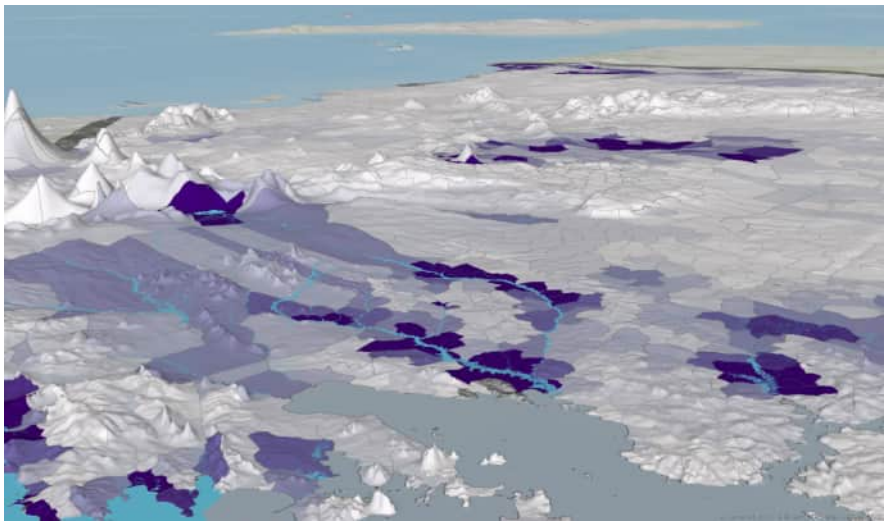

Figure 2
Example of one-time step in the definition of waterborne disease potential

Note: Areas covered by stagnant water, capturing WBD Potential, according to our hydrological simulation model are shown in light blue. The resulting share of stagnant water in each ward, $S_{w,t}$ is shown in shades of purple, where a darker colour indicates greater exposure to stagnant water.

While our paper is not the first to use flood data, it is unique in that we are able to create a highly disaggregated dynamic panel of areas under water at the ward level, by using variation over time and generating continuous time series of areas below water. We follow the hydrological engineering literature on large-scale flood simulations and numerically solve a simplified set of the differential equations that characterise water flow (Dazzi et al., 2021; Falter et al., 2013, 2016), with a custom-made integrated hydrological and hydraulic model built in the Python language. Similar to the literature on this methodology, we are less interested in resolving the detailed small-scale dynamics of water flow, and more so in the resulting large-scale allocation of water mass over time.

As input to the model, we use the global gridded geographic datasets described in Section 3.1 and Appendix A.1. Simulations are run at a spatial resolution of 90 m, which is orders of magnitude smaller than the ward level, while being feasible enough to enable simulation over the time scale of months at the country level[12]. For more details of the computational scheme, Appendix A.2

[12]The runtime of hydrodynamic simulations are very sensitive to the level of resolution and typically with the *cube* of the spatial resolution. However, a recent evaluation of large-scale 2D hydrodynamic simulations for several European rivers show that resulting flooded area and water level are insensitive to variations in spatial resolution, once it is finer than 100 m (Dazzi et al., 2021; Falter et al., 2013, 2016). This is important as it allows us to feasibly simulate the time evolution of stagnant water across a whole country such as Tanzania over time.

provides a schematic of the input data and algorithm used to model stagnant water.

Lastly, as a robustness and validation exercise, we construct an alternative treatment measure using an external hydrological model coupled to coarse occurrence data on surface water from satellite data. We find a relatively strong correlation between the two measures for urban areas, but weaker for rural areas, which could be due to a number of reasons, see Appendix A.3 for a discussion of this. In addition to capturing contemporaneous effects on children's learning, we are also interested in the effect on learning from the *accumulation* of this exposure over time. To this end, we use the external hydrological model and satellite data to generate a coarse measure of the total number of short-term shocks in the past n years of each observed child. See Appendix A.4 for more details on how this measure is constructed.

4.2 DiD model specification

We estimate the following difference-in-differences model for outcome Y of individual i residing in ward w, surveyed in calendar month m in the year y:

$$Y_{imwy} = \alpha_w^1 + \alpha_y^2 + \alpha_{wy}^3 + \delta S_{wy} + \gamma R_{wy} + X'_{iwy}\beta + \varepsilon_{iwy} \tag{3}$$

where δ is the main coefficient of interest: $S_{wy} \sim (0,1)$ is the average share of stagnant water in each ward in the two weeks prior to the date of the survey. Hence, δ captures the causal effect of an individual i being exposed to waterborne diseases in their ward w in year y. Since the WBD Potential share is given between 0 and 1, the value of the coefficient presented in results below is the value for the ward if the stagnant water share would cover 100% of the ward area – an unrealistic scenario. Thus, for interpretation of the results we refer to the coefficent scaled down by ten, reflecting if the WBD Potential is equal to 10%.

Next, as the share of stagnant water over time is potentially affected by local rainfall, and local rainfall may also affect the outcome, we include the last two weeks of rainfall at the ward level for each wave, R_{wy}, in our baseline specification. From the Uwezo surveys, we have survey waves from 2011, 2013, 2014, 2015 and 2017. We control for time-invariant unobserved differences across wards with ward fixed effects, α_w^1, as well as survey-year fixed effects α_y^2 (equivalent to survey wave) and calendar month fixed effects α_m^3.

Lastly, we also include a vector of controls in X on the individual level (gender, age, whether the mother had secondary or higher education, the mother's age, and an index for household wealth). We impute missing values within these variables as the sample mean. In the subsequent analysis we also distinguish between *rainy* and *dry* wards defined as the 20-year average yearly precipitation being above or below 1000 mm (which roughly corresponds to the mean in our sample). We do this because we expect that rainy wards are more likely to have a higher baseline of waterborne pathogens, as these are more likely to survive in more humid areas and resulting in more cases of waterborne diseases.

To explore potential heterogeneities we also interact the variable of interest, S_{wy} with an indicator for different facilities and habits regarding water, sanitation and hygiene (WASH) (Section 5.2).

Formally, we estimate

$$Y_{iwy} = \alpha_w + \alpha_{dy} + \delta_1 S_{wy} + \delta_2 WASH_{iwy} + \delta_3 S_{wy} \times WASH_{iwy} + \gamma R_{wy} + X'_{iwy}\beta + \varepsilon_{iwy} \quad (4)$$

where WASH is a binary indicator of sanitation status, such as whether the household has a toilet. Here, we are interested in δ_1, δ_2 and δ_3. The parameter δ_1 is the effect of WBD Potential on our outcome when WASH=0, δ_2 is the direct effect on Y_{iwy} from WASH. The key contribution of this model is δ_3, which is the coefficient on the interaction term. This estimate will give us the contribution to the outcome of the WASH variable through the effect of WBD Potential. We will carefully discuss potential sources of endogeneity in this interaction variable, but we nonetheless believe these results provide important, if sometimes only causally suggestive, evidence.

4.3 Descriptive statistics

Table 1 provides summary statistics of the Uwezo surveys and features of wards. Panel A summarises household and child characteristics. Test scores are standardised with mean zero and standard deviation one by wave, thus the total sample has a slightly lower mean. Children are on average 11.1 years old, and 46% are girls. Households are relatively large with seven people on average. Mothers have a similar education to other national surveys (see Section 2.3) where 24% have at least a secondary education. We also provide a wealth index based on normalised principal components of household assets, which is normalised to have mean zero and standard deviation one, by each wave. The components in the wealth index are the type of wall, whether the household has a radio, television, bicycle, motorbike, cattle, or electricity. Next, we see that 73% of households have any type of toilet. We also separate two sources of water: From nature and from tap, as there is evidence of either channel affecting waterborne disease risk[13] which could make this an important source of heterogeneity.

In panel B we provide ward-level characteristics on the 3876 wards. Overall, WBD Potential has a strong mode at zero, where the average level is 1.3%, but there is a large variation (see also Figure C.2 in Appendix C for graphical representation of distribution). More than 80% of the wards are rural. In our sample, a ward includes approximately 20 households sampled per ward, but there is a large variation in both the number of households and villages sampled. It should be noted that precipitation, elevation, soil suitability and temperature given in Table 1 are long-run means at the ward-level for the period 1970-2000 (Fick and Hijmans, 2017), and this variation will be absorbed by our within-ward analysis.

To get an overview of the spatial distribution of our treatment measure we also present Figure 3. While this is only representative of the time periods covering the Uwezo surveys, the distribution of treatment intensity compares favorably to the historical spatial distribution of cholera incidence, which has shown a higher incidence in the Lake Victoria region, southeast, and northwest part of

[13] Remaining households get their water from a public source, well, bottles, and tanks.

Table 1
Summary statistics, wards and Uwezo households

	Mean	SD	Min	Max
Household characteristics				
Test score (std)	-0.020	0.997	-4.481	4.239
Woman	0.46	0.499	0	1
Age	11.1	2.792	6	16
Mother's Age	36.3	5.903	16	60
Mother's Edu $\geq$ Sec	0.24	0.408	0	1
Wealth (index)	-0.0064	1.584	-2.230	24.58
Persons in household	7.03	3.363	1	78
Children in household	3.32	1.873	1	20
Household has toilet	0.73	0.444	0	1
Water from nature	0.13	0.340	0	1
Water from tap	0.43	0.495	0	1
Observations	386005			
B. Ward characteristics				
Num. wards	3876	0	3876	3876
WBDP (%)	0.013	0.0360	0	0.812
Rural ward	0.84	0.365	0	1
WBDP $>$ 5%	0.073	0.260	0	1
Num. households	19.7	16.99	0	297
Num. villages	1.72	1.173	0	29
Ward area (sq. km)	253.4	601.8	0.111	11437.0
Precipitation (mm/yr)	1016.0	277.1	428.7	2459.1
Elevation (mean metres)	1043.4	540.1	8	3105.4
Soil suitability (%)	0.71	0.218	0.0604	1.002
Temperature (C)	22.1	2.682	10.60	27.09
Observations	8841			

Note: Summary statistics of mean, standard deviation, minimum and maximum of each variable. Panel A displays statistics across individuals, while Panel B displays statistics at the ward level, by wave for the wave-variant characteristics.

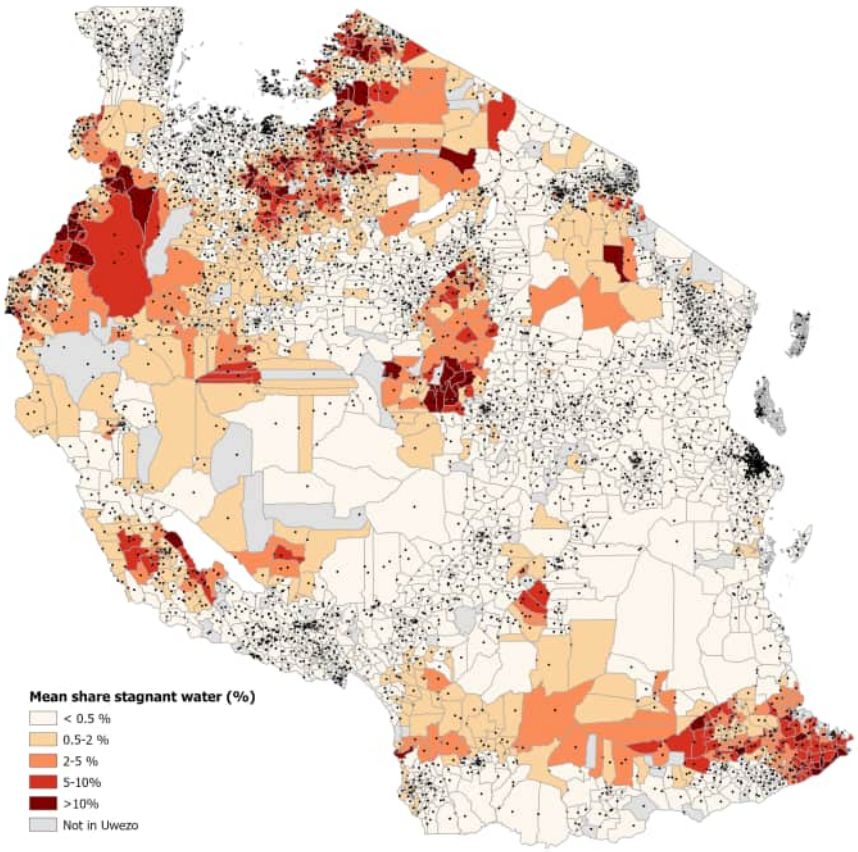

Figure 3
Spatial distribution of WBD Potential

Note: This map shows the geographic distribution of wards in Tanzania, together with population density (each dot represents 10,000 people) and the mean share of stagnant water, our main treatment measure, in each ward in our sample period. Note that the mean share of stagnant water is only representative of the specific months and years represented in our Uwezo survey waves, which is the period for which we ran our hydrological model, and is not indicative of the long-term mean values for Tanzania.

the country, where we also find greater likelihood of stagnant water forming (Nkoko et al., 2011). There is a large variation in the ward area, which correlates negatively with population density. We also see that WBD Potential (measured as the average over waves) tend to concentrate in smaller wards.[14]

5 Results

5.1 Effects of waterborne disease potential on test scores

We first present our main results, the effect of WBD Potential on test scores. As discussed in Section 4.2, in our main specification WBD Potential is the share of ward area covered by stagnant water, given as the average in the two weeks preceding the survey and the test of children. Test scores are standardised averages of the child's score in English, Maths and Swahili.

Table 2 displays the estimates of the effect of WBD Potential on test scores. We successively

[14] See Appendix C for a longer discussion.

Table 2
Effect of WBD Potential on test scores

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent: Test score (std)					
WBD potential	-1.390***	-1.254***	-0.934***	-0.660**	-0.647**	-0.742**
	(0.280)	(0.221)	(0.334)	(0.324)	(0.319)	(0.315)
Obs.	368,446	368,446	368,444	368,444	368,444	368,444
Clusters	3,844	3,844	3,842	3,842	3,842	3,842
Covs		✓				✓
Ward FE			✓	✓	✓	✓
Wave FE				✓	✓	✓
Month FE					✓	✓

Note: Results is the effect of WBD Potential on test scores. Standard errors in parentheses clustered on ward. WBD Potential is the average share of ward covered in stagnant water in the two weeks preceding the date of the survey. Covariates include child's gender and age, and mother's age and secondary education, a wealth index, local past two weeks of ward precipitation. Ward is the treatment unit of observation. Wave is the year of the survey and waves included are 2011, 2013, 2014, 2015 and 2017. Month is the calendar month of treatment.

add more controls and fixed effects to an initial parsimonious specification. Throughout columns 1-6, all estimates are negative and statistically significant. In column 1, we report a pooled regression with no controls. Here, the coefficient implies when WBD Potential is 10%, children's test scores are reduced by 0.14 standard deviations. In column 2 we add household-level controls, and then through columns 3-5 add ward, wave and calendar month fixed effects. Both ward and wave fixed effects reduce the coefficient size substantially, although calendar month does not affect the coefficient much, suggesting the estimate is not very sensitive to seasonal variation. Column 6 is our preferred specification: We include household and individual covariates, ward fixed effects, wave fixed effects and month fixed effects. The estimate implies that if WBD Potential is 10%, student test scores are reduced by 0.074 standard deviations, which is equivalent to one standard deviation increase in WBD Potential reducing test scores by 0.028 standard deviations. Compared to the observational difference in column 1, this effect is reduced by almost half. This suggests there are important time-invariant or ward-invariant unobservable characteristics which influences the way in which WBD Potential affects test scores.

We hypothesise that our results reflect a higher incidence of waterborne diseases among tested children, which affects their capacity to learn. However, we cannot here distinguish between whether the child performs worse because they have been absent from school, or they are ill or just recovering from illness when they take the test and thus performing worse than they would otherwise. In Section 6 we further explore how sensitive our results are to the time dimension of the definition of WBD Potential, and in Section 7 we provide further evidence that the effect on test scores is through health and provide some empirical support for a discussion on potential long-run effects.

For now, to further motivate how we see the connection between WBD Potential and character-

istics of waterborne disease contagion, in Table D.1 in Appendix D.1 we further divide the result by the long-run precipitation (measured as the mean annual precipitation in the years 1970-2000) of each ward, motivated by the fact that overall wetter climates are both more susceptible to waterborne disease, but also more familiar with outbreaks (and potentially how to prevent them) (Cann et al., 2013). We find that the main effect is driven mainly by dry wards, where the effect is larger and statistically significant (0.81 standard deviations). For rainy wards, the coefficient (-0.0005) is not statistically different from zero but imprecisely estimated.

One of our contributions is the ability to capture the effect of waterborne disease potential in the areas where stagnant water settles, as opposed to modelling the effect on outcomes from local precipitation (as studied in e.g. Shah and Steinberg, 2017; Maccini and Yang, 2009; Ide et al., 2021). In our main estimation we define and include a variable *Local Precipitation* which is equal to the sum (in cm) of rain in the two weeks preceding the survey date.[15] However, local precipitation may drive the estimated effect nonetheless. To address this concern we next run our main specification both with and without WBD Potential and local precipitation to estimate how both variables affect the children's test scores.

Table 3 summarises the result of this exercise. Panel A displays the correlation between precipitation and WBD Potential. While the coefficient is statistically significant, the contribution to WBD Potential is small: 1 cm of rainfall contributes 0.001 to the share of WBD Potential; approximately one per cent. This is reassuring, as we have intended to model the emergence of stagnant water pools as a separate phenomenon from precipitation, and we expect a large part of the formation of water pools to originate from non-local precipitation. Panel B and C run WBD Potential and precipitation as separate explanatory variables on test scores, and the last panel repeat our main specification from column 6 in Table 2. "Local Precipitation" as an explanatory variable is statistically significant for rainy wards, and including it only marginally changes the coefficient of the effect of WBD Potential on test scores. Thus, we are reassured that our results are driven by WBD Potential and not local rainfall, however, we include it as a control variable throughout.

5.2 Heterogeneity analysis: Mitigation with water and sanitation practices

Existing literature both from the economic and microbiological field emphasise the importance of the faecal-oral route in spreading waterborne diseases, and how improved water, sanitation and hygiene (WASH) practices mitigate this effect. To better understand to what extent the effect is driven by WASH practices we next run an interaction of our WBD Potential measure with whether the household has a toilet (as described by Equation (4)). These results are presented in Panel A in Table 4.

For the full sample, we estimate a negative effect: In households without a toilet, when WBD Potential is 10% it decreases test scores by 0.15 standard deviations and the effect is highly statistically significant. The effects are similar for dry and rainy wards, although more precisely estimated

[15] The unit of this measure reflects that the mean precipitation is approximately 0.5 cm, and that we use two weeks to mirror the two weeks we include in the measure of WBD Potential.

Table 3
Correlation between local precipitation and WBD Potential

	(1)	(2)	(3)
	All	Dry	Rainy
	Panel A. Dependent: WBD Potential		
Local precipitation (cm)	0.00102**	0.00355**	-0.000499*
	(0.000413)	(0.00172)	(0.000261)
Mean precip (cm/2 weeks)	0.44	0.34	0.53
Obs.	7,240	3,648	3,588
Clusters	2,558	1,319	1,238
	Panel B. Dependent: Test scores		
WBD potential	-0.716**	-0.831**	-0.0209
	(0.314)	(0.348)	(0.734)
Obs.	368,444	178,449	189,995
Clusters	3,842	1,669	2,173
	Panel C. Dependent: Test scores		
Local precipitation (cm)	0.0310***	-0.0357	0.0401***
	(0.0118)	(0.0234)	(0.0135)
Obs.	368,444	178,449	189,995
Clusters	3,842	1,669	2,173
	Panel D. Dependent: Test scores		
WBD potential	-0.742**	-0.812**	-0.00542
	(0.315)	(0.349)	(0.729)
Local precipitation (cm)	0.00318***	-0.00334	0.00401***
	(0.00117)	(0.00235)	(0.00135)
Obs.	368,444	178,449	189,995
Clusters	3,842	1,669	2,173

Note: Standard errors parentheses clustered on ward. Dry ward if mean annual precipitation $<$ 1000 mm, rainy ward if $\geq$ 1000 mm. District*Wave, Calendar month and Ward fixed effects included in all estimations.WBD Potential is the two-week average share of area of ward covered in stagnant water, $\sim$(0,1). 'Local precipitation' is the sum of precipitation in mm the past two weeks, by ward and wave. Panel A runs only rain on the share of test scores to provide a correlation measure between the two variables. Panel B displays the effect of stagnant water on test scores, and panel C the effect of two-week precipitation on test scores. For completion, Panel D mirrors the main specification. All estimations include individual covariates: Child's gender and age, and mother's age and whether she has secondary education or above, and an index for household wealth.

for dry wards. There is a positive effect on test scores of having a toilet, which is consistent with wealthier households having a toilet also being able to invest more in the education of children. The interaction term is the effect of WBD Potential for households with toilets. Only for rainy wards do we find a precisely estimated effect: The coefficient is positive and two thirds that of the negative effect for households with no toilet. This is consistent with epidemiological evidence that WASH practices are particularly important in rainy areas where pathogens proliferate more freely. We interpret this as evidence showing that good sanitation facilities can at least partially offset the higher risk of contracting waterborne diseases when WBD Potential is high.

However, the toilet of the *household* may be a poor indicator to get a full picture of the importance of WASH practices. This is in part because there are large spillovers between your house and your neighbour's house in the spread of waterborne diseases (Magana-Arachchi and Wanigatunge, 2020; Duflo et al., 2015; Kremer et al., 2022), and also because having a toilet is endogenous to the household's preferences or capacity which may affect children's learning in other ways. To address these points, for each household we construct an average of the share of households with a toilet in each village, leaving out the toilet status of the household itself from the mean. We then perform the same estimation and summarise these results in Panel B in Table 4.

We estimate for households in villages with no toilet a similar effect of WBD Potential on test scores as for households with no toilet. However, when splitting by rainy dry and rainy wards, a new pattern emerges: For households living in villages with low number of village toilets implies 10% WBD Potential lowers test scores by 0.54 standard deviations. This effect is diminished by households living in villages with a higher share of toilets: The coefficient WBD Potential (10%) offset the negative effect of test scores by 0.46 standard deviations relative to villages with no sanitation if everyone in the village has a toilet (i.e. share=1), suggesting large marginal returns the more households in a village have a toilet. In contrast, the interaction effect in dry wards is an order of magnitude smaller and statistically insignificant, while the direct effect of the village toilet share suggests if everyone in the village has a toilet, test scores are 0.29 standard deviations higher.

While your neighbours' choice of having a toilet is plausibly more exogenous than your own choice of a toilet, there are still important sources of endogeneity which could explain our results. For instance, an individual might buy a toilet because their neighbour has one, and both the village level and household level presence of toilets likely correlates heavily with wealth, so the channel through which the toilet acts through WBD Potential could be due to e.g. a higher general degree of economic development. While we cannot test this directly, one way to examine this wealth channel is to perform the same estimation with household wealth. These results are summarised in Panel C. As expected, children from wealthy households have higher test scores. In contrast to the results with toilets, the interaction between WBD Potential and household wealth is small and statistically insignificant for both dry and rainy wards, suggesting wealth does not directly affect how WBD Potential causes a reduction in test scores. Additionally, the direct effect of household wealth on test scores is similar across both types of wards. This suggests the wealth channel does not depend on whether the ward has a drier or wetter climate, which is in stark contrast to estimations of sanitation. Thus, wealth does not seem to explain the majority of the effect estimated from toilets

Table 4
WBD Potential and test scores, mitigation with sanitation

	Dependent: Test score (std)		
	All	Dry wards	Rainy wards
	Panel A: Household toilet		
WBD Potential	-1.474***	-1.399***	-1.528*
	(0.431)	(0.460)	(0.850)
HH toilet	0.105***	0.122***	0.0830***
	(0.00906)	(0.0127)	(0.0127)
WBDP*HH Toilet	0.235	-0.0632	1.092***
	(0.194)	(0.221)	(0.348)
Obs.	217,766	117,266	100,500
Clusters	3,090	1,594	1,496
	Panel B: Village toilets		
WBD Potential	-1.403*	-0.713	-5.393***
	(0.745)	(0.749)	(1.917)
Vill toilet	0.218***	0.288***	0.0778
	(0.0404)	(0.0517)	(0.0642)
WBDP*Vill Toilet	0.230	-0.665	4.555**
	(0.747)	(0.738)	(1.952)
Share has toilet	0.80	0.75	0.85
Obs.	217,766	117,266	100,500
Clusters	3,090	1,594	1,496
	Panel C: Household wealth		
WBD Potential	-1.465**	-1.851***	-0.320
	(0.598)	(0.679)	(1.930)
HH wealth	0.0872***	0.0832***	0.0903***
	(0.00263)	(0.00407)	(0.00336)
WBDP*HH Wealth	-0.0383	-0.0293	0.0148
	(0.0688)	(0.0742)	(0.146)
Obs.	368,444	178,449	189,995
Clusters	3,842	1,669	2,173

Note: Standard errors parentheses clustered on ward. WBD Potential is the two-week average share of area of the ward covered in stagnant water, $\sim$(0,1). HH toilet is an indicator of whether household has a toilet. Vill toilet is the share of households within the household's village that has a toilet (excluding theirs). HH Wealth is an index combining household assets, excluding toilets. Dry ward if long run precipitation $<$1000 mm per year on average, Rainy ward if $\geq$ 1000 mm. Wave, Calendar month and Ward fixed effects, Past two-week sum of local precipitation included in all estimations.

and these results are indicative of the importance of good WASH practices to stem the costs of waterborne diseases.

Finally, another important determinant of the spread of waterborne diseases is the quality of water in the household. While we lack detailed data on water quality, we know the main source of water in households. In Table D.2 in Appendix D we estimate the same specification as above, now with an indicator for the household water source interacted with WBD Potential. We find that when WBD Potential is 10% households *not* getting their water from nature have a 0.16 reduction in test scores, but households who get their water from nature have relatively higher test scores (0.8 standard deviations), which is indicative of water from nature being less associated with contracting waterborne diseases. For households, we instead find that the effect of WBD Potential is relatively stronger and more negative for households with tap water, although these results are not statistically significant. These findings are consistent with studies which show that the growth of waterborne disease pathogens often originates from poor water management (in cities) rather than a direct natural source (e.g. Knutsson, 2020; Cutler and Miller, 2005; Ling et al., 2018).

5.3 Additional heterogeneity Analysis

To better understand the effect we estimate, we next conduct heterogeneity analysis both by key individual characteristics and by type of ward.

Child age and gender: Existing literature shows that flooding affect children's outcomes differently for boys and girls (Maccini and Yang, 2009). As summarised by Figure D.1 we estimate negative effects overall and the coefficients of WBD Potential on test scores for boys and girls are never close to statistically significantly different from one another. Regarding age, we find a U-shaped pattern of the estimated effects of WBD Potential on test scores with respect to the child's age, where the main negative effect is driven by 9-14 year-olds. These results are summarised in Figure D.2 While we would suspect that younger children are more vulnerable to waterborne disease, there is also less variation in test scores for children in the ages 7-8 since most of the observations in these age groups tend to cluster around the lowest level of capabilities. However, at ages 9 and older more children learn higher-level skills such as multiplication and division which may enable us to better observe how children fall behind relative to their peers due to exposure to waterborne disease.

Urban and rural dimension: We next analyse if there are differences between urban and rural areas since most existing literature find that there are more severe disease effects in urban areas that are densely populated (Alsan and Goldin, 2019; Troeger et al., 2018). We summarise our results in Figure D.3, where we separately estimate the effect of WBD Potential on test scores for rural and urban wards. Within each subplot, we also split the estimate by dry and rainy areas. The effect on test scores is more precisely estimated for rural wards, but the estimates are consistently negative and not statistically different from one another.

Distance to coast: Relatedly, we also analyse if there is important variation across wards by distance to the closest coast, whether it is the sea or Lake Victoria, as evidenced by documented

disease outbreaks. We divide wards into quartiles ranging from 0-250 km, 250-525 km, 525-711 km and 711-1022 km away from the coast. The results are summarised by Figure D.4 in Appendix D.1. We find that all estimated coefficients are negative, but the effect of WBD Potential on test scores appears driven by wards close to the coast, where the effect is the largest, while the effect of WBD Potential is the most imprecisely estimated for areas far from the coast. This is consistent with empirical evidence, showing that for instance cholera is more common in coastal than inland regions across Sub-Saharan Africa (Rebaudet et al., 2013), and especially in Tanzania (Lugomela et al., 2014). We find these results plausible for two reasons: First, since coastal regions are by definition furthest downstream, stagnant water in these regions is likely more contaminated than water further inland since it will have passed through more potential sources of contamination before stagnating. Second, coastal regions in Tanzania tend to be more flat and conducive to exposure to stagnant water than inland areas, which are generally more rugged. Third, wards close to the coast are typically more humid and also more densely populated, and thus more conducive to harbouring waterborne pathogens.

6 Robustness checks

Non-linearities in treatment: In our main specification we estimate the effect of a continuous share of WBD Potential (which can take any value between zero and one) and estimate this on test scores. This assumes a linear relationship between WBD Potential and test scores. In this section, we explore how strong an assumption that imposes. We start by first redefining WBD Potential to be a binary treatment by creating a dummy which is equal to one for the wards in waves where WBD Potential is greater than 5%, a relatively rare event (see Table 1). Including this in our main specification instead of our continuous variable, we again estimate the effect on test scores. As summarised in Table D.3 in Appendix D.2 we find that treatment leads to -0.10 standard deviations lower thresholds.[16] This is larger but comparable to our main result of -0.7 standard deviations, consistent with the fact that 5% is a relatively severe or rare shock.

Next, we estimate the effect on our main WBD Potential measure but also include a squared term. As displayed in Table D.4, the linear term of the effect of WBD Potential on test scores increases from .74 standard deviations in our main specification to now being -1.33, and still statistically significant. The coefficient for the squared term, although it is statistically insignificant, is large and positive (1.52), which implies decreasing marginal effects of WBD Potential on test scores. While this may seem counter-intuitive, since more stagnant water would lead to greater probability of a disease outbreak which could potentially create positive feedback loops, attenuation happens mostly for treatment values near 1, indicating a fully water-covered ward, which never occurs in practice. One would also expect that, as the share of stagnant water grows, contaminated surface water becomes more diluted, thereby decreasing the probability of an outbreak. Several studies have shown that the probability of outbreaks of diarrhoea due to heavy rainfall is typically

[16] All results in this section can be found in Appendix D.2.

higher following a dry period than a wet period, which would generate a smaller but less contaminated amount of surface water, suggesting that rainfall dynamics affect not only the quantity but also the quality of surface water (Levy et al., 2016).

To further investigate how the intensity of our treatment affects children's test scores, we create additional dummies of WBD Potential which is equal to one for a value of two per cent or higher, and in the next estimation a value equal to one if WBD Potential is four per cent or higher. These results are presented in Figure D.5. We find that the effect is U-shaped relative to the per cent counted as treated, with the strongest effect on test scores being when stagnant water constitutes 8-14% of the ward area, although with stricter definitions of treatments the estimated effect becomes more imprecisely estimated, as expected since these are rare occasions (as implied by studying the distribution of WBD Potential in e.g. Figure C.2).

Comparison with GWS data: By hydrological standards we use a parsimonious algorithm to simulate the emergence and disappearance of stagnant water over time, both for transparency and computational reasons. However, by combining flood risk data on 30-by-30-metre cells in Tanzania from a combination of satellite data on surface water (GWS) and external hydrological model data we can generate an alternative treatment measure. These results and a comparison to our main specification are provided in Table D.5. In contrast to our simulations, no estimates using the GWS data are statistically significant, which could be due to several reasons (see discussion in Appendix A.3). Correspondingly, the correlation between the measures is typically positive but low, although where the correlation is the highest (.52), for urban wards, we estimate negative effects of WBD Potential on test scores.

Alternative estimator robust to heterogeneous treatment effects: As we rely on variation in treatment timing there is a danger that our estimates are biased due to heterogeneous treatment effects (Goodman-Bacon, 2021). In our setting, the likeliest source of such heterogeneity is that children exposed to higher waterborne disease potential are sicker and more vulnerable, putting them on a permanently more negative trajectory compared to non-treated children. A comparison against such wards would attenuate our effect to zero. To address this issue, we implement the estimator developed by de Chaisemartin and D'Haultfoeuille (2018) (denoted DCDH). From the potential other estimators we implement this estimator since our setting is non-staggered, in that wards can go from treated to untreated and back (de Chaisemartin and D'Haultfoeuille, 2022). Moreover, since the correct estimators for continuous treatments where treatment is non-staggered are not well-established as of yet, we redefine our treatment, WBD Potential, to a binary variable. This implies a ward becomes treated if the share of the stagnant water of the ward area exceeds 5%. Figure D.6 summarises these results. The estimate of the effect of WBD Potential on test scores with two-way fixed effects (TWFE) is -0.095 standard deviations while the estimate with the DHDC estimator is -0.104 and it is not statistically significant at the 95th level, although it is relatively precisely estimated. Since standard errors are bootstrapped for the DHDC estimator, we also provide a comparison to the TWFE with bootstrapped standard errors instead of clustered as in our main specification. The similarity in the estimates by estimators suggest heterogeneous treatment effects are not important confounders in our results.

Weeks included in the definition of WBD Potential: In our main specification we define WBD Potential to be the average share of stagnant water which covers the ward area over the past two weeks. In Figure D.7 we vary the number of weeks included in this average and re-estimate our main result, the effect of WBD Potential on test scores. We find that the coefficient of the treatment is remarkably stable across weeks included in the treatment, and statistically significant throughout. To better understand which weeks drive this result, we next compute a treatment variable which measures the stagnant water share by week, discretely. That is, in one estimation we only include the stagnant water share in the third week since the date of the survey, and in the next, we only include the stagnant water share in the fourth week since the date of the survey. This means there is likely a high serial correlation between these definitions of our treatment since the same pool of water can linger across weeks. The estimation results are summarised in Figure D.8. Here we see that the largest effect originates from the first week since the date of the survey, which is the only measure which is statistically significant. However, the coefficient is relatively stable and confidence intervals remain statistically significant up to the ninth week since the date of the survey, which is similar to the stability and pattern as seen in Figure D.7. Together, the results suggest that our estimated effect on test scores is the strongest for the weeks closer to the survey date but persists for the period observed. Importantly, we also see that the effect is measured also in the week immediately following the date of the survey, which is consistent with studies on outbreaks and the microbiological literature suggesting waterborne pathogens can contaminate and spread quickly.

Placebo checks across time: Given that the results show high autocorrelation over time, one concern is that our model captures a phenomenon in the future, which we then estimate effects for due to the correlation with past events. Such would be the case if past WBD Potential captures that children expect more favourable environmental conditions or events in the future in a way which affects test scores. If this is the case, *future* WBD Potential (in terms of the date children are surveyed) should better predict the change in test scores than past events. However, due to the high correlation (95%) between weeks, models including weekly WBD Potential in the same estimation are likely to induce problems associated with multicollinearity. Figure D.9 displays two attempts to disentangle the potential role of future WBD Potential while reducing the influence of multicollinearity. First, we estimate a model with our main WBD Potential (average two weeks pre survey) together with an indicator for the *change* in WBD Potential compared to the WBD Potential two weeks after the survey. In a second estimation we also aggregate the measure of WBD Potential in the 2-9 weeks pre survey, the two weeks that lie on either side of the week of the survey, and the 2-3 weeks after the survey. For both estimations, the estimates related to past events are both more negative and estimated with more precision, in line with our main specification.

7 Mechanisms: Different health channels

7.1 Health and Sanitation

Thus far we have not directly tested whether the channel we expect WBD Potential to act through is indeed waterborne diseases and the effects on health. To explore the health mechanism we make use of three waves (1999, 2010, 2015) of the Demographic and Health Surveys since the Uwezo surveys have no data on health.

We first explore how well the measure we have called WBD Potential captures waterborne disease incidence. While DHS does not test for waterborne diseases specifically, they measure important physical attributes of children, such as weight for height, and ask about recent health issues and symptoms of ill health. We make use of these variables to test our main hypothesis that WBD Potential affects waterborne disease incidence. We also conduct placebo checks as additional support for our hypothesis. In Table 5 we estimate our main DiD specification. In Panel A, each column represents a different outcome which we estimate on the full sample. The outcomes in the first two columns are outcomes most plausibly affected by waterborne diseases: Diarrhoea, which is the symptom most closely associated with waterborne disease, and weight for age (since severe diarrhoea tends to affect weight). The following four columns in Panel A are outcomes that should not be affected by the waterborne disease to the same degree: Fever, cough, anemia and height. Fever might e.g. be positive if our results are driven by increased malaria incidence, as fever is one of the main short-term symptoms of malaria. Cough is mainly associated with respiratory disease, while anemia and height are both long-run outcomes that should not be affected by WBD Potential since this measure is defined as the change in stagnant water in the past two weeks. All health outcomes are equal to one if the child has had the health issue at any point in the past two weeks, while weight, anemia and height are tested at the date of the survey.

We find a large and statistically significant effect on the probability that the child has had diarrhoea: A ten per cent WBD Potential increases the probability of the child having had diarrhoea recently by 2.75%. In other words, one standard deviation increase in WBD Potential increases the probability that a child has had diarrhoea the past two weeks by 1.4 percentage points, which is an increase of 11% relative to the mean diarrhoea incidence. The coefficient on weight is negative as expected but not statistically significant. The remaining outcomes that represent placebo checks are reassuringly estimated close to zero.

As further placebo checks, in Panel B of Table 5 we estimate the effect of WBD Potential on diarrhoea related to how likely it is that the child received contaminated water. In columns 1-2 we estimate the effect of WBD Potential on diarrhoea and split the sample by whether the child is breastfeeding. We hypothesise that children who are breastfeeding are less likely to have received food or water from other sources which should minimise the risk of contracting diarrhoea from waterborne diseases. In these estimations we restrict the sample to children 24 months and younger since children breastfeeding at an older age is uncommon. We estimate that for children who do *not* currently breastfeed, 10% of WBD Potential increases the probability of the child having had

Table 5
WBD Potential effect on health and water

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Health and physical outcomes					
	Waterborne diseases		Placebo			
	Diarrhoea	W.Age	Fever	Cough	Anemia	Height
WBDP	0.275**	-8.192	-0.0499	-0.0587	0.0722	27.35
	(0.113)	(6.244)	(0.140)	(0.144)	(0.0961)	(23.33)
Mean DV	0.13	88.49	0.22	0.21	0.40	92.12
Obs.	15,956	15,021	16,016	16,021	16,085	15,550
Clusters	242	242	242	242	242	242
	Panel B. Health and water					
	DV: Diarrhoea				DV: Fever	
	No Breastf	Breastfeeds	Not water	Water	Not water	Water
WBDP	0.506***	0.0844	-0.0822	0.300**	0.130	-0.0719
	(0.152)	(0.135)	(0.202)	(0.121)	(0.391)	(0.235)
Mean DV	0.11	0.14	0.13	0.13	0.22	0.22
Obs.	6,317	9,637	2,089	8,897	2,089	8,912
Clusters	241	241	185	240	185	240

Note: Standard errors in parentheses clustered on DHS cluster/village level. WBD Potential is the share of the area covered in stagnant water in the two weeks preceding date of survey. In panel A: Columns 1-2 displays outcomes plausibly affected by waterborne disease. In column 1, the dependent variable = 1 if child has suffered from diarrhoea the past two weeks, and column 2 weight fo age as percentile of wave median. Columns 3-5 are placebo outcomes of health issues less associated with waterborne disease: Fever, cough and anemia (note anemia = 1 for mild to severe measure, thus the high mean). Column 6 records height in cm, also less likely to be affected by current waterborne disease. All health issues (diarrhoea, fever and cough = 1 if child has had in past two weeks). In Panel B, the first four columns record the effect of WBD Potential on diarrhoea with different subsamples: By whether the child is breastfed (columns 1-2, only $\leq$ 24 months children used in this sample), and whether the child was given plain water in the past 24 hours. Columns 5-5 estimates the effect on fever from WBD Potential when dividing the results by whether the child has been given water. Cell, Wave, Calendar month fixed effects and sum of past two weeks precipitation used in all estimations. Individual level controls include birth order, multiple birth, gender, age, mother's age, total fertility of mother, toilet type.

diarrhoea recently by 5.1pp (one sd of WBD Potential increases diarrhoea by 2.6pp). For children who breastfeed the coefficient is small and statistically insignificant. While breastfeeding is not randomised across children and these mothers or children likely differ in more than this aspect, the large difference in the effect of WBD Potential is highly suggestive of a link between WBD Potential, water and diarrhoea, and consistent with public health advice to breastfeed to reduce diarrhoeal risk.

In columns 3-4, we perform the same estimation of WBD Potential on diarrhoea but instead divide the sample by whether the child has been given plain water the past 24 hours. As is consistent with WBD Potential leading to an increase in the likelihood of water to contain pathogens, children who have been given water recently have a 30pp higher likelihood of having had diarrhoea recently and this effect is statistically significant. For the children who have not been given water recently, we estimate a coefficient of -0.08, and it is not statistically significant. Here too, the results suggest

a direct link between WBD Potential, water and the likelihood of contracting diarrhoea. As a final placebo check, in columns 5-6 we again run the same specification and split the sample by whether the child was given water, but change the outcome to whether the child has had a fever recently. If the effect on fever looks similar, it might suggest a general issue with health in the community, or some interaction with the probability of contracting malaria. However, for both subsamples of children we find that the coefficient is small and not statistically different from zero. We take this as further support that WBD Potential accurately measures risk factors in water becoming contaminated with waterborne pathogens.

Next, we provide further evidence that WBD Potential accurately affects water by utilising an indicator which is collected in DHS: The amount of time (in minutes) the household has to their main source of water. We estimate the effect of WBD Potential on the time it takes to the water source, and also divide the sample by the source of water of the household: Whether they preliminarily obtain water from a tap, well or from a natural source of surface water. Intuitively, if WBD Potential accurately captures an increase in stagnant water, the amount of nearby surface water in the ward should increase, reducing the time it takes to get to the closest water source *only for households who collect water from nature*. The estimation results summarised in Table 6 are consistent with such an effect: Overall and for households who mainly derive their water from a tap or well, there is no statistically significant effect on the time to water from WBD Potential. However, in column 4 we see that the effect of WBD Potential reduces the time to water by 68 minutes for households who obtain their water from nature, and the coefficient is statistically significant. This suggests that WBD Potential increases stagnant water share to the extent that households notice this when they collect water.

Table 6
Time to water by water source

	(1)	(2)	(3)	(4)
	Dependent: Time to water (minutes)			
	All	Tap	Well	Nature
WBD Potential	-4.967	-37.56	13.25	-68.38**
	(23.80)	(33.73)	(34.46)	(28.98)
Mean DV	40	33	42	49
Obs.	13,546	3,479	4,617	2,514
Clusters	241	155	202	176

Note: Standard errors parentheses clustered on DHS grid-cell level. All estimations use calendar month, grid cell and wave fixed effects. WBD Potential is the average per cent share of the area covered in stagnant water the two weeks prior to the date of survey. Each column represents a DiD estimation with a different subsample. The first column includes the full sample. We then divide the sample by which source of water the household states: Piped or tap water (2), from any type of well (3), from nature i.e. a river, dam, lake, stream, canal, pond (4). Individual level controls include birth order, multiple birth, gender, age, mother's age, total fertility of mother, toilet type.' Sum of past two weeks precipitation per village/cluster also included.

7.2 Mitigation with sanitation

Having more empirical support for how WBD Potential increases the probability of contracting waterborne diseases, we next return to how WASH practices can worsen or mitigate the spread of waterborne diseases. Another advantage of DHS relative to Uwezo is that not only whether the household has a toilet is collected, but also what type. Based on the UN Sanitation ladder we categorise household toilet types into four categories: No facilities, Unimproved sanitation, Shared facilities, Improved sanitation (Kvarnstrm et al., 2011).[17] We estimate the effect of WBD Potential on diarrhoea in one estimation, interacting WBD Potential with each category of the sanitation ladder, letting the lowest rung of sanitation – no facilities and open defecation – to be the baseline. As in previous analyses of WASH practices, toilet types are not exogenously determined to the household. However, given the existing evidence on the importance of the faecal-oral channel in spreading waterborne diseases, and our estimated results on the link between WBD Potential and diarrhoea, we expect a direct link between sanitation and waterborne diseases. Furthermore, we are mostly interested in the interaction effect between sanitation quality and time-varying exposure to waterborne disease, which is arguably a more exogenous measure than sanitation quality alone.

The estimation results are summarised in Figure 4. Since both the existing literature and our previous results suggest that the WASH practices channel differ by urban and rural contexts, we also divide the sample by whether the households were defined as urban or rural. As expected, we see that WBD Potential increases the probability of the child having had diarrhoea recently for children living in both urban and rural areas, although the effect is both larger and more precisely estimated for urban areas. The direct effect of the sanitation ladder on diarrhoea is precisely estimated to be close to zero for all types of toilets. In contrast, we estimate negative effects for all but one of the three interaction terms between WBD Potential and each rung on the sanitation ladder. For the full sample, it is only households on the highest rung of the ladder where we can precisely estimate a mitigation effect of WASH practices from WBD Potential. The coefficient suggests that households who have these sanitation facilities can completely offset the increased risk from WBD Potential. This is consistent with existing evidence on the link between sanitation and diarrhoea, which suggests that not any sanitation but mainly *high-quality* facilities are important to stop the faecal-oral channel of the spread of waterborne diseases (Troeger et al., 2018; Magana-Arachchi and Wanigatunge, 2020).

Furthermore, we see that the main negative effect of WBD Potential and the mitigation with improved sanitation originate from urban areas, which is consistent with observational literature and historical cases. The large heterogeneity between wards and households with higher and lower types of sanitation also provides an explanation for why only analysing the urban sample yields an imprecisely estimated effect, since there are large heterogeneities. For urban wards, we also see that the step from no facilities to 'unimproved sanitation' also contributes to mitigating the negative effect of WBD Potential, but less so than the high-quality facilities. Interestingly, the effect of the

[17] To better understand the differences between these categories, Figure E.2 in Appendix E.2 illustrates some examples.

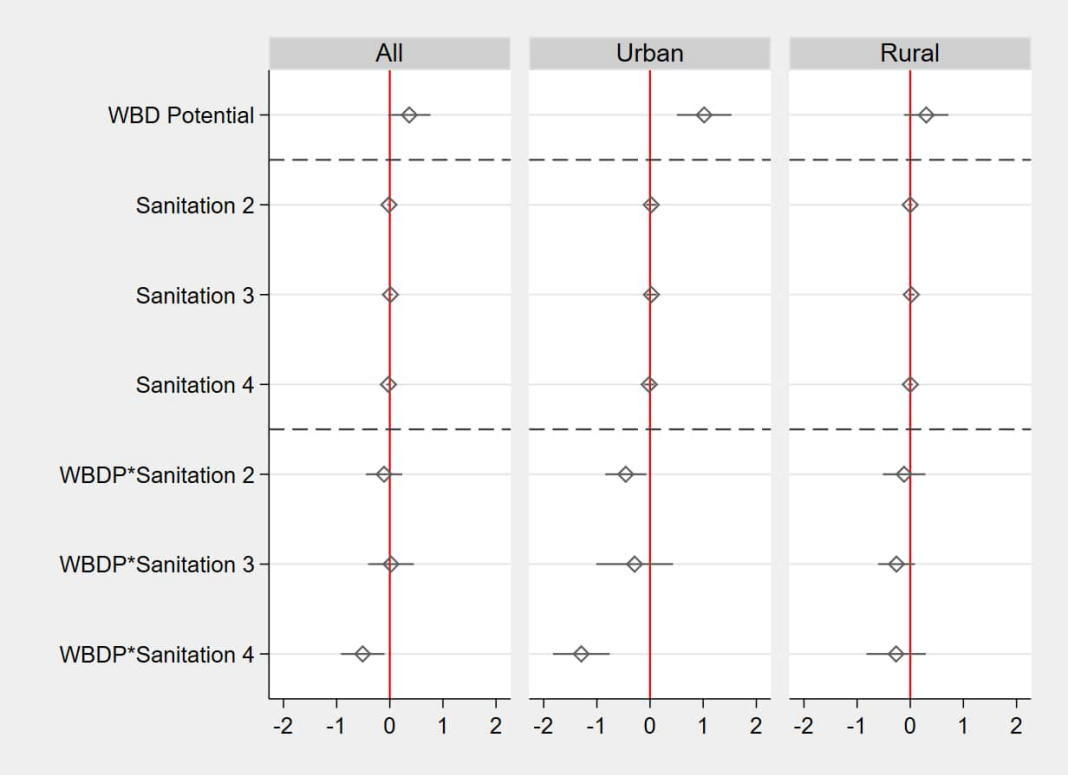

Figure 4
UN Sanitation ladder: Effect of WBD Potential on diarrhoea

Note: The three subgraphs represent separate triple-difference estimations of our main treatment, WBD Potential, interacted with sanitation ladder and the effect on the probability the child has had diarrhoea the past two weeks. The sanitation ladder is a categorical variable representing four types of sanitation facilities: No facilities, Unimproved sanitation, Shared facilities, Improved sanitation. Here, the baseline is households with the first sanitation ladder category: No facilities. The left subgraph presents the results using the whole sample of DHS children, while the remaining two divide the sample by whether the household lives in an urban or rural area (whether the DHS cluster is classified as urban or rural). See Figure E.2 for more explanation of the sanitation categories.

third category, shared facilities, is estimated with the most imprecision. This may reflect that shared facilities may be a poor indicator of better sanitation as there is large heterogeneity in how well they meet the sanitation needs of households (Magana-Arachchi and Wanigatunge, 2020). Indeed, we find that sharing a toilet with other households increases the risk of contracting WBD when WBD potential is high, which is reported in Table E.2.

7.3 Accumulation of waterborne disease on learning

In our main results, we have estimated that WBD Potential defined as the past two weeks of stagnant water negative affects children's test scores. While we find evidence in support of the effect of WBD Potential decaying after some weeks, we have so far made no attempts to characterise how repeated exposure to waterborne disease incidence affects children in the longer run. This is particularly critical in the face of existing evidence showing that adverse events in childhood have lasting effects on cognitive ability (e.g. Maccini and Yang, 2009). On the side of learning, relatively short absences in school can accumulate to form significant negative effects on learning and labour market performance (Cattan et al., 2023). Thus, if children are repeatedly affected by waterborne

diseases, these effects may accumulate in the long run in their overall capacity to learn.

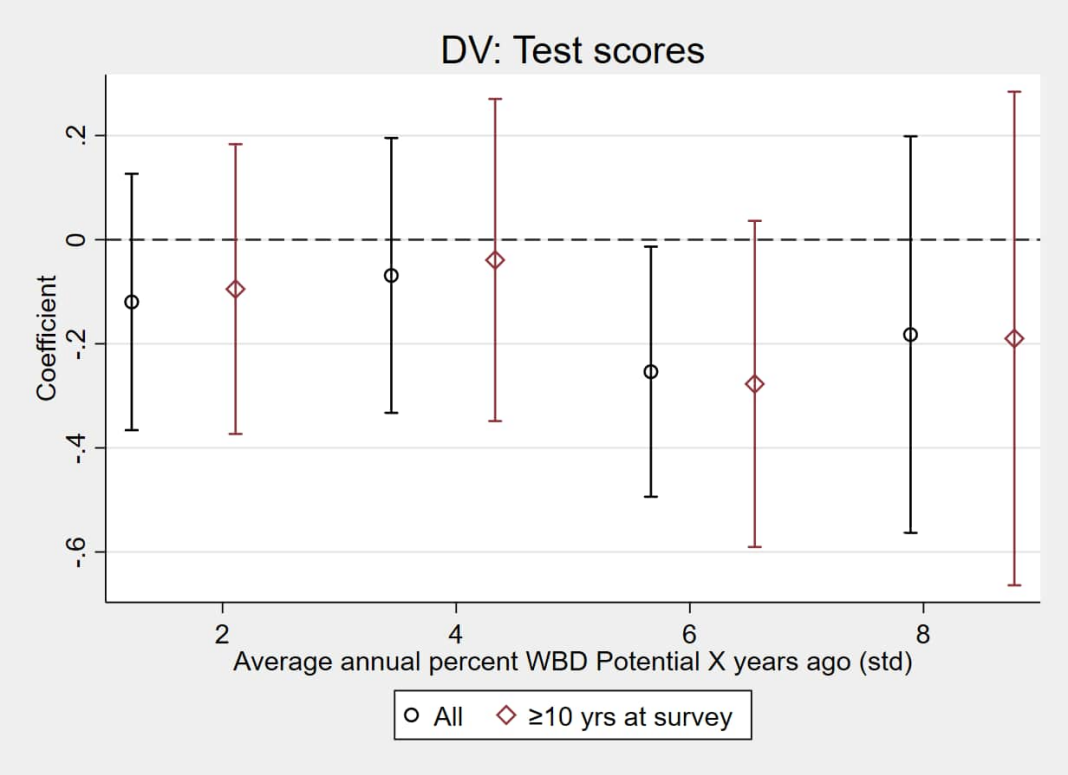

Figure 5
Long-run exposure to WBD Potential and test scores

Note: This figure summarsies the results of WBD from year-accumulation of WBD Potential. For each indicator, we construct an average WBD Potential treatment

In Figure 5 we return to our analysis of the effect of WBD Potential on test scores, but here we take a more long-run perspective. For each year, we calculate the average share of the ward covered in stagnant water. We then take the average of these years at two, four, six and eight years. This measure is then related to the year before the year of the survey, to minimise the influence of short-term effects. For ease of interpretation, we also standardise the treatment. However, going further back in time introduces two issues. First, we may not have comparable samples in the long run since there may be selective mortality and migration.[18] We have not adjusted our estimations for such effects. Second, we cannot distinguish between the years of exposure we measure and the age of children when they are first exposed to the shock. Since young children are more vulnerable to waterborne diseases, we might overestimate the effect on test scores. As one attempt to adjust for this, we exclude the youngest children in our sample and only include ten-year-olds and older in the estimation.[19]

Each estimate in Figure 5 represents a separate estimation. Excluding the younger children

[18]For example, an increase in mortality may lead to a "culling" effect such that stronger (or wealthier) children survive, which would lead us to underestimate our effects. Furthermore, repeated shocks with waterborne diseases may lead to the most vulnerable (wealthy) migrating out of the ward, leading us to under (over) estimate the effects on test scores.

[19]Auxiliary analysis focusing on single age groups yield imprecise estimates, likely because such analysis is underpowered.

increases imprecision but the estimates are otherwise similar. We consistently estimate negative coefficients, but it is only statistically significant for accumulation six years back: One standard deviation of accumulated WBD Potential in the past six-year period implies children's test scores are -0.25 standard deviations lower. We take this as suggestive evidence that there are some negative accumulation effects to having been exposed to WBD Potential in the past in contrast to other children's diseases where immunity may develop. Since the estimates are overall imprecisely estimated, we consider these findings suggestive.

8 Concluding remarks

In this paper we have estimated the effect of WBD Potential, an environmentally determined risk factor of waterborne diseases, in Tanzania across a small geographical unit: Wards. We do this by developing a novel hydrological model where we simulate the share of stagnant water in each ward relative to a known date of testing of children, which we have called WBD Potential. We hypothesise that the environmental risk factors that cause stagnant water pools to grow encourage a host of waterborne pathogens to proliferate: Contact with these pathogens cause a local increase in waterborne diseases which affects children's health and capacity to learn.

Applying a DiD specification to extract the causal effect of waterborne diseases through WBD Potential we find that children exposed to higher waterborne disease potential have lower test scores. In our main specification, we find that children who live in wards where stagnant water covers 10% of the ward area (WBD Potential=0.1) have 0.074 lower standard deviation mean test scores. Put into other words, one standard deviation increase in WBD Potential reduces test scores by 0.03 standard deviations. To compare against another environmental shock, Hyland and Russ (2019) estimate that early childhood droughts reduces years of educational attainment by 0.437 years, or 0.1 standard deviations of years of schooling for children in Sub-Saharan Africa. In contrast to their long-run results, our main estimate measures the effect of a contemporaneous shock directly on student performance and learning, test scores. Our estimated short-term effect is smaller than they estimate on years of schooling but in line with their findings that environmental factors have significant effects on children's education. Moreover, our suggestive evidence for the long-run effect of WBD Potential on test scores suggests large effects on learning, in line with or above their estimated effect sizes.

By moving on to another set of surveys, the Demographic and Health Survey, we find evidence in support of the health mechanism we have hypothesised: 10% WBD Potential increases the incidence of diarrhoea, the main symptom of waterborne diseases by 2.75pp, which is equivalent to 11% of the mean incidence. These effects are large but consistent with existing evidence on the importance of the faecal-oral transmission of waterborne diseases as exemplified by existing literature on the expansion of WASH practices in diarrhoeal incidence (Duflo et al., 2015; Kremer et al., 2022). By exploring both the effect on health and education, we analyse how WASH practices interact with WBD Potential and its effects. We find that the results on test scores are stronger in dry wards, while for rainy and urban wards WASH practices are important co-determinants of the

effect on both health – in the incidence of diarrhoea – and test scores. This is consistent with both historical events and documented disease outbreaks, where sanitation is key in combatting waterborne diseases when there is high population density and in general wetter climate.

We generate two key insights. Firstly, that not only the household's own sanitation facilities matter, but there are important spillovers within a village such that the sanitation practices in the whole village is at least as important as the household, implying that policies to address sanitation practices may also want to address the communities' attitude and capacity for improved WASH practices as a whole. Secondly, that it in particular high-quality facilities offset the increased risk of waterborne diseases, consistent with null results from RCT's that have adopted a piecemeal approach. Thus, WASH programs in Tanzania and other developing countries that battle reoccurring outbreaks of waterborne disease will likely have to invest in a higher-quality (and more expensive) toilet and sanitation facilities to reduce the incidence of waterborne diseases to that of the developed world. One way to make this more cost-efficient is to target areas where WBD Potential is likelier to be higher or more volatile.

While large-scale investment in WASH infrastructure is likely necessary, albeit expensive, in order to improve the local disease environment, an intermediate and potentially cost-efficient step in reducing the worst consequences of an outbreak could be improved targeting of medical treatment to areas at higher risk of outbreaks. To reiterate, we find that one standard deviation increase in WBD Potential reduces test scores by 0.03 standard deviations. This effect size is equivalent to the effect on test scores estimated by Mbiti et al. (2019) who over a two-year period incentivise teachers in Tanzania with 5,000 TZS (3 USD) per student's passing grade (although their estimate is not statistically significant). The cost of this program can be compared to diarrhoea rehydration treatment which costs only 0.56 USD. Despite this high cost effectiveness, only about one in six children receive the treatment in rural Tanzania. We thus leave it to future research to explore demand and supply-side policies to increase children's access to treatment.

Our exploratory analysis on the accumulation of exposure to WBD Potential implies that not only extreme events like severe droughts and floods affect education, but that there are significant losses in learning owing to repeated shocks of waterborne diseases. These results are especially concerning as access to water will be even more strained with the onset of more severe climate change, as water shortage may increase the propensity and need of communities to draw water from unsafe sources. Our analysis of local climactic risk factors of waterborne diseases provide hopeful evidence for how the costs of waterborne diseases can be combatted with sanitation policies if targeted to vulnerable areas.

References

Alsan, Marcella, and Claudia Goldin. 2019. "Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 18801920." *https://doi.org/10.1086/700766* 127 586. 10.1086/700766.

Bandyopadhyay, Sushenjit, Shireen Kanji, and Limin Wang. 2012. "The impact of rainfall and temperature variation on diarrheal prevalence in Sub-Saharan Africa." *Applied Geography* 33 63–72.

Bietenbeck, Jan, Sanna Ericsson, and Fredrick M. Wamalwa. 2019. "Preschool attendance, schooling, and cognitive skills in East Africa." *Economics of Education Review* 73 101909. https://doi.org/10.1016/j.econedurev.2019.101909.

Briceño, Bertha, Aidan Coville, Paul Gertler, and Sebastian Martinez. 2017. "Are there synergies from combining hygiene and sanitation promotion campaigns: evidence from a large-scale cluster-randomized trial in rural Tanzania." *PloS one* 12 (11): e0186228.

Cann, KF, D Rh Thomas, RL Salmon, AP Wyn-Jones, and D Kay. 2013. "Extreme water-related weather events and waterborne disease." *Epidemiology & Infection* 141 (4): 671–686.

Cattan, Sarah, Daniel A Kamhofer, Martin Karlsson, and Therese Nilsson. 2023. "The Long-Term Effects of Student Absence: Evidence from Sweden." *The Economic Journal* 127.

Chae, Sae Rom, Haji Lukupulo, Sunkyung Kim et al. 2022. "An Assessment of Household Knowledge and Practices during a Cholera Epidemic Dar es Salaam, Tanzania, 2016." *The American Journal of Tropical Medicine and Hygiene* 107 766–772. 10.4269/AJTMH.21-0597.

de Chaisemartin, Clment, and Xavier D'Haultfoeuille. 2018. "Fuzzy differences-in-differences." *Review of Economic Studies* 85 999–1028. 10.1093/restud/rdx049.

de Chaisemartin, Clment, and Xavier D'Haultfoeuille. 2022. "Difference-in-differences estimators of intertemporal treatment effects." `http://www.nber.org/papers/w29873`.

Cutler, David, and Grant Miller. 2005. "The role of public health improvements in health advances: The twentieth-century United States." *Demography* 42 1–22. 10.1353/DEM.2005.0002/METRICS.

Dazzi, Susanna, Iuliia Shustikova, Alessio Domeneghetti, Attilio Castellarin, and Renato Vacondio. 2021. "Comparison of two modelling strategies for 2D large-scale flood simulations." *Environmental Modelling & Software* 146 105225.

Deryugina, Tatyana, Garth Heutel, Nolan H Miller, David Molitor, and Julian Reif. 2019. "The mortality and medical costs of air pollution: Evidence from changes in wind direction." *American Economic Review* 109 (12): 4178–4219.

Dinkelman, Taryn. 2017. "Long-run health repercussions of drought shocks: Evidence from South African Homelands." *The Economic Journal* 127.

Duflo, Esther, Michael Greenstone, Raymond Guiteras, and Thomas Clasen. 2015. "Toilets Can Work: Short and Medium Run Health Impacts of Addressing Complementarities and Externalities in Water and Sanitation." 10.3386/W21521.

Eisenberg, Marisa C, Suzanne L Robertson, and Joseph H Tien. 2013. "Identifiability and estimation of multiple transmission pathways in cholera and waterborne disease." *Journal of theoretical biology* 324 84–102.

Falter, Daniela, NV Dung, Sergiy Vorogushyn et al. 2016. "Continuous, large-scale simulation model for flood risk assessments: proof-of-concept." *Journal of Flood Risk Management* 9 (1): 3–21.

Falter, Daniela, Sergiy Vorogushyn, Julien Lhomme, Heiko Apel, Ben Gouldby, and Bruno Merz. 2013. "Hydraulic model evaluation for large-scale flood risk assessments."

Farr, Tom G, Paul A Rosen, Edward Caro et al. 2007. "The shuttle radar topography mission." *Reviews of geophysics* 45 (2): .

Fick, Stephen E, and Robert J Hijmans. 2017. "WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas." *International journal of climatology* 37 (12): 4302–4315.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics* 225 254–277. 10.1016/j.jeconom.2021.03.014.

Hersbach, H, B Bell, P Berrisford et al. 2018. "ERA5 hourly data on single levels from 1979 to present." *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* 10.

Hyland, Marie, and Jason Russ. 2019. "Water as destiny The long-term impacts of drought in sub-Saharan Africa." *World Development* 115 30–45. 10.1016/J.WORLDDEV.2018.11.002.

Ide, Tobias, Anders Kristensen, and Henrikas Bartuseviius. 2021. "First comes the river, then comes the conflict? A qualitative comparative analysis of flood-related political unrest." *Journal of Peace Research* 58 83–97. 10.1177/0022343320966783.

Ilomo, Authors Onesto, and Bernard Mlavi. 2016. "The Availability of Teaching and Learning Facilities and Their Effects on Academic Performance in Ward Secondary Schools in Muheza-Tanzania." *International Journal of Education and Research* 4, www.ijern.com.

Kahabuka, Catherine, Gunnar Kvåle, Karen Marie Moland, and Sven Gudmund Hinderaker. 2011. "Why caretakers bypass Primary Health Care facilities for child care-a case from rural Tanzania." *BMC health services research* 11 (1): 1–10.

Knutsson, Daniel. 2020. "The Effect of Water Filtration on Cholera Mortality."

Kremer, Michael, Stephen Luby, Ricardo Maertens, and Brandon Tan. 2022. "Water treatment and child mortality: a systematic review and meta-analysis For the current version click here. Summary Background." https://www.socialscienceregistry.org/trials/5977.

Kvarnstrm, E., J. McConville, P. Bracken, M. Johansson, and M. Fogde. 2011. "The sanitation ladder a need for a revamp?" *Journal of Water, Sanitation and Hygiene for Development* 1 3–12. 10.2166/WASHDEV.2011.014.

Leclerc, H, L Schwartzbrod, and E Dei-Cas. 2002. "Microbial agents associated with waterborne diseases." *Critical reviews in microbiology* 28 (4): 371–409.

Levy, Karen, Andrew P Woster, Rebecca S Goldstein, and Elizabeth J Carlton. 2016. "Untangling the impacts of climate change on waterborne diseases: a systematic review of relationships between diarrheal diseases and temperature, rainfall, flooding, and drought." *Environmental science & technology* 50 (10): 4905–4922.

Ling, Fangqiong, Rachel Whitaker, Mark W. LeChevallier, and Wen-Tso Liu. 2018. "Drinking water microbiome assembly induced by water stagnation." *Multidisciplinary Journal of Microbial Ecology*.

Lo Iacono, Giovanni, Ben Armstrong, Lora E Fleming, Richard Elson, Sari Kovats, Sotiris Vardoulakis, and Gordon L Nichols. 2017. "Challenges in developing methods for quantifying the effects of weather and climate on water-associated diseases: A systematic review." *PLoS neglected tropical diseases* 11 (6): e0005659.

Lugomela, Charles, Sabrina Moyo, Thomas J Lyimo, Lucy A Namkinga, Ralf Goericke, and Sara Sjöling. 2014. "Co-variations of cholera with climatic and environmental parameters in coastal regions of Tanzania." *Western Indian Ocean Journal of Marine Science* 13 (1): 93–105.

Maccini, Sharon, and Dean Yang. 2009. "Under the weather: Health, schooling, and economic consequences of early-life rainfall." *American Economic Review* 99 (3): 1006–26.

Magana-Arachchi, D.N., and R.P. Wanigatunge. 2020. "Ubiquitous waterborne pathogens." *Waterborne Pathogens* 15. 10.1016/B978-0-12-818783-8.00002-5.

Masanja, Honorati, Pyande Mongi, Jitihada Baraka et al. 2019. "Factors associated with the decline in under five diarrhea mortality in Tanzania from 1980-2015." 10.7189/jogh.09.020806.

Masunga, Daniel Stephen, Anushree Rai, Mortada Abbass et al. 2022. "Leptospirosis outbreak in Tanzania: An alarming situation." *Annals of Medicine and Surgery* 80 104347. 10.1016/J.AMSU.2022.104347.

Mayala, BK, LEG Mboera, and F Gwacha. 2003. "Mapping of cholera risks using geographical information system in Ilala district, Tanzania." *Tanzania Journal of Health Research* 5 (1): 8–12.

Mbiti, Isaac, Karthik Muralidharan, Mauricio Romero et al. 2019. "Advance Access publication on." *The Quarterly Journal of Economics* 1627–1673. 10.1093/qje/qjz010.

Mellon, Jonathan. 2021. "Rain, Rain, Go Away: 176 potential exclusion-restriction violations for studies using weather as an instrumental variable." *Available at SSRN 3715610*.

Mgema, B. 2022. "Teacher Absenteeism and Curriculum Implementation in Primary Schools in Nyamagana District in Mwanza Region of Tanzania." *East African Journal of Education Studies* 5 116–124. 10.37284/EAJES.5.2.679.

Moore, Sean M, Andrew S Azman, Benjamin F Zaitchik et al. 2017. "El Niño and the shifting geography of cholera in Africa." *Proceedings of the National Academy of Sciences* 114 (17): 4436–4441.

Nkoko, Didier Bompangue, Patrick Giraudoux, Pierre-Denis Plisnier et al. 2011. "Dynamics of cholera outbreaks in Great Lakes region of Africa, 1978–2008." *Emerging infectious diseases* 17 (11): 2026.

Poggio, Laura, Luis M De Sousa, Niels H Batjes, Gerard Heuvelink, Bas Kempen, Eloi Ribeiro, and David Rossiter. 2021. "SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty." *Soil* 7 (1): 217–240.

Prüss, Annette, David Kay, Lorna Fewtrell, and Jamie Bartram. 2002. "Estimating the burden of disease from water, sanitation, and hygiene at a global level.." *Environmental health perspectives* 110 (5): 537–542.

Rebaudet, Stanislas, Bertrand Sudre, Benoît Faucher, and Renaud Piarroux. 2013. "Cholera in coastal Africa: a systematic review of its heterogeneous environmental determinants." *The Journal of infectious diseases* 208 (suppl_1): S98–S106.

Rinaldo, Daniele. 2020. "Dams, Snails and Poverty Traps." https://github.com/daniele-rinaldo/economic_impact_schisto.

Rosling, Hans, Ola Rosling, and Anna Rosling Rnnlund. 2018. *Factfulness: Ten reasons we're wrong about the world - and why things are better than you think*.

Schellenberg, Joanna Armstrong, Cesar G Victora, Adiel Mushi et al. 2003. "Inequities among the very poor: health care for children in rural southern Tanzania." *The Lancet* 361 (9357): 561–566.

Shah, Manisha, and Bryce Millett Steinberg. 2017. "Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital." *Journal of Political Economy* 125 (2): 527–561.

Tanzania Ministry of Health and Social Welfare. 2008. "Health Sector Strategic Plan III July 2009 June 2015." https://extranet.who.int/nutrition/gina/en/node/8824.

Troeger, Christopher, Brigette F Blacker, Ibrahim A Khalil et al. 2018. "Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016." 10.1016/S1473-3099(18)30362-1.

UNESCO. 2014. "Education for all (EFA) report for Tanzania Mainland." *Education for all 2015 National review*.

UNICEF. 2022a. "Oral rehydration salts and zinc: Market and supply update." https://www.unicef.org/supply/media/13851/file/ORS-and-Zinc-Market-Supply-Update-September-2022.pdf.

UNICEF. 2022b. "Water and climate change - 10 things you should know." https://www.unicef.org/stories/water-and-climate-change-10-things-you-should-know.

Vos, Theo, Stephen S Lim, Cristiana Abbafati et al. 2020. "Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden

of Disease Study 2019." *The Lancet* 396 (10258): 1204–1222.

WHO. 2010. "Combating waterborne disease at the household level." *World Health Organization*, https://apps.who.int/iris/rest/bitstreams/51808/retrieve.

WHO. 2022a. "Drinking water and health." https://www.who.int/news-room/fact-sheets/detail/drinking-water.

WHO. 2022b. "WHO responds to a Cholera outbreak in Katavi Region." https://www.afro.who.int/countries/united-republic-of-tanzania/news/who-responds-cholera-outbreak-katavi-region.

World Bank. 2016. "Service Delivery Indicators SDI Results for Tanzania."

Appendices

Table A.1
Data sources used in the WBD potential algorithm

Input category	Dataset	Resolution and accuracy of input data
Topography	Digital Elevation Model from the Shuttle Radar Topography Mission (Farr et al., 2007)	30 m horizontal resolution, aggregated into 90 m. 6 m vertical accuracy.
Rainfall	ERA5-reanalysis data for Precipitation (Hersbach et al., 2018)	27 km horizontal resolution. 1 hour time resolution.
Evaporation	ERA5-reanalysis data for Potential Evapotranspiration (Hersbach et al., 2018)	27 km horizontal resolution. 1 hour time resolution.
Soil data	ISRIC 2.0 global soil database (Poggio et al., 2021)	1 km horizontal resolution.

Note:

A Geographic data and modelling stagnant water

A.1 Data sources

We use four different categories of data for our algorithm to compute the WBD potential, as summarised in Figure 1 and Table A.1, which are: topographical data, rainfall data, evaporation data and soil infiltration data.

We use topographical data from the Shuttle Radar Topography Mission (Farr et al., 2007). This is one of the most commonly used high-resolution global datasets on topography in scientific research, and is assembled by satellite data gathered in February 2000. Since we are not aiming to resolve detailed features in urban areas but rather large-scale runoff processes, we believe that this data is representative of the topography in Tanzania for the time period of our survey data. The resolution for each grid cell is approximately 30 m, with a vertical accuracy of around 6 m. Since we run our algorithm at the country level, we aggregate this data to a 90 m resolution in order to make the computation feasible. This still leaves us with more than 120 million grid cells for the area of Tanzania and likely provides a detailed enough horizontal resolution for the large-scale analysis we run. Recent evaluations of large-scale 2D hydrodynamic simulations for several European rivers show that resulting ooded area and water level are insensitive to variations in spatial resolution once it is finer than 100 m (Dazzi et al., 2021; Falter et al., 2013, 2016). The topography is used to determine the flow direction of each cell during the simulation, which may change dynamically as a function of the water depth over the cell, and to determine the water depth across cells at each point in time.

For rainfall input, we use data from ERA5, which provides global gridded dataset with hourly estimates of a multitude of atmospheric variables, including precipitation (Hersbach et al., 2018). We use the reanalysis data, which is based on an ensemble of forecast models which take both

satellite data and local weather station data into account and updates predicted atmospheric variables at a 1-hour resolution. The spatial resolution is 0.25 decimal degrees, which approximately translate to 27 km at the equator. This gives us close to 1400 data points for Tanzania, which is vastly greater than the number of rainfall stations in the country. Gridded rainfall data derived from satellite observations is especially useful in developing countries, where there is often a relative scarcity and lower quality of data from rainfall gauges. Moreover, rainfall stations may vary systematically with local development, which means that measurement error may become systematically correlated with the local level of development. The advantage with using satellite-adjusted data is that measurement error is arguably orthogonal to local development. With regards to hydrological simulations, rainfall is especially important since it tends to be one of the main sources of uncertainty. Fortunately, a recent study investigating 22 global gridded rainfall datasets systematically found that ERA5 reanalysis data provided one of the best calibration scores and lowest inaccuracies when used operationally in hydrological models (Beck et al., 2017). The rainfall data provides all the input of water in the model, and thus drives the resulting surface runoff, depending on local infiltration rates, topography and evaporation.

For evaporation we also rely on the ERA5 reanalysis data, for the same reasons given above (Hersbach et al., 2018). This also lends consistency to the rainfall data, since these two datasets are produced jointly and dependent on each other. Specifically, we use the *potential* evapotranspiration rate which is applied to cells with a water depth greater than zero.

Lastly, for soil infiltration, we use soil data from the ISRIC 2.0 global soil database (Poggio et al., 2021). This is a state-of-the art high resolution soil dataset which provides a resolution as fine as 1 km and provides a distribution of the content of clay, silt and sand at different soil depths. We depth-integrate this data and use the distribution of soils in each cell to classify each cell as a soil type according to the USDA classification system. From this classification we can then derive soil parameters such as saturated infiltration capacity and soil porosity. The current version of our algorithm uses a simplified infiltration measure by applying only the saturated conductivity of the soil to account for losses due to soil infiltration. Due to the high resolution of the topography we do not apply any slope-adjusted infiltration rates as some low-resolution hydrological models do. Instead a greater slope will translate into a faster runoff process, which will reduce the resulting infiltration. Future iterations of this algorithm could potentially be improved by applying a full soil infiltration mode commonly use in state-of-the-art hydrological models, such as the Green-Ampt method, and also by simulating the groundwater storage layer as a separate entity able to refeed the infiltrated water as groundwater seepage into rivers and streams, which as of current is not handled by the algorithm.

A.2 Algorithm for WBD Potential

The purpose of the algorithm is to model the time evolution of stagnant water surfaces over time, which we wish to aggregate to a weekly-level treatment measure, using a combination of hydrological and hydraulic calculations. For this purpose, we run the model with 5-minute temporal

resolution. Before the model is run, input data are processed to cover the same extent and transformed to the same coordinate system. Below is a schematic explaining how the algorithm, which is implemented in Python, works.

1. Initialize and run the model starting > 3 months before the first interview date until the last interview date.
   (a) Initalize a new day d
   (b) For every 5-minute timestep $t \in \{1, ..., 288\}$ in each day:
      i. Accumulate the last 5 minutes of precipitation on all grid cells
      ii. For each cell i, j identify the immediately neighboring cell with the lowest current water level, where water level is the sum of the current water depth and elevation of the cell.
      iii. If this cell has a lower water level than cell i, j, then transfer the 5-minute accumulated volumetric flow V from cell i, j to the receiving cell, otherwise do nothing. The volumetric flow rate $V = f(\Delta h)$ is an increasing function in the difference in energy levels (Δh) between the emitting and receiving cells, based on the Manning formula for open channel flow.
      iv. Finally, remove the 5-minute accumulated volumetric infiltration rate from cell i, j: $\min(d_{i,j}, f(K))$ where $d_{i,j}$ is the water depth of cell i, j and K_s is a soil infiltration parameter. The current version, $f(K) = K_s$ where K_s is the saturated hydraulic conductivity of soil s in cell i, j.
   (c) At the last time step of each day, $t = 288$, subtract the volumetric actual evaporation rate E from all cells where $E = \min(d_{i,j,t}, E_d)$ where $E_{i,j,d}$ is the potential evaporation rate in cell i, j for day d.
2. Export water depth for all cells to daily georeferenced arrays, which are then aggregated into weekly-level treatment measures at the ward-level using ArcGIS and the ArcPy package for Python.

A.3 Validation with satellite data

Optimally, one would want to to have a measure of the actual surface water. One such way could be through satellite imagery. There now exists a global database of surface water down to a 30 m resolution, released by the Joint Research Centre of the European Commission and spanning the time period 1984-2021 (Pekel et al., 2016). One problem, however, is that data is only available at the monthly level, which is too aggregated for the short-run effects we analyze, which is at the weekly level. Moreover, with the temporally disaggregated data (at the month level) missing data due to e.g. incomplete satellite coverage and cloud cover is common, at least for Tanzania. Lastly, there is also the potential issue that observed surface water is endogenous. It could, for instance,

be that areas where water is cleared away faster have better access to functioning infrastructure and are more developed. By instead simulating surface water using static topography and using time variation in climate variables, as we do in our baseline algorithm, we arguably get an exogenous source of variation in surface water.

One way to get around the potential endogeneity issue and also deal with missing data is to use the long-term occurence data, which reports the percent of months a cell was covered by surface water (p_{sw}), and then use variation only in hydrological input (rainfall, infiltration, evaporation) to predict whether a specific cell is covered by water. This method would rely on only simulating variation in hydrology through the simulation of local runoff, and then to infer whether a given cell is covered by water or not in that scenario.

Since our algorithm integrates both hydrological and hydraulic computation, one alternative, in order to isolate hydrological variation, is to combine the surface water frequency data with an external purely hydrological model. To this end, we use the GloFAS-ERA5 model, which builds on ERA5 data similar to our algorithm, but is a much more advanced hydrological model with the purpose to simulate river discharge at the local level (Harrigan et al., 2020). One advantage is that it is calibrated and validated, and used operationally around the globe. The resolution of this model is much coarser however, approximately 11 km, so cannot be used in itself to infer which cells become covered by stagnant water. However, from long-term output by the model we can generate a hydrological frequency distribution for each cell. We can then run the model at a daily timestep and elicit the percentile value of the hydrological situation for each day (p_h). Assuming that inland surface water occurrence is mostly determined by the current hydrological situation, this implies that a cell will be considered covered by water only if $p_h > 1 - p_{sw}$. For example, a cell that is observed to be covered by water only 1 % of the time ($p_{sw} = 0.01$) would require a local river discharge percentile value greater than or equal to 99 % ($p_h \geq 0.99$) to count as flooded, since 99 % of the time, it should be "dry". We use this as an alternative measure to validate the findings from our baseline algorithm, as well as an robustness exercise, which we report in Table D.5. Moreover, we also use it for the long-term accumulation of stagnant water shocks.

In terms of correlation between the two treatment measures, it is consistently positive but low. It is the lowest for rural areas (0.02) and highest for urban areas (0.52), where we also get the most consistent estimates between the two methods (both yield negative effects on learning, albeit insignificant). While it is reassuring that there is some degree of positive correlation, the fact that it is generally low may imply that these methods are largely complementary to each other. The satellite occurrence data is based on monthly observations, which means that it will fail to capture areas that are only covered by water during short periods of time, such as a few days. Our main algorithm, which is run with a time resolution in minutes but exported at the daily level, is thus potentially able to capture more cells covered by stagnant water. Any additional area that is identified by our algorithm but not the satellite data will thus reduce the correlation between the two measures, which could be sensitive due to the frequency of zero-valued data (dry cells) regardless of the method used, but this would mostly be a positive feature of our model rather than an inaccuracy. In terms of the algorithm, there are a number of potential inaccuracies. The

most straightforward one is topography. We use 30 m resolution data and aggregate into 90 m for computational feasibility, which is likely to further decrease the vertical accuracy. Evaluation of topography data from the SRTM tends to show substantially larger inaccuracy in rural than urban areas, since urban areas by design tend to have less topographic variation, and thus be less sensitive to measurement and aggregation error. The area flooded by water is likely sensitive to variation in the topography, since a small change in slope would lead water to run off rather than stagnate. A secondary, typically large source of uncertainty in any hydrological model, are soil-water processes. Generally, runoff models work very well in urban areas, where infiltration is small and most runoff is due to rainfall, whereas rural areas, resulting surface water will to a larger degree be determined by soil composition, vegetation and groundwater depth, all which could be sources of inaccuracies in our model. In terms of magnitudes, we find our algorithm produces an average for the simulation periods of 1.3 %, whereas the satellite imagery generates a mean of 0.6 %. As argued before, this could be partly due to the fact that satellite imagery uses a monthly resolution, so would likely capture a strictly smaller area of surface water than our algorithm, and further that infiltration and runoff processes may be systematically underestimated by the algorithm. Comparing urban areas only, where these inaccuracies should play a significantly smaller role, we find that our hydrological model generates a mean of 1.4 % and the satellite data 0.9 %, which hints at the effects of inaccuracies due to topographical and soil infiltration uncertainty. Hence, the remaining systematic difference in magnitude may be driven by our ability to capture the short-term variation in surface water in urban areas to a larger degree.

Since most of our results, from learning to health outcomes, seems largely driven by urban areas, it is reassuring that the overlap of the methods is strongest for urban areas. This is consistent both on the independent variable side, with surface runoff models showing less uncertainty in urban areas, and on the dependent variable side, where urban areas have typically found to be at a greater risk of waterborne disease outbreaks.

A.4 Generating the accumulation measure

In addition to short-run effects we are also interested in analyzing the accumulated effects of many short-term shocks, i.e. over several years. For this, we need a measure of daily stagnant water all across Tanzania starting from the year where the oldest children in the earliest wave were just born – 1995. This is computationally very expensive for our baseline algorithm. Instead, we use our alternative treatment measure defined above, where we combine the occurrence data of satellite data on surface water with daily percentile values from the GloFAS hydrological model. We run this for the period 1995 to 2017. This accumulation runs orders of magnitudes faster and should arguably give similar results to our baseline algorithm at the scale that we run our analysis at.

We then create the accumulation measure by first, for each day, aggregating the share of water-covered cells in each ward, identically to our short-term treatment. We then calculate the average stagnant water share of each year as our baseline accumulation measure. Hence, a ward that experiences 1 % stagnant water over 100 days count the same as a ward that experiences 10 % stagnant

Table B.1
Common waterborne diseases and symptoms

Pathogen type	Example	Common symptoms
Bacteria	Cholera, Salmonella (Typhoid fever), Shigella (Dysentery), E.coli, Legionella	Diarrhoea, Fever, blood in stool
Viruses	Rotavirus, Adenovirus, Astrovirus, Hepatitis A and E	Diarrhoea, Gastroenteritis, Fever
Protozoa parasites	Cryptosporidia, E. histolytica	Diarrhoea, Gastrointestinal illness
Parasitic worms (Helminths)	Roundworms, Hookworms, Trematodes (flat worms), Schistosomiasis	Fever, Abdominal pain, Diarrhoea, Gastrointestinal illness, Malnutrition,

Note: Sources: Magana-Arachchi and Wanigatunge (2020), Hedley and Wani (2015), WHO (2019).

water over 10 days. Based on the age of each observed child and the survey year, we can then create a measure that reports the share of time in the past n years, prior to the survey year, that a child was subject to these short-term treatments. In Figure 5 we show the results of this exercise looking back 2, 4, 6 and 8 years and separate the effects between all children and those that are at least 10 years old at the time of the survey.

B Details on Waterborne Diseases

B.1 Mechanisms for contagion and disease symptoms

Waterborne diseases are adverse health conditions caused by pathogens that are transmitted by the intake of or contact with pathogen-polluted water, such as by the intake of harmful bacteria or worms. These pathogens include bacteria, viruses and worms, and common diseases are cholera, typhoid fever, and dysentery. Table B.1 summarises some of the most commonly occurring diseases and symptoms, by type of pathogen. Symptoms vary depending on its cause, but the most common one by far is diarrhoea and other issues relating to the gastrointestinal issues, such as abdominal pain (Magana-Arachchi and Wanigatunge, 2020).

For clarification, neither malaria nor chemically polluted water that causes health issues are typically considered waterborne diseases. First, malaria is not a waterborne disease. Malaria is a serious and sometimes fatal disease which infects humans via a parasite carried by mosquitoes (and not through contact with contaminated water per se) (WHO, 2019). Malaria causes flu-like symptoms such as fever and vomiting, but is in contrast to waterborne diseases not associated with diarrhoea. Second, chemically polluted water is not a waterborne disease. While water polluted with e.g. pesticides (Boedeker et al., 2020) or arsenic (Mandal et al., 1996) lead to large adverse health effects, they are not waterborne diseases since this term typically is reserved for pathogen-

induced disease.

B.2 Spread, infection and recovery of waterborne diseases

Waterborne diseases spread via water, either by directly drinking water infected with waterborne disease pathogens or by swimming in contaminated water. For most waterborne diseases, the so-called faecal-oral channel where contaminated human feces spread to water or food which then is ingested by another human is the most common source of an outbreak (Magana-Arachchi and Wanigatunge, 2020). For example, a person could become infected by eating food prepared from agricultural output using human faeces as fertiliser. Thus, the spread of waterborne disease is inextricably associated with Water, Hygiene and Sanitation Practices (WASH) and improving these can break the faecal-oral transmission channel.

Waterborne pathogens occur naturally in rivers and lakes, but grow exponentially under conditions with stagnant water that become contaminated. Poor sanitation means an outbreak can spread across persons and households via the fecal-oral channel. In a lab-controlled environment, exponential growth of bacteria causing waterborne disease can occur within hours (Farhat et al., 2018; Zlatanovi et al., 2017). Ling et al. (2018) found that the whole water supply of a Chinese city became contaminated within six days under conditions where water stayed stagnant in the city's plumbing system.

For most contraction of waterborne disease, the time from first contact to disease outbreak is a few days. For example, cholera takes between 2 hours to 5 days for a person to show symptoms after first ingesting contaminated food or water (Azman et al., 2013). Symptoms can last from days (e.g. a virus infection) to years (some types of worms), but most common is days or few weeks (Percival et al., 2014). In severe cases waterborne disease can be fatal, in particular for young children, with severe diarrhoea leading to dehydration and death if left untreated.

B.3 Waterborne disease burden in Sub-Saharan Africa

According to the UN, the Sustainable Development Goal to achieve safe drinking water for all will fall short of 1.6 billion people by 2030 (UN, 2022). The consequence of this shortfall is dire: The United Nations estimate that 829,000 people die from lack of safely managed drinking water each year (UN, 2022). Strikingly, the burden of waterborne disease caused by unclean water fall almost exclusively on developing countries and in particular Sub-Sahran Africa (Anthonj et al., 2018). For instance, Black et al. (2010) estimate that each day 2,000 children under the age of five die in Africa due to diarrhoea – the second largest source of child mortality[20]. However, the mortality number masks an even higher incidence of disease: Troeger et al. (2018) found that average diarrhoea episodes per person in Sub-Saharan Africa were 1.05 per year. Thus, waterborne disease does not only pose a fatal risk, but also when less severe is likely to affect the vast majority of people in the region several points in their lives.

[20] 18% of total mortality, the largest source of mortality being neonatal conditions in the first 28 days of life for infants (Black et al., 2010)

C Descriptive statistics about treatment

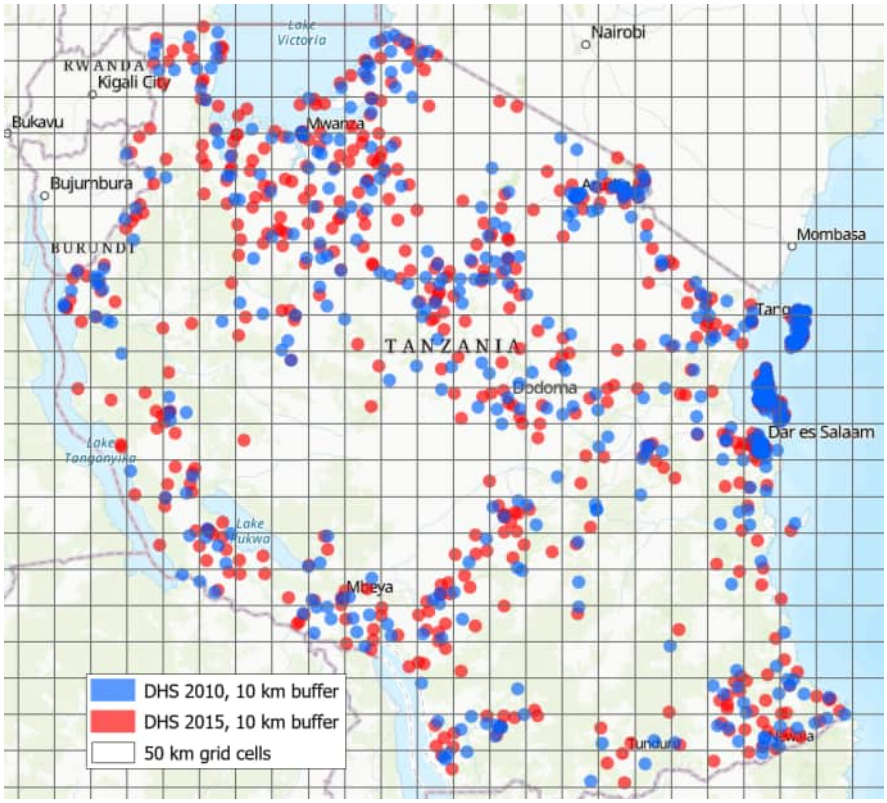

Figure C.1
DHS clusters

Note: The map shows DHS clusters for the survey years 2010 and 2015 with a 10 km buffer, for which share of stagnant water is computed. To account for unobserved geographic variables, we use 50 km grid cell fixed effects, as indicated by the grid shown in the map.

In the descriptive statistics presented in Table 1, we see that while the treatment measure has a relatively low mean at 1.3 %, the variation is substantial, with a standard deviation of 3.57 %. Figure C.2 shows the full histogram of the treatment and provides more details on the full variation of the treatment magnitude. The distribution is indicative of a power law, not uncommon for hydrological distributions, with approximately 3 % of ward-year observations having more than 30 % of their area covered by stagnant water in a given two-week period.

Figure 3 in the main text shows the spatial distribution of the mean value of the treatment at the ward-level. We see that the spatial heterogeneity is quite substantial, with most of the treatment coming from the south-east of Lake Victoria in northern Tanzania, in northwestern Tanzania and along the Ruvuma river system, near the border to Mozambique in the south. Relatively little of the areas along the eastern coast, such as the area around Dar-es-Salaam, get any substantial treatment in the survey periods, which in this map represents the Uwezo data. Fortunately, the areas that are likely to contribute to our estimated treatment effect are relatively dispersed across the country, which reduces the risk of spatial autocorrelation biasing our standard errors. Note that this map is indicative only of the specific months and years covered by the Uwezo survey data. The situation could look very different for other time periods, if rainfall, evaporation and other climatological

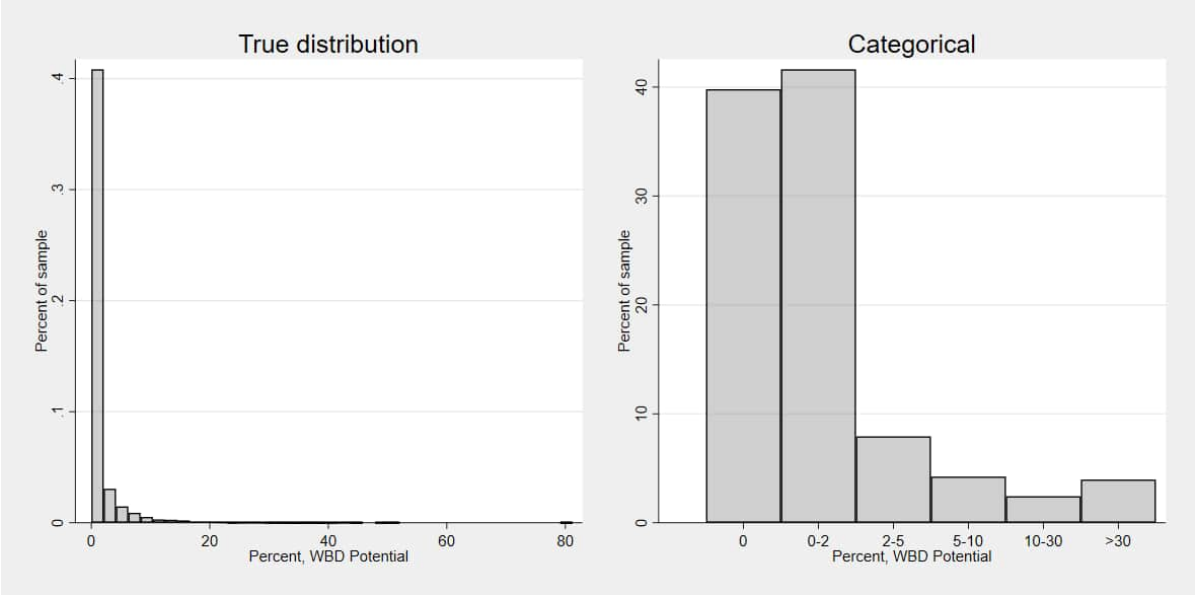

Figure C.2
Distribution of WBD Potential

Note: This figure depicts the distribution of waterborne disease potential (WBD Potential), where a ward w during wave and survey year y has one simulated value between 0 and 1, which is the share of the ward area covered by stagnant water. The left subplot depicts the distribution of of WBD Potential, and to provide additional clarity the right subplot displays the distribution across five categories of shares of stagnant water.

conditions would change, and hence should not be interpreted as a static measure of waterborne disease risk in Tanzania.

D Additional results with test scores

While we test the interaction between sanitation practices and the share of stagnant water, another aspect of WASH might be the source of water. We thus interact the stagnant water share with an indicator for whether the households take their water supply from a natural source such as a river, lake or spring. It should however be noted that the registered water source could also be a reaction from more contaminated water, which would bias our result negatively.

We find large negative coefficients on test scores from the share, and the effect is statistically significant at the 10% level of significance for dry wards. However, there is no significance on the interaction effect between stagnant water share and the source of water. One reason might be that the source of water is an outcome of treatment. Another reason may be that our indicator does not distinguish between type of natural water source: As our model by design cleared out flowing water bodies from the modelling of stagnant water, this provides evidence for the effect on education to not be from households receiving water from flowing water bodies. Thus, we need to find an indicator which allows us to separate water source from stagnant water versus flowing water.

Table D.1
Effect of WBD Potential on test scores: By Long-run precipitation

	Dependent: Test score (std)		
	All	Dry wards	Rainy wards
WBD potential	-0.742**	-0.812**	-0.00542
	(0.315)	(0.349)	(0.729)
Obs.	368,444	178,449	189,995
Clusters	3,842	1,669	2,173
Wave FE			
Ward FE	✓	✓	✓
Month FE	✓	✓	✓

Note: Standard errors in parentheses clustered on ward. WBD Potential is two-week average share of area of ward covered in stagnant water, $\sim$(0,1). Dry ward if mean precipitation $<$ 1000 mm precipitation. Rainy ward if $\geq$ 1000 mm precipitation. Wave, Calendar month, Ward fixed effects, and ward-level 2-week sum of precipitation included in all estimations. Household covariates included are child's gender and age, and mother's age and whether secondary education or above.

Table D.2
WBD Potential and test scores, by water source

	Dependent: Test score (std)		
	All	Dry wards	Rainy wards
	Panel A: Water from river, lake or spring		
WBD Potential	-1.331**	-2.032***	1.413
	(0.613)	(0.666)	(1.085)
Water from nature	-0.0374*	-0.0245	-0.0442*
	(0.0192)	(0.0281)	(0.0263)
WBDP*Water from nature	0.823**	0.687	0.965
	(0.389)	(0.486)	(0.677)
Mean Water	0.14	0.14	0.13
Obs.	167,201	87,298	79,903
Clusters	2,991	1,506	1,485
	Panel B: Water from tap		
WBD Potential	-0.693	-1.341*	1.750
	(0.744)	(0.757)	(1.803)
Water from tap	0.0720***	0.0874***	0.0589***
	(0.0139)	(0.0192)	(0.0201)
WBDP*Water from tap	-0.383	-0.557	-0.324
	(0.367)	(0.448)	(0.704)
Mean Water	0.41	0.39	0.43
Obs.	138,691	72,981	65,710
Clusters	2,908	1,476	1,432

Note: Standard errors parentheses clustered on ward. WBD Potential is two-week average share of area of ward covered in stagnant water, $\sim(0,1)$.*Water from nature* is an indicator variable =1 if household gets their water from a river, lake or spring. *Water from nature* is an indicator variable =1 if household gets their water from the tap. *Mean Water* is the share of households for each water source across the whole sample. Rainy ward if $\geq$ 1000 mm precipitation. Wave, Calendar month and Ward fixed effects included in all estimations. Past two week sum of precipitation by ward used as control.

Figure D.1
WBD Potential on test scores, by sex

Note: This figure summarises DiD estimate of WBD Potential on test scores. We separately present results for boys and girls, and further display estimates where we divide wards by their long-run precipitation. In all estimations we include calendar month, Wave and ward fixed effects as well as individual and household covariates. Standard errors are clustered on ward-level.

D.1 Results with test scores: Heterogeneity

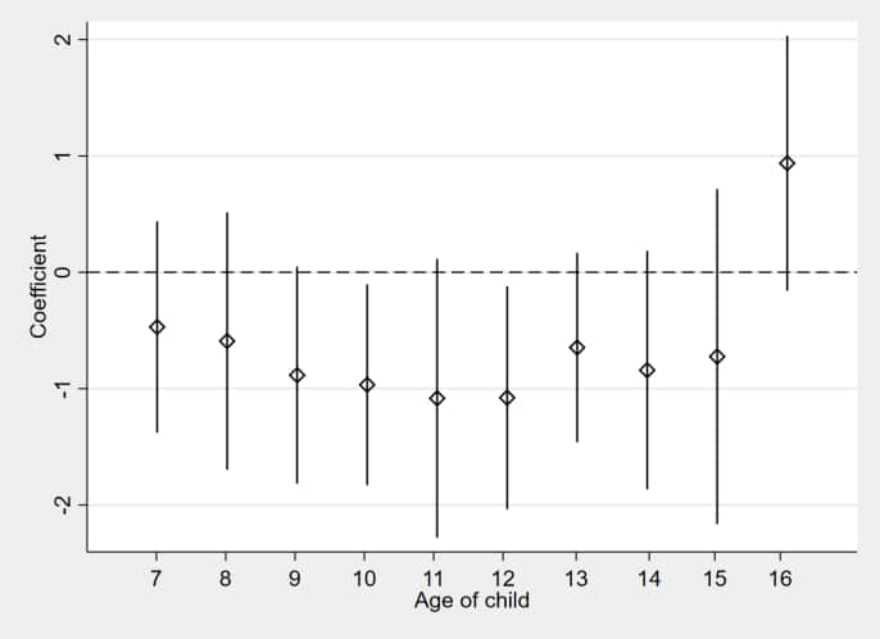

Figure D.2
WBD Potential on test scores, by age

Note: This figure summarises DiD estimate of WBD Potential on test scores. We separately present results by the age of the child at survey. In all estimations we include calendar month, Wave and ward fixed effects as well as individual and household covariates. Standard errors are clustered on ward-level.

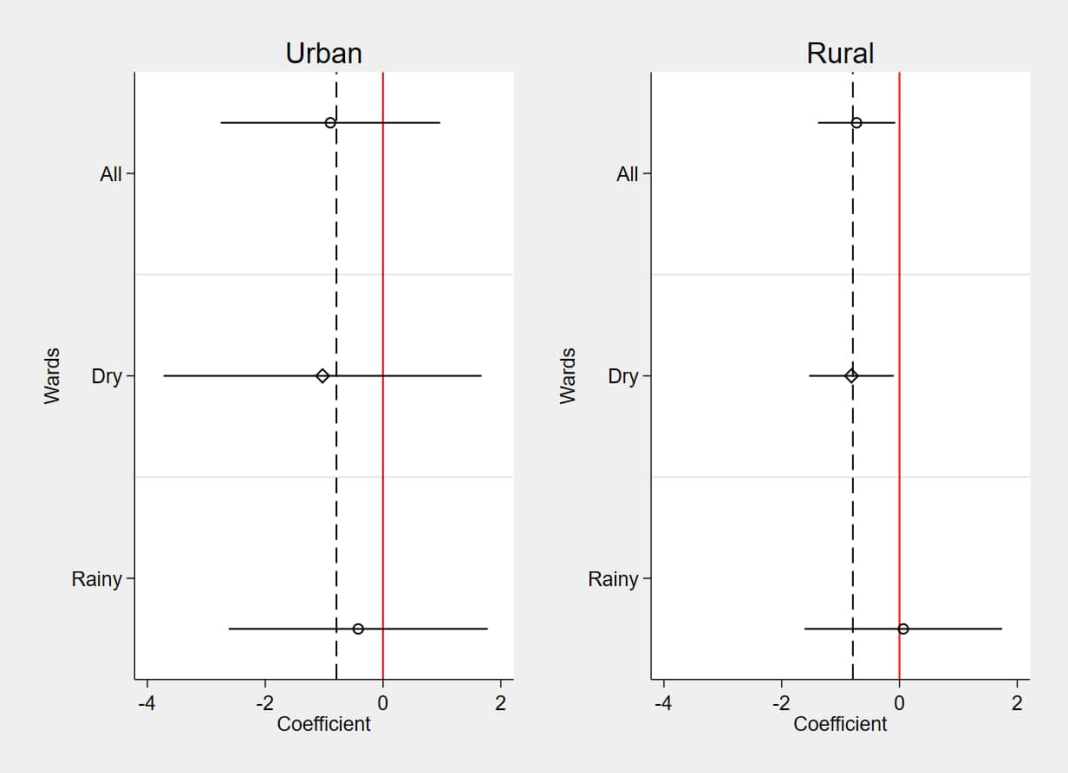

Figure D.3
WBD Potential, by urban-rural wards

Note: This figure summarises DiD estimate of WBD Potential on test scores, splitting each subplot by whether the ward is urban or rural. The solid red line represents a zer ocoefficient of WBD Potential on test scores, and the dashed black line the estimated coefficient size for the full sample. Within each subplot we further split wards by their long-term precipitation: *Dry* ward if long run precipitation $<$1000 mm per year on average, *Rainy* ward if $\geq$ 1000 mm.

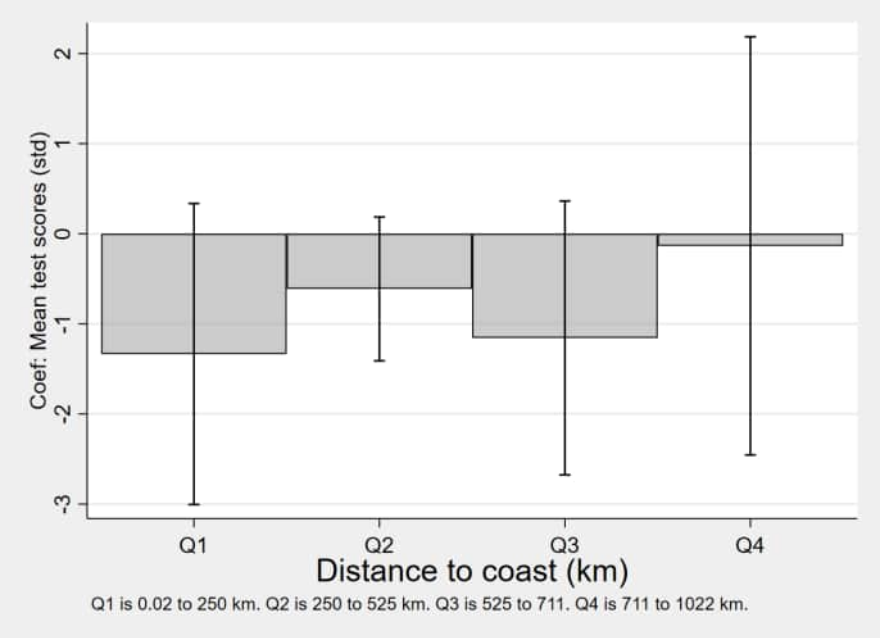

Figure D.4
WBD Potential on test scores, by distance to coast

Note: This figure summarises DiD estimate of WBD Potential on test scores, splitting the sample by the ward's distance to the coast in four separate estimations. In all estimations we include calendar month, Wave and ward fixed effects as well as individual and household covariates. Standard errors are clustered on ward-level.

D.2 Results with test scores: Robustness

Table D.3
WBD Potential on test scores, binary treatment definition

	Dependent: Test score (std)		
	All	Dry wards	Rainy wards
WBDP $\geq$ 5%	-0.101**	-0.102**	-0.0932
	(0.0406)	(0.0457)	(0.0890)
Obs.	368,444	178,449	189,995
Clusters	3,842	1,669	2,173
Wave*District FE			
Ward FE	✓	✓	✓
Month FE	✓	✓	✓

Note: Standard errors in parentheses clustered on ward. WBD Potential $>$ 5% is in indicator for whether the two-week average share of area of ward covered in stagnant water exceeded 5%. Dry ward if mean precipitation $<$ 1000 mm precipitation. Rainy ward if $\geq$ 1000 mm precipitation. Wave, Calendar month, Ward fixed effects, and ward-level 2-week sum of precipitation, squared WBDP included in all estimations. Household covariates included are child's gender and age, and mother's age and whether secondary education or above.

Table D.4
Exploring non-linearities: Including squared WBDP

	Dependent: Test score (std)		
	All	Dry wards	Rainy wards
WBD potential	-1.133**	-1.423**	-0.346
	(0.557)	(0.645)	(1.785)
WBDP Squared	1.152	1.560	3.272
	(0.974)	(1.072)	(8.976)
Obs.	368,444	178,449	189,995
Clusters	3,842	1,669	2,173

Note: Standard errors in parentheses clustered on ward. WBD Potential is two-week average share of area of ward covered in stagnant water, $\sim$(0,1). Dry ward if mean precipitation $<$ 1000 mm precipitation. Rainy ward if $\geq$ 1000 mm precipitation. Wave, Calendar month, Ward fixed effects, and ward-level 2-week sum of precipitation. Household covariates included are child's gender and age, and mother's age and whether secondary education or above.

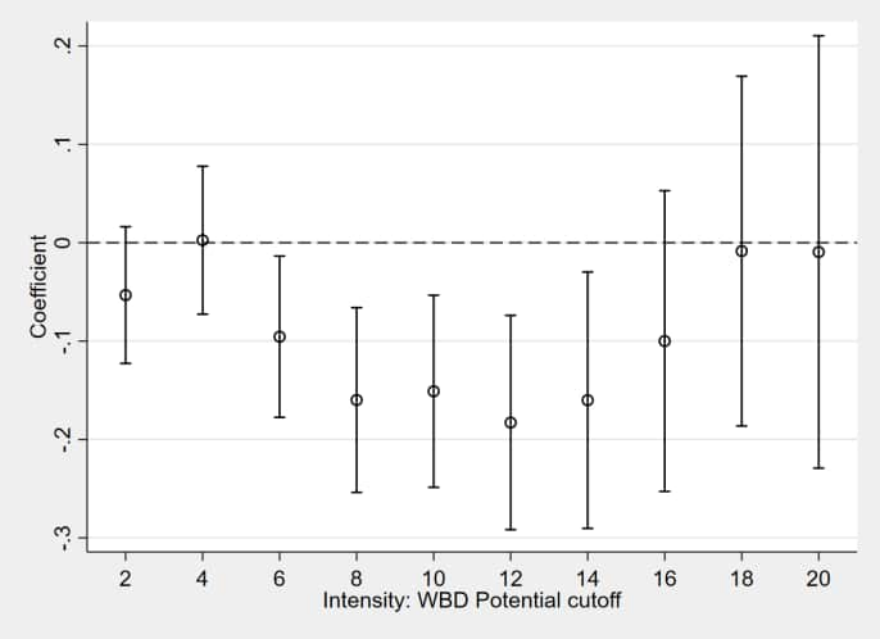

Figure D.5
WBDP as indicator with different thresholds

Note: This figure summarises DiD estimate of WBD Potential on test scores. Here, we redefine WBD Potential into a dummy which, for the first estimate, is a dummy equal to 1 for wards in waves where the simulated stagnant water share exceeds 2%. In the next estimation we define a ward as treated if the stagnant water share exceeded 4%, and so on. In all estimations we include calendar month, wave and ward fixed effects as well as individual and household covariates. Standard errors are clustered on ward-level.

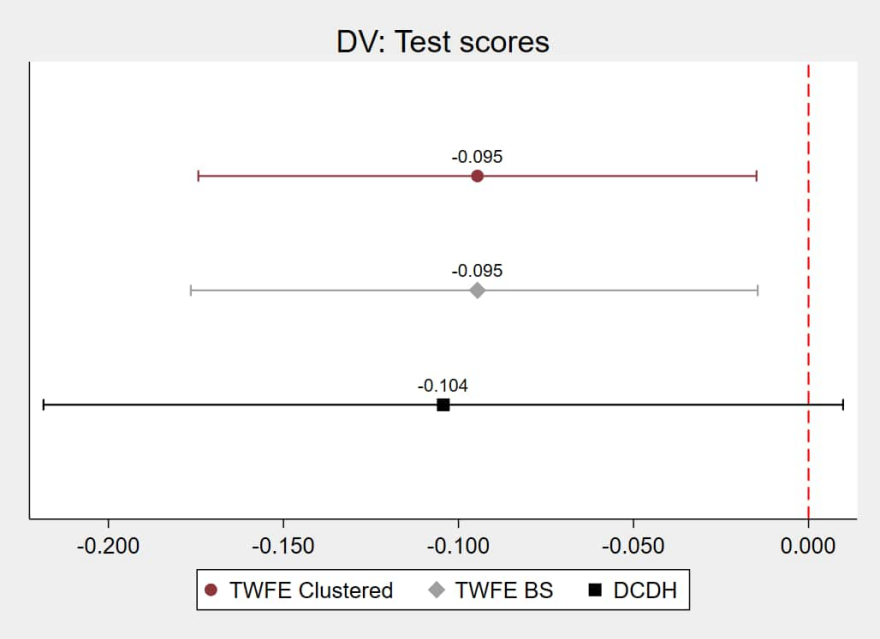

Figure D.6
Comparison of TWFE and alternative estimator robust to heterogeneous treatment effects

Note: This figure reports the DiD estimate and 95% confidence intervals for two different estimators when running the effect of WBD Potential on test scores: Two-way fixed effects (denoted TWFE) and the group-time DiD estimator robust to heterogeneous treatment effects as discussed in and developed by de Chaisemartin and D'Haultfoeuille (2018) (denoted DHDC). For this comparison, we re-frame our treatment as a binary indicator as in Table D.3, such that a ward is treated during wave in year y if the share of stagnant water exceeds 5%. Here, we redefine WBD Potential into a dummy which is equal to 1 for wards in waves where the simulated stagnant water share exceeds 5%. Since the standard errors are bootstrapped for the DHDC estimator, we also provide a comparison to the TWFE with bootstrapped standard errors instead of clustered as in our main specification. Bootstrapped standard errors produced through 999 replications.

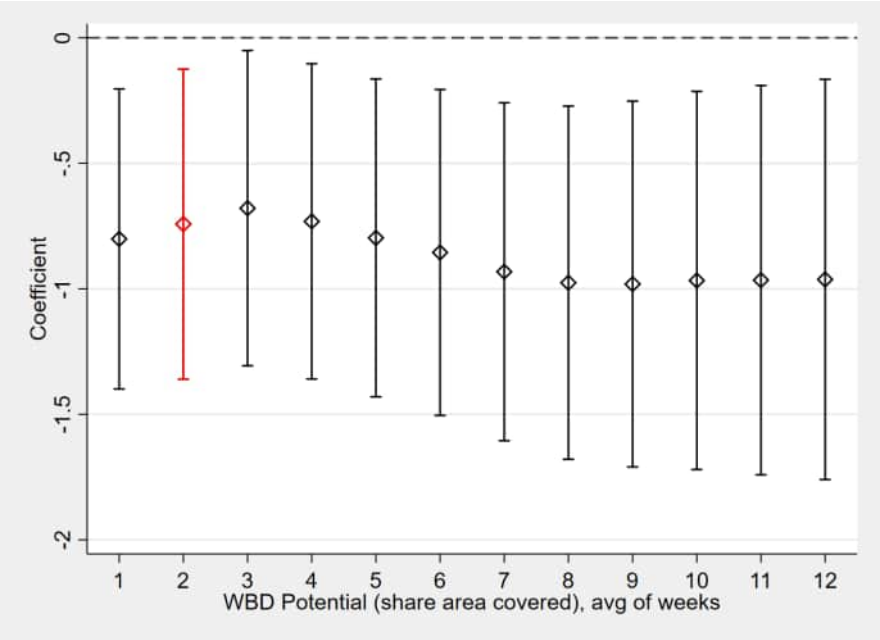

Figure D.7
Varying weeks in average measure of WBDP

Note: This figure summarises DiD estimate of WBD Potential on test scores. Here, for each estimation we redefine the aggregation average of WBD Potential. In the first estimation, we define WBD Potential as the share of stagnant water in the one week preceding the date of survey when the child is tested. In the second estimate, we instead take the average share of stagnant water in the two weeks preceding the date of survey, and so on. The highlighted red estimate and confidence intervals reflects the definition we use in our main results, where we take the average of the two weeks. In all estimations we include calendar month, wave and ward fixed effects as well as individual and household covariates. Standard errors are clustered on ward-level.

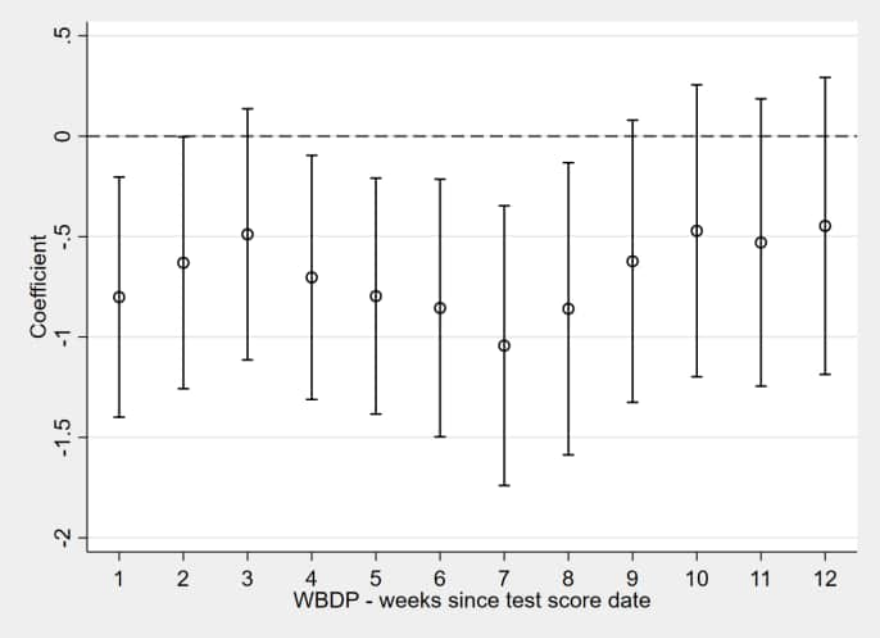

Figure D.8
WBDP, by week discretely

Note: This figure summarises DiD estimate of WBD Potential on test scores. Here, for each estimation we redefine which week we include as the measure of WBD Potential. In the first estimation, we define WBD Potential as the share of stagnant water in the one week preecding the date of survey when the child is tested. In the second estimate, we instead take the share of stagnant water in the week which starts two weeks before the date of the survey and ends the week before the survey. Similarly, in the estimation labelled "3" we define WBD Potential as the share of stagnant water in (only) the week three weeks prior the date of survey. In all estimations we include calendar month, wave and ward fixed effects as well as individual and household covariates. Standard errors are clustered on ward-level.

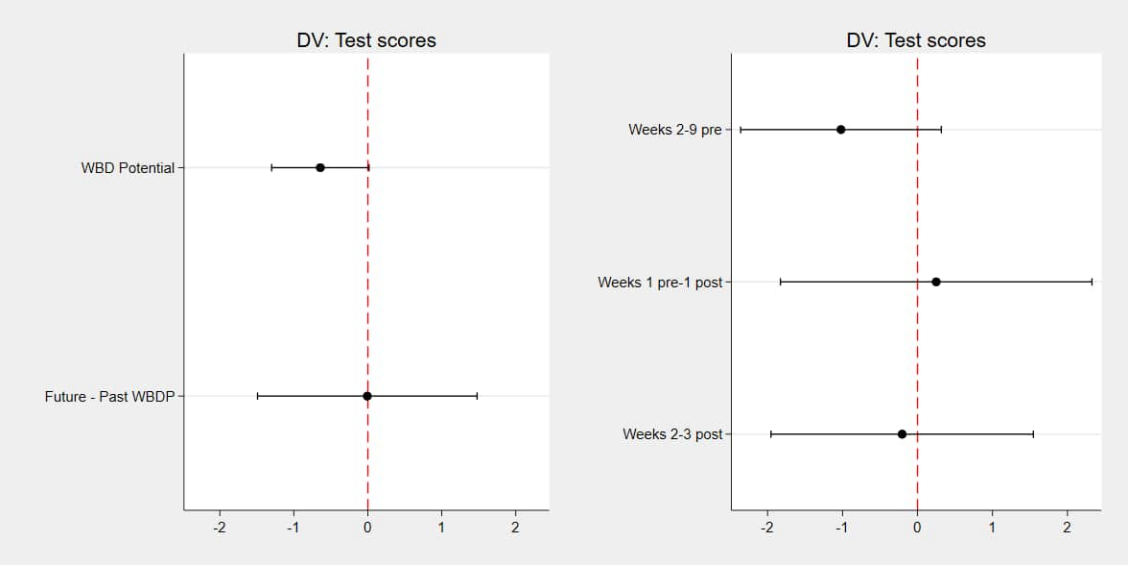

Figure D.9
Placebo: Effect of future WBD Potential

Note: This figure summarises DiD estimates of WBD Potential on test scores. Here, the two subplot represent one estimation each. On the left, we estimate with the main specification WBD Potential, and include the effect of the change in WBD Potential between the average stagnant water share *pre* survey, and the average stagnant water share *post* survey. On the right, we instead estimate longer-run means: Weeks 2-9 is the WBD Potential as the average 2-9 weeks prior to the survey. Weeks 1 pre- 1 post is the average WBD Potential the two weeks around the date of the survey, and the last estimate similarly is the average WBD Potential in the two weeks after the survey.

Table D.5
WBD Potential, simulation vs GWS

	Dependent: Test score (std)				
	All	Dry	Rainy	Rural	Urban
	Panel A. Treatment from main simulation				
WBD potential	-0.742**	-0.812**	-0.00542	-0.732**	-0.896
	(0.315)	(0.349)	(0.729)	(0.335)	(0.953)
Obs.	368,444	178,449	189,995	308,177	60,267
Clusters	3,842	1,669	2,173	3,354	488
	Panel B. Treatment from GWS satellite data				
WBD potential	-13.63	-22.57	14.45	6.820	-468.1
	(56.36)	(87.82)	(65.27)	(55.83)	(286.8)
Corr	0.13	0.16	0.12	0.02	0.52
Obs.	368,444	178,449	189,995	308,177	60,267
Clusters	3,842	1,669	2,173	3,354	488

Note: Standard errors parentheses clustered on ward. Dry ward if mean annual precipitation $<$ 1000 mm, rainy ward if $\geq$ 1000 mm. Rural and Urban also divide the sample by ward. Wave, Calendar month and Ward fixed effects included in all estimations.WBD Potential is two-week average share of area of ward covered in stagnant water, $\sim$(0,1). Panel A replicates the main result, where WBD Potential is the output from the hydrological simulation; two week average share of ward area covered in stagnant water. Panel B outputs a treatment generated from the GWS model and data, but is similarly the two week average share of ward area covered in stagnant water. "Corr" the correlation between the Panel A and Panel B definitions of WBD Potential within each estimation sample. All estimations include individual covariates: Child's gender and age, and mother's age and whether she secondary education or above, and an index for household wealth.

E Empirical strategy and results with DHS

E.1 DHS descriptive statistics

Table E.1
DHS Summary statistics

	Mean	SD	Min	Max
HH characteristics				
Birth Order	3.75	2.544	1	17
Twin birth	0.034	0.181	0	1
Girl	0.50	0.500	0	1
Age	1.93	1.416	0	4
Mother's age	29.3	7.108	15	49
Mother total fertility	4.12	2.588	1	17
type of toilet facility	23.3	4.768	11	41
Household wealth index	2.88	1.382	1	5
B. Cluster/Cell characteristics				
WBDP (%)	0.017	0.0522	0	0.512
above5	0.093	0.291	0	1
Urban share of clusters	0.22	0.414	0	1
Num. villages/wave	493.6	146.5	176	608
Num. households/wave	324.4	21.98	303	351
Mean precipitation, cell	1069.3	298.0	481.4	1682.4
Obs	21471			

Note: Raw summary statistics of mean, standard deviation, minimum and maximum across three waves of DHS: 1999, 2010m 2015. Panel A statistics across individuals and households, while Panel B displays statistics at cluster or cell, by wave.

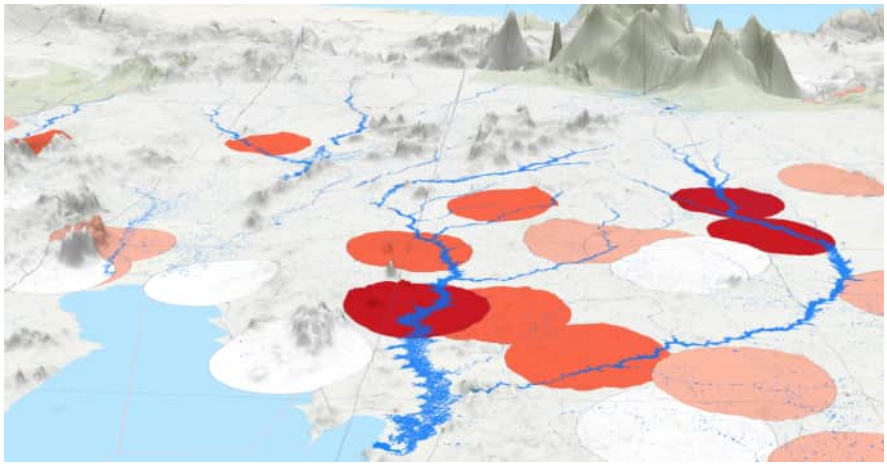

Figure E.1
WBD Potential – Aggregation with DHS

Note: To be added

E.2 Results with DHS

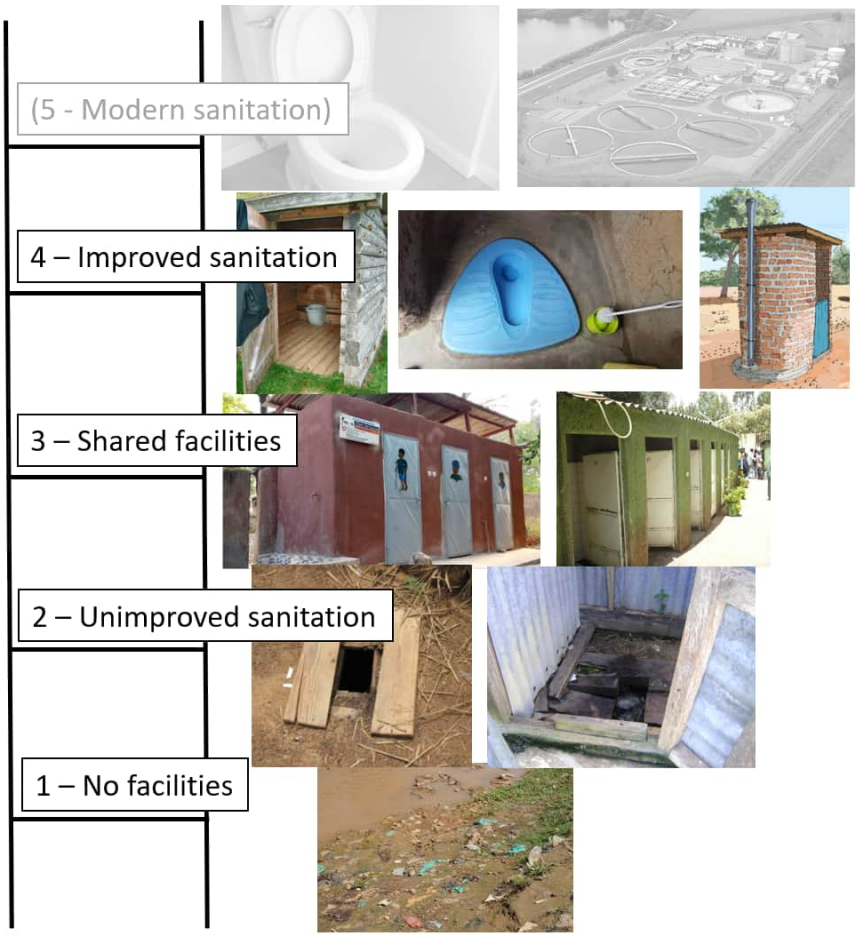

Figure E.2
UN Sanitation ladder: Examples

Note: In this image we provide examples of how the different steps in the UN sanitation may look like. In reality, there are five steps of the ladder, the highest rung being modern sanitation. However, this category is not represented in our sample so for our case we consider a four-step ladder of sanitation. In 1, there are no facilities which includes open defecation. In 2, there are unimproved sanitation facilities that are not integrated to a well-functioning sanitation system. In 3, we include shared facilities. In 4, we have improved sanitation that are well-maintained and of higher standard and technology.

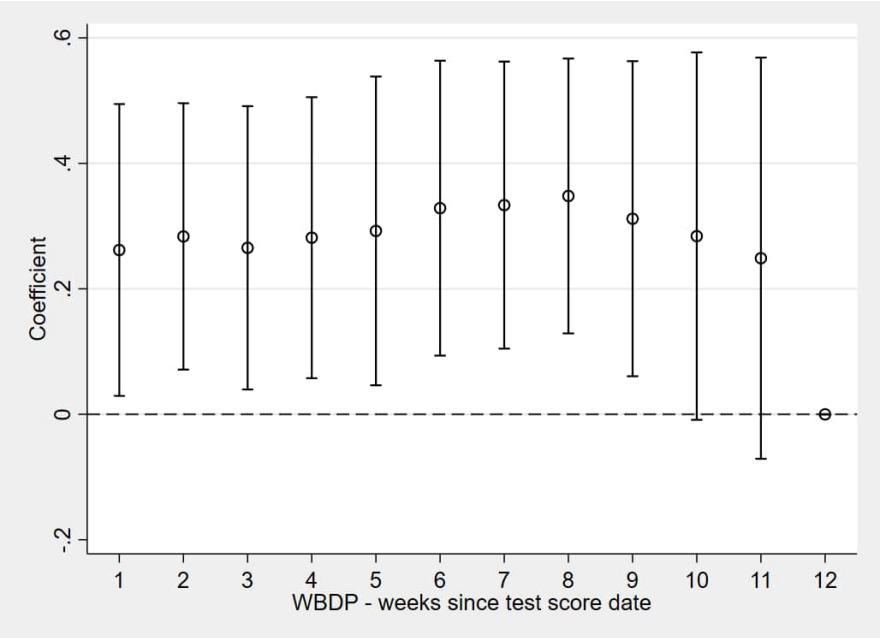

Figure E.3
WBDP by week: Effect on diarrhoea

Note: This figure summarises DiD estimate of WBD Potential on the share of children with recent diarrhoea. Here, for each estimation we redefine which week we include as the measure of WBD Potential. In the first estimation, we define WBD Potential as the share of stagnant water in the one week preceding the date of survey when the child is tested. In the second estimate, we instead take the share of stagnant water in the week which starts two weeks before the date of the survey and ends the week before the survey. Similarly, in the estimation labelled "3" we define WBD Potential as the share of stagnant water in (only) the week three weeks prior the date of survey. In all estimations we include calendar month, wave and ward fixed effects as well as individual and household covariates. Standard errors are clustered on ward-level.

Table E.2
Shared toilets and the effect of waterborne disease on diarrhoea

	(1)	(2)	(3)
	Dependent: Child has had diarrhoea		
	All	Urban	Rural
WBD Potential	0.080	0.018	0.036
	(0.12)	(0.22)	(0.21)
Shared toilet	0.027***	0.017	0.024**
	(0.0093)	(0.018)	(0.010)
WBDP*Shared toilet	0.16	0.36**	0.016
	(0.15)	(0.18)	(0.17)
Obs.	13,088	3,219	9,869
Clusters	237	98	228

Note: Standard errors in parentheses clustered on DHS gridcell level. All estimations use calendar month, gridcell and wave fixed effects. WBD Potential (%) is the average percent share of area covered in stagnant water the two weeks prior to date of survey. 'Shared toilet'=1 for households who use shared toilet facilities.

References

Anthonj, Carmen, Sophie Githinju, and Thomas Kistemann. 2018. "The impact of water on health and ill-health in a sub-Saharan African wetland: Exploring both sides of the coin." *Science of the Total Environment* 624 1411–1420.

Azman, Andrew S., Kara E. Rudolph, Derek A.T. Cummings, and Justin Lessler. 2013. "The incubation period of cholera: A systematic review." *The Journal of infection* 66 432. 10.1016/J.JINF.2012.11.013.

Beck, Hylke E, Noemi Vergopolan, Ming Pan et al. 2017. "Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling." *Hydrology and Earth System Sciences* 21 (12): 6201–6217.

Black, Robert E., Simon Cousens, Hope L. Johnson et al. 2010. "Global, regional, and national causes of child mortality in 2008: a systematic analysis." *Lancet (London, England)* 375 1969–1987. 10.1016/S0140-6736(10)60549-1.

Boedeker, Wolfgang, Meriel Watts, Peter Clausing, and Emily Marquez. 2020. "The global distribution of acute unintentional pesticide poisoning: estimations based on a systematic review." *BMC Public Health 2020 20:1* 20 1–19. 10.1186/S12889-020-09939-0.

de Chaisemartin, Clment, and Xavier D'Haultfoeuille. 2018. "Fuzzy differences-in-differences." *Review of Economic Studies* 85 999–1028. 10.1093/restud/rdx049.

Dazzi, Susanna, Iuliia Shustikova, Alessio Domeneghetti, Attilio Castellarin, and Renato Vacondio. 2021. "Comparison of two modelling strategies for 2D large-scale flood simulations." *Environmental Modelling & Software* 146 105225.

Falter, Daniela, NV Dung, Sergiy Vorogushyn et al. 2016. "Continuous, large-scale simulation model for flood risk assessments: proof-of-concept." *Journal of Flood Risk Management* 9 (1): 3–21.

Falter, Daniela, Sergiy Vorogushyn, Julien Lhomme, Heiko Apel, Ben Gouldby, and Bruno Merz. 2013. "Hydraulic model evaluation for large-scale flood risk assessments."

Farhat, Nadia, Frederik Hammes, Emmanuelle Prest, and Johannes Vrouwenvelder. 2018. "A uniform bacterial growth potential assay for different water types." *Water Research*.

Farr, Tom G, Paul A Rosen, Edward Caro et al. 2007. "The shuttle radar topography mission." *Reviews of geophysics* 45 (2): .

Harrigan, Shaun, Ervin Zsoter, Lorenzo Alfieri et al. 2020. "GloFAS-ERA5 operational global river discharge reanalysis 1979–present." *Earth System Science Data* 12 (3): 2043–2060.

Hedley, Lucy, and Roberty L Serafino Wani. 2015. "Helminth infections: diagnosis and treatment." *The Pharmaceutical Journal*, https://pharmaceutical-journal.com/article/ld/helminth-infections-diagnosis-and-treatment.

Hersbach, H, B Bell, P Berrisford et al. 2018. "ERA5 hourly data on single levels from 1979 to present." *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* 10.

Ling, Fangqiong, Rachel Whitaker, Mark W. LeChevallier, and Wen-Tso Liu. 2018. "Drinking water microbiome assembly induced by water stagnation." *Multidisciplinary Journal of Microbial Ecology*.

Magana-Arachchi, D.N., and R.P. Wanigatunge. 2020. "Ubiquitous waterborne pathogens." *Waterborne Pathogens* 15. 10.1016/B978-0-12-818783-8.00002-5.

Mandal, Badal K, Roy Chowdhury, Gautam Samanta et al. 1996. "Arsenic in groundwater in seven districts of West Bengal, India The biggest arsenic calamity in the world." *Current Science* 70 976–986, https://about.jstor.org/terms.

Pekel, Jean-François, Andrew Cottam, Noel Gorelick, and Alan S Belward. 2016. "High-resolution mapping of global surface water and its long-term changes." *Nature* 540 (7633): 418–422.

Percival, Steven L, Marylynn V Yates, David W Williams, Rachel M Chalmers, and Nicholas F Gray. 2014. *Microbiology of waterborne disease: Microbiological Aspects and Risks*. Elsevier Academic Press, , 2nd edition.

Poggio, Laura, Luis M De Sousa, Niels H Batjes, Gerard Heuvelink, Bas Kempen, Eloi Ribeiro, and David Rossiter. 2021. "SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty." *Soil* 7 (1): 217–240.

Troeger, Christopher, Brigette F Blacker, Ibrahim A Khalil et al. 2018. "Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016." 10.1016/S1473-3099(18)30362-1.

UN. 2022. "Sustainable Development Goals Report." https://sdgs.un.org/goals/goal6.

WHO. 2019. "Surveillance and outbreak management of water-related infectious diseases associated with water-supply systems." http://apps.who.int/bookorders.

Zlatanovi, Lj., J.P. van der Hoek, and J.H.G. Vreeburg. 2017. "An experimental study on the influence of water stagnation and temperature change on water quality in a full-scale domestic drinking water system." *Water Research* 123 761–772. https://doi.org/10.1016/j.watres.2017.07.019.