Working paper

International Growth Centre

Cholera in times of floods

Weather shocks and health in Dar es Salaam

Nathalie Picarelli
Pascal Jaupart
Ying Chen

July 2017

When citing this paper, please use the title and the following reference number:
C-40404-TZA-1

DIRECTED BY

FUNDED BY

Cholera in Times of Floods[1]

Weather Shocks & Health in Dar es Salaam

Nathalie Picarelli, Pascal Jaupart, Ying Chen[2]

July 31st, 2017

Abstract: Climate change is making extreme weather events more frequent around the world. Urban residents in developing countries have become more vulnerable to health shocks due to poor sanitation and infrastructure. This paper is the first to empirically measure the relationship between weather and health shocks in the urban context of a developing country. Using unique high-frequency datasets of weekly cholera cases and accumulated precipitation for wards in Dar es Salaam, we find robust evidence that extreme rainfall has a significant positive impact on weekly cholera incidence. The effect is larger in wards that are more prone to flooding, have higher shares of informal housing and unpaved roads. We identify limited spatial spillovers. Time-dynamic effects suggest cumulated rainfall increases cholera occurrence immediately and with a lag of up to 5 weeks.

Keywords: health shocks, weather shocks, infrastructure, developing cities

JEL-Codes: I15, O18, O13, Q54

[1] This research has benefited from valuable comments and conversations with Gharad Bryan, Ed Glaeser, Vernon Henderson, Aurélie Jandron, Guy Michaels, Henry Overman, as well as seminar participants at the LSE, the 7th European meetings of the Urban Economics Association, and the 2017 Columbia University Ph.D. Workshop in Sustainable Development. We are also grateful to Mark Iliffe and Tanner Reagan for sharing datasets, and the Tanzanian Regional Medical Officers and Municipal Officers for assistance with accessing confidential cholera datasets, specifically Ndeniria Swai, Alex Mkamba, Michael Gweba, and Lawrent Chipatta, as well as Theophil Malibiche. Joshua Chipman and Claire Lwehabura were also very helpful during the process of this research. This research was funded by the International Growth Centre (IGC) grant 40404. All errors are our own.

[2] Correspondence: n.picarelli@lse.ac.uk. Affiliations: London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom.

1 Introduction

Climate change will have a significant impact on the lives of the poor in the years ahead as extreme weather events such as floods, heavy precipitation and droughts are expected to become more frequent (Harrington et al. 2016). Populations at different stages of development are affected differently by the same weather variations (Dasgupta, 2010; Burgess et al. 2017). As cities in developing countries continue to urbanize at an unprecedented pace, the question of how their urban dwellers are impacted by weather shocks is becoming increasingly relevant. On the one hand, urban residents may seem better prepared than their rural counterparts against weather extremes; their livelihoods are for instance less dependent on weather phenomena. Yet, rapid urbanization has often led to *unplanned* cities with poor infrastructure, limited public service provision, and with large segments of the population living in informal settlements.

Empirical evidence is scant concerning the impact of weather shocks in developing country cities. This paper tries to make progress on this issue by looking at the effect of rainfall and flooding on cholera incidence in Dar es Salaam. Looking at health outcomes is important. Contagion is one "downside of density" (Glaeser & Sims 2015). Throughout history, cities with low-quality infrastructure and poor sanitation have been pockets of epidemics (for instance 19th century London or Paris, Kesztenbaum & Rosenthal 2016). Poor health and disease not only lower productivity in the short-term, they also hinder long-term economic growth (Well 2007).

We examine this question in the context of a cholera outbreak in Dar es Salaam in 2015 and 2016, during which almost five thousand cases were recorded. Cholera is a water and food-borne disease and its transmission is closely linked to inadequate access to clean water and sanitation facilities. Weather shocks such as rainfall can affect health through two main channels. The first one is related to direct mechanisms operating via human physiology and disease. In this case, heavy rainfall and floods can increase exposure to vibrio cholera bacteria, which survives better in wet environments (Lipp et al. 2002, Osei et al. 2010). Droughts have also been linked to cholera outbreaks, as population use unfit water for their needs (Sasaki et al. 2008, Taylor et al. 2015). The second sets of mechanisms are indirect ones, through the effect weather shocks may have on real incomes. In the case considered here, there are many ways through which floods and rainfall can reduce accessibility to work. The resulting lower income may in turn lower the consumption of health-improving goods (i.e. safe drinking water, soap), increasing the exposure to the bacteria.

Our empirical analysis uses finely disaggregated ward-level panel data containing weekly recorded cholera cases and weekly accumulated precipitation for all the municipalities in the city. We are therefore testing whether exogenous weekly rainfall variation at the ward-level affects cholera occurrence. Sorting between neighbourhoods is not an issue with the specifications chosen in the short timeframe considered. Further, we find no evidence of a relationship between diarrheal diseases and wealth. The use of high-frequency data and the fine geographical detail thus allow us to estimate with precision our relationship of interest. We focus on reduced-form specifications; different spatial and time-lag models support our choice. Dar es Salaam is a city of more than 4 million people where close to 70% of its residents live in informal settlements. Access to improved sanitation is very low and only 37% of the city has regular refuse collection (World Bank 2017). We complement our data with ward-level infrastructure characteristics (i.e. roads, footways, drains, water wells and housing informality) to understand the relationship between infrastructure quality, precipitation and cholera.

We find robust evidence that weekly accumulated rainfall and flooding leads to higher ward-level cholera occurrence. On average, a 10 mm increase in weekly accumulated precipitation leads to an increase of between 1.5% to 3.5% of weekly recorded cholera cases, significant at 1% level. The effect is found to be much larger when considering a more flexible quartile specification. On average, a single additional week of rainfall falling above the 75th percentile of the total rainfall distribution (extreme rainfall), increases the number of effective cholera cases by up to 20.3% relative to a week with very light rain (<0.30 mm). Further, we find that the impact of heavy rainfall (>75th percentile) is close to 20 percentage points higher in wards at greater risk of flooding (i.e. higher flood-prone area), all else equal. A dry week (0 mm rainfall) is also positively related to cholera incidence, but the coefficient is not statistically significant. These findings are consistent with the direct and indirect mechanisms put forward earlier. Particularly, the inundation of drains, water systems, and pit latrines greatly enhances the risk of exposure to contaminated water and food. These results are robust to alternative estimators and specifications, including an instrumental variable strategy and controlling for the spatial autocorrelation of standard-errors.

Remarkably, we find little to no spatial spillovers from precipitation in neighbouring wards. Only when considering the relative elevation of contiguous wards there is a small significant effect from precipitation in downhill wards. That is, a 10 mm increase in weekly accumulated rainfall in downhill neighbouring wards increases cholera cases in the ward by 0.01% to 0.03%. This finding is consistent with water source contamination. In contrast, our results reveal moderate time-dynamic effects. Using a distributed lag model, we find significant positive effects of past rainfall on cholera incidence to up to five preceding weeks. Contemporaneous rainfall remains the largest determinant, with a stable size close to 3% and statistical significance at or above 5% level. All lags decrease in size the farther in time.

We explore the non-linear relationship of rainfall and cholera incidence with respect to various ward-level infrastructure characteristics as well. We find the effect of weekly rainfall on cholera cases to be consistently higher for wards with larger shares of informal housing and a higher density of footways (i.e. informal roads). These results are consistent with the two possible mechanisms outlined earlier. Neighbourhoods with limited access to sanitation and low-quality infrastructure are likely to be more exposed to the cholera bacteria when surfaces are washed and drains are overflown by severe precipitation. Vulnerable populations in these wards are also more likely to suffer from negative income shocks during extreme weather events.

Our findings relate to three different bodies of research. The first one is a large economics literature looking at weather shocks and economic events mostly in advanced economies (Munshi 2003; Miguel et al. 2004; Barrios et al. 2010;), and including health outcomes of human populations (Deschenes & Greenstone 2007; Deschenes & Moretti 2009; Burgess et al. 2011; Deschenes 2011). Our paper follows their empirical methods (for a review see Dell et al. 2014). The second is the large literature in development economics and public health studying policy interventions, mechanism of transmission and health outcomes in developing countries (Banerjee et al. 2004; Miguel & Kremer 2004; Dunkle et al. 2010; Penrose et al. 2010; Devoto et al. 2012). Here, we contribute to that literature by focusing on urban areas and by studying general mechanisms beyond specific policy interventions. Our findings use robust econometrics techniques to discern the relationship between disease, infrastructure and weather shocks. Finally, we particularly contribute to the nascent literature in urban economics studying the effects of weather phenomena on urban areas in developing country cities (Kocornik-Mina et al. 2015; Glaeser & Henderson 2017; Henderson et al. 2017). To the best of our knowledge, our paper is the first to empirically study the effect of weather shocks on disease transmission within a city of a developing country. Our findings have important policy implications as extreme weather events become more frequent in the next decades. Cities in developing countries need to address infrastructure gaps to contain the risk of recurrent epidemic outbreaks in vulnerable populations and neighbourhoods. Investing in resilient infrastructure, with the proper servicing of informal settlements or related measures such a regulating waste dumping may prove to be more beneficial in the long-term than the use of short-term palliative measures during outbreaks. Evidence on large-scale policy interventions in urban areas is still limited and more is needed to understand how to prevent contagion of treatable diseases in developing cities.

This paper is organized as follows. The next section formalizes the relationship between health and weather shocks, specifically here cholera and rainfall in cities. Section 3 describes the context of Dar es Salaam, the cholera disease and the data. In section 4 we present our empirical strategy. Section 5 presents the estimates of the effect of rainfall and flooding on cholera incidence and different extensions and robustness checks. Section 6 concludes.

2 Theoretical framework: Weather & Health

In this section, we describe a simple theoretical framework to examine the various channels through which weather shocks (i.e. here, heavy precipitation) can affect cholera incidence in an urban setting. It relies on endogenous health models in a simplified fashion (for more details see Becker 2007; Deschenes 2012; Burgess et al. 2017).

Cholera is an acute diarrheal infection caused by ingestion of food or water contaminated with the bacterium vibrio cholera. It can affect both children and adults and can kill within hours if left untreated. The main reservoirs of the cholera bacterium are people and warm salt water bodies such as estuaries and coastal areas. Cholera transmission is closely linked to inadequate access to clean water and sanitation facilities (WHO 2017)[3]. Here we assume that weather shocks such as extreme rainfall, droughts and flooding can affect human health (i.e. cholera prevalence) both directly (through higher exposure to the bacteria, for instance), and indirectly (due to the negative income-shocks that may arise through the weather shock).

Consider a city with a large number of agents indexed by *i*. Agents seek to maximize their lifetime utility u_i depending on consumption c_{it} and health status h_{it}. These two are complements. This city is partitioned into several neighbourhoods and each agent lives in a given neighbourhood (or ward) indexed by *n*. We assume that agents are exogenously allocated to neighbourhoods with different characteristics such as better infrastructure, and the latter are exogenously determined. It follows that agents have limited scope for affecting local infrastructure, as well as other public goods and services, and take these as given (mobility is limited in the short-term scenario considered here).

[3] http://www.who.int/mediacentre/factsheets/fs107/en/, accessed on 28 June 2017.

Agent *i*'s health status, resident of neighbourhood *n* at time *t*, (h_{int}), is thus determined by her consumption c_{it} and random health shocks z_{int}. We do not specify a precise relationship between health outcomes and consumption, but it is assumed that an individual can improve her health by increasing her consumption, particularly by purchasing health-improving goods. All told, the agent's health status is thus given by:

$$h_{int} = h_i(c_{it}, z_{int}) \qquad (1)$$

where $h_i(\cdot)$ is increasing in c_{it} and decreasing in z_{int} (adverse health shocks). The function h_i is unrestricted and can differ across heterogeneous individuals. The vector z_{int} includes weather-shocks w_t such as flooding, droughts, and heavy rainfall or temperature extremes. We assume further that the effect of the weather shock is conditional on the quality of local neighbourhood infrastructures (q_n) such as access to drinking water and waste water systems, or road and pavement material. The adverse health shock is thus a function of the weather shock that varies with local infrastructure, $z_{int}(w_t(q_n))$.

Consumption[4] c_{it} is financed through labour income in period *t*, which depends on the agent's productivity a_{it}. Productivity is agent-specific but also depends on weather shocks w_t:

$$a_{it} = a_i(w_t) \qquad (2)$$

Here the effect on productivity can stem from the weather shock's impact on z_{int}, hampering the agent's ability to work efficiently, or deter accessibility to jobs or other factors of production (i.e. machinery, location)[5]. A given weather shock w_t thus affects an agent's health status through both consumption (via productivity) and health idiosyncratic shocks (via z_{int}). In other words, there are two fundamental mechanisms through which weather shocks (extreme precipitation for the purpose of our empirical analysis) can potentially harm an agent's health status here.

First, through direct health effects: random shocks z_{int}, enter the agent's health status directly as in equation (1). That is, holding constant the agent's income, location and consumption decisions, we expect a negative weather shock to impact this agent's health adversely (w_t impacts z_{int} in the language of our model). In the case considered here, heavy rainfall and flooding can directly impact one's health status through greater exposure to and contact with contaminated water and food. An extensive public health literature discusses the potential for cholera prevalence in wet environments and in cases of heavy precipitation and flooding (see for example, Osei et al. 2012). Further, the magnitude of the effect can be expected to depend on the relationship the weather phenomena and the disease pathogens have with local infrastructure characteristics. Cholera strives in stagnant water and poor hygiene conditions as explained earlier.

The second, more indirect mechanism through which weather can affect health in this model is through the agents' productivity in equation (2). This term depends on weather shocks that may affect the

[4] We assume the consumption good is produced using an aggregate production function that requires capital and labour inputs; it exhibits decreasing return to scale. Goods can be bought and sold at the market price, which is exogenously determined. Agents are subject to budget constraints in each period which are a function of the labour income (in turn dependent on productivity and adverse weather shocks), as well as prices and quantities of goods consumed. We assume imperfect credit and savings markets which prevents agents from smoothing their consumption in time.

[5] We assume $a_{it} = a_i(z_{int}(w_t), Q_{nt}(w_t))$, where Q_{nt} refers to complements to work such as accessibility to jobs, machinery or location, that can be affected by weather shocks.

agent's ability to work via z_{int}. Flooding and heavy rainfall may also significantly affect work-places and accessibility in contexts where poor roads and infrastructure is widespread (see footnote 4). Reduced productivity can translate in lower earnings and reduced consumption of healthier quality goods such as clean water and fresh food. The dependency of productivity and hence, labour income, on this type of shocks is extremely likely in low-income countries where informal jobs dominate employment. It is estimated that close to 80% of jobs in the services sector are informal in Tanzania (UNDP 2015). Note that this assumes imperfect credit and savings markets preventing agents from smoothing their consumption when hit by economic hardship. Given the Tanzanian context, this assumption does not seem unrealistic.

The main implication of this exercise is to expect an increase in cholera cases due to a weather shock such as extreme precipitation and flooding. The increase should be larger in neighbourhoods with poorer infrastructure. Conversely, the impact of heavy rainfall should be mitigated in areas with a supply of higher-quality local public goods. Linking this section with our empirical analysis, we expect to see non-linear effects of rainfall on cholera occurrence.

Finally, one additional implication of this simple formalization concerns policy interventions. In the face of potential weather shocks, any agent *i* would seek to minimize the damage that the negative shock has on their utility, $u_i(c_{it}, h_{it})$ by consuming health-improving goods or potentially by reallocating resources between periods. This latter option here is limited due to credit and savings constraints. These potential shock-minimizing strategies have strong implications for policy. In the theoretical framework considered here, governments can reduce the adverse effect of weather on health outcomes by directly increasing the quality of infrastructure that is related to the pathogens' transmission (i.e. pavement, sanitation, water drainage, sewage), and thus directly limit the potential effect of an adverse health shock, z_{itn}. But they can also intervene by supporting the agent's shock minimization strategies through subsidized health goods or direct transfers.

3 Background & Data

This section provides further details on cholera-specific characteristics as well as Dar es Salaam's context. It also describes the data in detail and provides basic summary statistics.

3.1 Cholera

Cholera is an acute diarrhoeal infection of fecal-oral transmission. It is caused by the ingestion of food or water contaminated with the bacterium vibrio cholera. It takes between twelve hours and five days for a person to show symptoms after ingesting contaminated food or water. It can affect both children and adults and can kill within hours if left untreated; there is a 50% death rate if untreated, but all deaths are avoidable otherwise. Main treatments include antibiotics and Oral Rehydration Salts (ORS). Roughly 1.3 to 4.0 million cases are recorded worldwide every year, and the disease is endemic to many parts of sub-Saharan Africa and South Asia (WHO 2017).

There are multiple pathways for cholera transmission (Clasen et al. 2007). The disease is closely linked to inadequate access to clean water and sanitation facilities. Risk factors are also considered to be high population density and crowding, all of which are often common in urban slum areas (Penrose et al.

2010). Cholera incidence has been found to be highest in highly urbanized areas (Osei & Duker 2008; Sur et al. 2005). The main reservoirs of the cholera bacterium are people and warm salty water bodies such as estuaries and coastal areas. Global warming and rising sea levels are believed to create a favourable environment for cholera bacterium growth (WHO 2017). Heavy rainfall and flooding have all been associated with a higher likelihood of cholera outbreak. Surface runoff from point sources (pit-latrines, waste dump site, water wastes) may cause increased contamination of water sources, while stagnation and slow flowing of waterways may lead to increased exposure to cholera vibrios (Osei et al. 2010).

3.2 Dar es Salaam

Dar es Salaam is one of the largest cities in eastern Africa. It is located in the east of Tanzania by the Indian Ocean. Its urban population grew at 6.5% yearly between 2002 and 2012 (Wenban-Smith 2014), and today the city counts with more than 4.4 million people. Since 2016, it is divided in five municipal districts: Ilala, Temeke, Kinondoni, Kigamboni and Ubungo[6]. These municipalities are further divided up into 90 wards.

The rapid pace of urbanization has led to the city suffering from large infrastructure deficits. Close to 70% of Dar es Salaam's residents live in informal settlements without adequate access to clean water, proper drainage system and waste collection (UN-HABITAT 2010; Natty 2013). Only 13% of the city's residents have adequate sewage systems and 37% of the solid waste is properly collected. The World Bank (2015) estimates that only 50% of residents have access to improved sanitation. The most common form of improved sanitation is improved pit-latrines (other forms are rare). About two-thirds of households in the city share their toilet facilities. Access to piped water is also very limited, with only 17% of city-centre dwellers having piped-water.

Dar es Salaam's geography and coastal location makes it vulnerable to climatic hazards, particularly floods, sea level rise and coastal erosion (Kebede and Nicholls 2010). There are two rainy seasons every year, the short (October to December) and long (March to May) seasons, and average annual precipitation is above 1,000 mm. The combination of high informality and climatic vulnerability makes flood risk one of the main challenges for sustainability, exposing infrastructure and residents to safety and health hazards from vector-borne diseases such as malaria and cholera (World Bank 2017).

Cholera has been endemic in Tanzania since the 1970s and Dar es Salaam has historically been the most affected region[7]. During the 2015-2016 outbreak there were over 24,000 cases recorded nationally, with more than one fifth in Dar es Salaam (Figure A1 in appendix)[8]. Previous outbreaks occurring between 2002 and 2006 reported over 30,000 cases nationally, with nearly 18,000 in the capital city (WHO 2008). Given the city's poor sanitary conditions, high population density, lack of access to safe drinking water, and limited drainage, continuous heavy rainfall makes stagnant and unsanitary water a widespread health risk for common water borne diseases. The lack of storm water drains, frequently blocked by

[6] In the analysis, we only use 3 municipal districts as these were the ones that existed at the time cholera cases were recorded during the last outbreak. Ubungo and Kigamboni were created in 2016 from dividing Kinondoni and Ilala further so this does not impact our findings in any way.

[7] The largest cholera epidemic in Tanzania to date took place in 1997 where 40,000 cases were reported. The epidemic is said to have started in Dar es Salaam. Dar es Salaam has had the most cholera cases since 2002 of all regions of the country (Penrose et al. 2010).

[8] All tables and figures indexed by A# are in Appendix I.

unregulated waste dumping, means that heavy rainfall quickly leads to flooding and contaminates water wells (Pan-African START 2011).

3.3 Data

To examine the relationship between weather variation and cholera incidence outlined in our theoretical framework, we collect data from several sources and put together a comprehensive ward-level panel dataset for each week between the first week of March 2015 and the first week of September 2016. The choice of the timeframe is data constrained – that is, we use the first week for which precipitation data is available and the last week for which cholera cases were recorded to avoid measurement error from unrecorded cases. We cover all the 90 wards of the city[9]. The use of high-frequency data and the fine geographical detail allow us to estimate with precision our relationship of interest. The basic panel thus consists of weekly cholera cases registered according to the ward of residence. We combine this data with weekly accumulated precipitation and weekly air-temperature in these wards. Further, we add data on ward-level infrastructure, geographical characteristics (i.e., elevation, flood-prone surface) and population (census). We outline below the different data sources. Main summary statistics are in Tables 1 and 2.

3.2.1 **Cholera cases**

The key data in this analysis are the new ward-level cholera cases collected from the Regional Medical Office and Municipal Health Officers for all the wards of Dar es Salaam and covering the entire 2015-2016 outbreak. The data was registered daily for each individual presenting symptoms of severe diarrhoea in a medical facility. It includes basic socio-demographic characteristics (age, sex) of the individual, the ward and sub-ward of residence, as well as the date of the first symptom and registration at the hospital. Cases were tested for the vibrio cholera bacteria, and the dataset also includes lab results. We exclude all cases tested negative and focus on effective cholera cases only[10]. No positive case is reported earlier than mid-August 2015 (epidemic week zero). The outbreak officially lasted from August 2015 until May 2016. We aggregate the daily cases by week to account for the fact that the incubation period is between 12 hours to 5 days.

Measurement error is a potential problem. The biggest threat concerns the possibility that not all cholera cases are reported in the non-outbreak period. It is also possible that not all registered cases during the peak of the epidemic are effective cholera cases (see footnote 8). There are mitigating factors against both these risks. First, cholera is one of the few diseases that require reporting to the World Health Organization (WHO) by the International Health Regulations as it can quickly spread if left untreated and result in explosive outbreaks. This implies careful monitoring of the disease as well as frequent laboratory testing. Further, we focus our analysis during an outbreak where monitoring is more likely to be enforced. Finally, our baseline estimates are weighted by the population of the ward, to account for the difference in precision concerning cholera measurement from larger and smaller

[9] Since mid-2016 there are 5 municipalities in Dar es Salaam as two municipalities were further sub-divided. We use the original administrative units at the time of the outbreak in our regression analysis for simplicity and coherence with the recorded cholera dataset. This should not affect any of the results.

[10] We include both positive and untested cases. Most untested cases are at the peak of the outbreak when all patients presenting symptoms are treated as cholera patients. Measurement error is possible but should be limited as tests are frequently carried, particularly at the beginning and end of the outbreak period.

populated wards. While bias from measurement error in our dependent variable is still possible, it should not be large.

Overall, close to 5 thousand cases of cholera were reported positive in Dar es Salaam in the period analysed (4964 of total 5698 tested), with the bulk taking place during the first 10 weeks of the outbreak (Figures 1 & 2). On average, during the period covered there were 0.72 effective cases weekly per ward. The number is larger during the first 10 weeks of the epidemic (3.16) as well as the first 20 weeks of the epidemic (2.54) (Table 2). Cholera cases were more pronounced in Kinondoni and Ilala, reporting totals of 2428 and 1796, respectively. Temeke was the least affected (Figure A2). Most cases took place within 15 km from the Dar es Salaam CBD (Figure 3); only 2 of the 90 wards reported zero positive cases throughout the period.

3.2.2 Weather & Geography

Rainfall - The weather datasets in this paper are from NASA. The daily precipitation measures by ward are derived from the Integrated Multi-satellite Retrievals (IMERG) for Global Precipitation Mission (GPM), where rainfall is comprehensively measured at the highest accuracy and finest spatial resolution to date (Huffman, 2016). We use the near-real-time total daily rainfall defined as precipitation accumulated in the past 24 hours by 23:59pm (Coordinated Universal Time) of each day. We calculate weekly accumulated precipitation from the daily data. In terms of the spatial resolution, rainfall is measured at squared pixels of 0.1 ◦ × 0.1 ◦ (roughly 120km2).

As ward boundaries are irregularly shaped, we compute ward-level daily rainfall accumulation by weighting recorded rainfall with the ward overlay with satellite pixels. We first union these two layers to create polygons at the ward-pixel level. These ward-pixel polygons all have consistent rainfall measurement, and their respective area is computed. We then sum up the ward-pixel rainfall measures for each ward by weighting by their ward area share. This gives us the area-weighted weekly rainfall accumulation at the ward level (Figure A3). The choice of focusing on rainfall accumulation stems from the fact that precipitation is 'readily stored' in the soil, tanks or water wells. It is stagnant water that might breed cholera and thus, measuring average rainfall instead would fail to take this important dimension into account.

Because satellite data are subject to error (Dell et al. 2014), we also use an additional and independent gridded data set to address potential measurement issues and obtain instrumental variables (IV) estimates. We use precipitation obtained from IMERG's predecessor technology, the Tropical Rainfall Measuring Mission (TRMM) (Goddard Earth Sciences Data and Information Services Center 2016). Despite the fact that TRMM is less accurate (Shari et al. 2016; Chen and Li. 2016; Wang et al. 2017) and its resolution coarser, it has been widely used since 1997. Its algorithm intercalibrates all existing satellite microwave precipitation measures, microwave-calibrated infrared satellite estimates, and precipitation gauge analyses. The near-real-time data is chosen over the production data as it is recommended for flood and crop forecasting (NASA Precipitation Measurement Missions 2016). The instrumental variables approach is motivated by the fact that both satellite measures assign weather variables to grid points and contain measurement error in their 'true' representation of rainfall. In that case, the IV estimates can correct for measurement error bias under the assumption that errors in both variables are uncorrelated (Burgess et al. 2017).

Temperature - The daily temperature data also comes from NASA. We obtained near-surface air temperature (i.e., temperature at the height of most human activities) from the FLDAS Noah Land Surface Model (McNally 2016). The spatial resolution of this dataset is also $0.1^\circ \times 0.1^\circ$, so ward-level daily temperature is computed similarly to rainfall above. Average weekly temperature is later computed at the weekly level.

Elevation – The elevation calculation is based on the Japan Aerospace Exploration Agency (JAXA) global digital surface model. The measurement is at 30-meter spatial resolution, based on the most precise global-scale elevation data at this time acquired by the Advanced Land Observing Satellite. Mean ward-level elevation is computed across all grids that fall inside each ward.

Flood-prone surface – To estimate the surface of a ward that is prone to flooding, we use data collected by the NGO Dar Ramani Huria (RH) in OpenStreetMap (OSM) format. Using community-based mapping RH is able to create highly accurate maps of infrastructure and flood-prone areas in Dar es Salaam. We complement their detailed mapping of drainage, waterways and wetlands with GeoFabrik's OSM data for missing wards. The data is less accurate but allows us to have a larger coverage. We then use InaSAFE[11] to model build-areas prone to inundation and calculate the total share of the ward area that is flood-prone. We compare our estimates to the more precise-ones of RH for available wards. The pairwise correlation is 0.81.

Basic summary statistics of weather and geographical variables are displayed in Table (2). The average weekly rainfall in Dar es Salaam according to the meteorological agency amounts to 20.6 mm. This is consistent with our weekly accumulation from both TRMM and GPM's measures. On average, in the period covered there were 20.1 mm of accumulated rainfall weekly, with a median of 2.9 mm. The rainiest month is usually April, which is seconded by our dataset. There is little spatial variation of temperature across the city's' wards, the average recorded weekly is 26.7º C with a standard deviation of 0.37º C. On average 10% of the area of a ward is prone to flooding, but there are significant disparities across wards (the standard deviation being 16%).

3.2.3 Infrastructure & Population

Infrastructure - Infrastructure data at the ward-level is also obtained from data collected by RH's in OSM format, and complemented with GeoFabrik's for missing wards. We focus on the following characteristics which are likely to be correlated with cholera incidence: drains, roads, footways (i.e. unpaved roads) and water wells. For the first four variables, we use their density, calculated as the number of km per square km. Aside from roads where we can distinguish between roads and footways, we have no specific measure of quality of the infrastructure. A general assumption is to think that a higher density of roads and drains reflect higher-quality infrastructure, while a higher density of footways reflects lower-quality. The distinction in practice is hard to make, particularly for drains. Anecdotal evidence suggests drains often get clogged by unregulated waste dumping due to heavy rainfall and quickly contaminate

[11] InaSAFE is a free software that produces realistic natural hazard impact scenarios. It was developed by the government of Indonesia, the Australian government and the World Bank. For more details see http://inasafe.org/ (last accessed on July 21st 2017).

surfaces. We are thus agnostic concerning the expected signs of these coefficients. We have unfortunately no data on sewerages[12].

We obtained a dataset of formal and informal plots from the municipalities' database of surveyed plots, and are then able to estimate the share of the ward's area that houses informal settlements. Not all municipalities have mapped their informal plots fully[13] which explains the smaller sample when using this data. For the wards for which we have information, 34% of the total areas are on average informal. The large number reflects the fact that 70% of the population of Dar es Salaam lives in informal settlements.

Population- We make use of the population data from the Census 2012 to weight our regressions by ward population size. The interest in this is twofold. First, cholera incidence in wards with large populations is likely to be more precise, so weighting corrects for heteroskedasticity associated with these differences in precision (Burgess et al. 2017). Second, rather than on the average ward, the results reveal the impact on the average person, which is more meaningful here. We also use this data to calculate ward-level population density. The average ward of Dar es Salaam was populated with 48.5 thousands people in 2012; population density was 11.53 per square km (Table 1).

4 Empirical Strategy

In this section we describe the econometric methods we use to estimate the effect of precipitation on cholera incidence. As the relationship between rainfall and new cholera cases is expected to be non-linear, we adopt both parametric and flexible non-parametric specifications. We begin by presenting specifications measuring the contemporaneous effect of precipitation. We then consider models allowing for the effect of rainfall to be associated with local public goods provision and other ward characteristics. We also assess the importance of the spatial spillovers of precipitation. Lastly, the last sub-section details a more general dynamic model including various precipitation time lags.

4.1 Contemporaneous Effects

To quantify the contemporaneous effect of rainfall on cholera incidence in any given ward and week, we begin by estimating a baseline panel log-linear model relating the logarithm of cholera cases[14] to weekly rainfall accumulation for this ward:

$$C_{wmt} = \alpha . R_{wt} + \gamma . T_{wt} + \mu_w + \delta_t + \theta_m . t + \varepsilon_{wmt} \qquad (3)$$

where C_{wmt} is the outcome variable (log of cholera incidence) in ward *w* in week *t*. The key explanatory variable of interest is R_{wt}, measuring weekly accumulated rainfall. We also control for ward daily temperatures measured as weekly averages (T_{wt}) as temperature variation is likely to be correlated with rainfall variation. Since our focus is on precipitation and spatial variation in temperature in Dar es Salaam is limited, we model a linear temperature effect. The specification in equation (3) also includes a full set of ward fixed effects, μ_w, absorbing unobserved time-fixed ward idiosyncratic characteristics.

[12] Basic sanitation data in Table 1 is obtained from the 10% sample of the Census 2012. Unfortunately, these are only used in the descriptive section because of the lack of consistency in the sample.
[13] Only the Municipality of Kinondoni has.
[14] Since no cholera cases are recorded in several wards and weeks in our sample, we add one to the number of new cholera cases and take the logarithm of that expression. In mathematical terms: $C_{wmt} = \ln(1 + C_{wmt})$.

Permanent differences in access to healthcare for instance will therefore not confound the estimates. Their inclusion also addresses the potential issue of sorting across neighbourhoods. We also include week fixed effects, δ_t, to control for time-varying influences common across wards. The equation also includes municipality linear time trends to account for time-varying factors that differ across administrative boundaries and affect health. We also estimate equation (3) with municipality-week fixed effects to flexibly control for unobserved municipality-wide time shocks. We use only three municipalities in the analysis as these were the administrative divisions existing at the time of data collection. Further, the main three hospitals are located in these municipalities. As shown later, our estimations across these specifications are consistent and robust. ε_{wmt} is an error term clustered at the ward level. Finally, we weight our regressions by ward population as explained earlier. Unweighted regressions are in Appendix I (section II.1). Results are unchanged.

To take into account non-linear relationships more rigorously, we also estimate contemporaneous rainfall effects using the following flexible model:

$$C_{wmt} = \sum_{k=1}^{4} \beta_k . 1\{R_{wt} \text{ in quartile } k\} + \phi . T_{wt} + \mu_w + \delta_t + \theta_m . t + \eta_{wmt} \tag{4}$$

where the independent variables we are mainly interested in capturing are the distribution of weekly rainfall in Dar es Salaam. The regressors $1\{R_{wt}$ in quartile k$\}$ calculate whether the total amount of rainfall R_{wt} in week *t* and ward *w* was in the first, second, third, or fourth quartile of the rainfall distribution of our study period. We estimate a separate coefficient on each of these quartile variables and treat the second quartile as the omitted reference category. The other regressors are as defined in equation (3). This approach has two benefits. The first one is to allow for more flexibility in the response function. The second one, more relevant here, is that it also allows us to specifically distinguish the effect of intense and light rain. The upper quartile (>75th percentile) is generally used as a proxy for flooding (Chen et al. 2017).

The parameters in equations (3) and (4) are thus identified from ward-specific deviations in rainfall from the ward average remaining after controlling for week fixed effects and municipality linear trends. Given the relatively short time period of analysis we argue that this variation is as good as exogenous and uncorrelated with other unobserved determinants of cholera incidence.

Equations (3) and (4) make several important assumptions about the effect of rainfall on cholera. First, they assume that the impact depends on weekly accumulation alone. It ignores the possibility of within week variation in rainfall having an effect on health. In addition, equation (4) assumes that the impact of rainfall is constant within a given quartile. While this might be restrictive, we estimate separate quartile coefficients to improve on equation (3) and its parametric assumptions. Third, by estimating contemporaneous effects, we assume that past weekly rainfall does not affect health outcomes. We also ignore the possibility of neighbouring wards' rainfall influencing a given ward's cholera outcomes. We relax some of these assumptions in what follows.

A final concern is spatial dependence. In this case, within-cluster correlations in the specification of the error covariance matrix (i.e., standard-errors clustered at the ward level) may not be enough (Barrios et al. 2012). To account for this issue we also compute equations (3) and (4) using Conley (1999) spatial standard-errors[15]. The implicit assumption is that spatial dependence is linearly decreasing in the distance from the wards centroids up to a cutoff distance, for which we chose 50 km based on Dar es Salaam's extent. This technique ensures that uncertainty in α and β is adjusted to account for heteroscedasticity, ward-specific serial correlation, and cross-sectional spatial correlation. The size of our main point estimates are generally unchanged as is statistical significance. We consider these results as robustness checks in Appendix I section II.2.

We are interested in reduced-forms here. However, we are conscious that the true (unknown) relationship may include some time dependency in the dependent variable. That is, past cholera may determine contemporaneous cholera. To test the validity of our fixed-effects model, we compute a dependent-lagged model instead in Table A15 in Appendix I. While we find the effect of lagged cholera cases significant, and positive up to 5 weeks, the size of the coefficient for contemporaneous precipitation always remain stable and statistically significant.

4.2 Non-linear Effects and Spatial Spillovers

Ward characteristics, such as population density or the number of water wells, may affect the impact of rainfall on health as outlined in our theoretical framework. To account for this possibility, we estimate variations of equation (3) that include interactions between rainfall and ward features. While local public goods are not exogenously allocated to wards, there are several reasons to believe this is not a problem here. First, the use of ward fixed effects should deal with neighbourhood sorting. Further, the lack of proper infrastructure is widespread in Dar es Salaam and public health evidence suggests households from all income-levels may be affected by cholera. Using the 2015-16 Tanzania Demographic and Health Survey and Malaria Indicator Survey (DHS) we test the relationship between income, wealth, and incidence of diarrhoeal diseases in the city (Appendix II). We find no evidence in favour of a wealth bias regarding the risk of contracting a diarrheal disease.

We also measure whether contemporaneous precipitation in neighbouring wards affect cholera cases in a given ward. We focus on first contiguity wards and consider total neighbouring accumulated rainfall to begin with. We then distinguish between rainfall recorded in uphill and downhill neighbouring wards. We calculate the average elevation of each unit and classify as uphill the neighbouring wards with a relatively higher elevation. Downhill neighbouring wards have a lower or equal average elevation. Formally the model we estimate is as follows:

$$\mathrm{C_{wt}} = \rho_1 . \mathrm{R_{wt}} + \rho_2 . \mathrm{UR_{wt}} + \rho_3 . \mathrm{DR_{wt}} + \pi . \mathrm{T_{wt}} + \mu_w + \delta_t + \varsigma_{wt} \quad (5)$$

where $\mathrm{UR_{wt}}$ and $\mathrm{DR_{wt}}$ measure weekly accumulated rainfall in uphill and downhill neighbours, respectively. The other regressors are defined as in equations (3) and (4).

4.3 Dynamic effects

The empirical approaches discussed so far do not address the possibility of a dynamic relationship between precipitation and new cholera cases. Rainfall in one week might result in increased cholera incidence in the following weeks due its incubation period and the manner in which the disease spreads.

[15] We use the Stata code developed by Fetzer (2010) and Hsiang (2010).

This delayed response would imply that the contemporaneous estimates from equation (3) underestimate the true impact of rainfall. We investigate this possibility by including a distributed lag structure in our models:

$$C_{wt} = \sum_{j=0}^{J} \lambda_j . R_{wt-j} + \rho . T_{wt} + \mu_w + \delta_t + \zeta_{wt} \qquad (6)$$

This model allows the effect of rainfall up to *J* weeks in the past to affect cholera incidence in a given week. In equation (6), the total dynamic effect of rainfall on cholera cases is obtained by summing the coefficients on the contemporaneous and lagged rainfall variables. Different lag structures potentially generate different estimates of the dynamic causal effect. As a consequence we experiment with several time lags and use up to 5 lagged weekly accumulated precipitation in our regressions.

5 Main Results

This section presents our empirical results on the relationship between precipitation and cholera incidence. We begin with discussing baseline contemporaneous estimates of both rainfall and flooding. We then assess the importance of non-linear effects, spatial spillovers and measure dynamic effects last.

5.1 Baseline Effects

Our baseline results concern the effect of rainfall and precipitation on weekly-ward cholera occurrence. Tables 3-6 report baseline estimates of population-weighted regressions. Unweighted regressions are in Appendix I section II.1 (Tables A.1-A4), while the same regressions with Conley HAC standard-errors are in section II.2 (Tables A.7-A9). Conclusions remain unchanged irrespective of the specification.

Table 3 reports estimates based on equation (3). The first column shows coefficients obtained with ward and week fixed effects only. Precipitation is found to have a positive and statistically significant effect on cholera. The point estimate suggests that a 10 mm increase in weekly accumulated rainfall causes a 2% increase in recorded cholera cases in a given ward. Including municipality linear trends does not affect the results much (column 2). Municipality-week fixed effects are controlled for instead in column 3. While the impact of precipitation remains statistically significant at the 1% level, its magnitude increases; that is, there are 3.4% additional cholera cases per ward every 10 mm increases in rainfall. Overall, these reduced form estimates consistently show a positive impact of precipitation on cholera incidence.

All subsequent tables are organized in the same fashion, with municipality trends added in column 2 and municipality-week fixed effects added in column 3. To test the sensitivity of our results to measurement error in the recorded rainfall data, we instrument our main precipitation variable with rainfall recorded by the TRMM satellite as explained in section 3.2.2. The potential sources of measurement error in these two datasets are likely to be unrelated, and therefore uncorrelated. Results are displayed in Table 4. The two satellite-based precipitation variables are strongly correlated, and first stage F statistics range between 24 and 37 across specifications (see fourth row). Our two-stages least squares coefficient estimates remain positive but become larger as attenuation bias theory would predict. Including municipality-week fixed effects results in a loss of statistical significance (column 3). The first stage F-statistic is also lower however, inflating standard errors to some degree. On the whole, our findings are supported by the IV results. There is a strong positive relationship between cholera occurrence in a given ward and precipitation. In the interest of proceeding conservatively we continue to stress the OLS results hereafter, but Table 4 suggests that the true impact of precipitation on cholera may be even larger.

Table 5 explores the impact of rainfall using the more flexible quartile specification detailed in equation (4). The second rainfall quartile (light rain or precipitation between 0 to 2.9 mm weekly) is used as omitted category. Notably, the non-parametric relationship between weekly accumulated rainfall and cholera occurrence show particularly large effects at the upper-end of the rainfall distribution. Indeed, the estimated coefficients in the three columns consistently indicate that extreme precipitation has a strong impact on cholera incidence. For instance, a single additional week with recorded rainfall falling in the fourth quartile (>75th percentile, between 26.9 mm and 408.6 mm weekly), relative to a week with light rain, increases the number of new cholera cases by 20.3% (column 2). The first quartile coefficients, measuring loosely speaking the effect of a dry week relative to little rainfall, are positive but not statistically significant. These results are key findings in our paper. Clearly, extreme rainfall has a higher incidence on ward-level cholera occurrence than light rain, suggesting not all ranges of precipitation are necessarily related to cholera occurrence. Moreover, upper-quartile rain has been consistently used in the literature as a proxy for flooding (Chen et al. 2017), and implies water stagnation may be a likely mechanism.

To explore further the role of extreme precipitation, Table 6 puts the emphasis on flooding and attempts to measure its impact in various ways. We begin with assessing whether the impact of rainfall is non-linear and depends on the extent to which a ward is prone to flooding. We use our measure of the share of the ward that is subject to flooding and interact it with weekly accumulated rainfall. Our results presented in panel A show a positive interaction term as theory would predict. The interaction is non-statistically significant however. The coefficient of the uninteracted precipitation measure remains in the same order of magnitude as the coefficients of Table 3. In panel B we measure the effect of the fourth quartile precipitation relative to the rest of the precipitation distribution. Here flooded is a dummy variable for weekly accumulated rainfall falling on the upper-quartile of the overall rainfall distribution. Our estimates are positive, significant at the 1% level, and stable across alternative specifications. In panel C we interact the flooded dummy with our flood-prone area share defined as above. The interaction term is now positive and significant at 5% level, implying that the impact of heavy rainfall is much higher in wards at greater risk of flooding all else equal.

Overall, the results of this section support the theoretical mechanisms described in section 2 and the channels put forward in the public health literature. There are various reasons why heavy rainfall and flooding could lead to an increase in cholera, as mentioned earlier. Not only the bacterium survives longer in wet humid surfaces, but the risk of increased contamination is higher. The inundation of drains, water systems, and pit latrines, greatly enhances the probability of exposure to contaminated water and food. Further, behavioural changes during periods of weather shocks may also increase the probability of contagion (WHO). Finally, indirect mechanisms through income-shocks due to the inundation of job locations or inaccessibility to the work-place may further contribute to the adoption of risky behaviour.

5.2 Non-Linear Effects: A Story of Infrastructure Quality?

We now explore further the relationship between rainfall and cholera incidence and assess potential non-linearities related to ward-level characteristics. As explained above, the size of the weather shock in a given ward is likely to depend on the quality of the infrastructure such as the availability of well-functioning drains, paved roads and improved sanitation. This section concentrates on the correlation that rainfall and several indicators of ward infrastructure and 'neighbourhood quality' has with cholera incidence. Our choice of ward characteristics is in part dictated by data availability. We focus on population density, road density, as well as the density of drains and footways, and the number of water wells. We also include the percentage of the ward's area that hosts informal and formal housing.

As mentioned earlier, we are constrained when it comes to measuring the quality of infrastructure and focus on quantity when no distinction is possible. Because of this, while we expect higher population density to increase the measured effect of rainfall on cholera through a heightened risk of contagion, we are agnostic with respect to the influence of road and water infrastructure measures. On the one hand, greater physical supply of water wells and drains could be negatively associated with cholera by efficiently evacuating used-water and rain. On the other hand, it could magnify the impact of heavy precipitation on cholera when the quality is low, for instance if because of unregulated dumping, drains and evacuation canals are clogged in times of heavy rain.

Table 7 reports baseline estimates of population-weighted regressions. Estimates of unweighted regressions are in Table A5, and results with Conley HAC standard-errors are in Tables A10, both in Appendix I. The size of the coefficients is stable across our different specifications, although the sizes are generally larger when accounting for the spatial-autocorrelation of errors. Further, contemporaneous precipitation remains consistently positive and statistically significant at between 1 to 5% levels.

The first seven columns of the tables separately estimate each interaction term, while in the last three columns we estimate all interactions jointly. Since we lose a large number of observations when we include certain interactions, we report results using three alternative samples. Overall, almost all characteristics considered individually are positive and significant at various levels of significance. Yet, only population and footway densities, as well as housing informality are consistently so across the different specifications. The mechanisms here are intuitive. For instance, in highly populated wards, close proximity between individuals can increase the probability of disease transmission. Similarly, footways are unpaved roads. Contaminated water might stagnate easier in muddy surfaces. At the same time, footways could just reflect informality. Indeed, informality displays the larger size, with on average weekly accumulated rainfall increasing cholera incidence by 1.7% to 4.5% more in wards with higher shares of informal housing.

Once we introduce all interaction terms together in columns (8) to (10) significance and signs are considerably changed suggesting individual interactions may be suffering from omitted variable bias. Nonetheless, some important patterns remain. First, the only consistently positive and statistically significant coefficient (at 5% level) across the various specifications is the non-linear informal housing correlation. While, the sample size is much smaller, results suggest living on informal housing *increases* cholera incidence due to weekly accumulated rainfall by between 2.3% and 5.7%. This finding is far from surprising. Informal settlements are usually located in flood-prone areas. They suffer from poor quality infrastructure and deprived water and sanitation conditions. Penrose et al. (2010) find similar patterns when investigating a previous cholera episode in Dar es Salaam. Two other infrastructure interactions display stable sizes and signs. The non-linear water wells correlation is negative but almost never statistically significant. Most relevant, wards with a higher density of footways have again a higher likelihood of accumulated rainfall affecting weekly cholera incidence. The size is small (between 0.2% to 2%), but it is always statistically significant when accounting for the spatial autocorrelation of errors.

These results support the theoretical mechanisms outlined earlier. First, informal housing and unpaved roads increase the effect of the weather shock. They are thus likely to affect individuals directly and indirectly through health and productivity shocks. Our results so far, despite being imperfectly measured, suggest the quality of infrastructure is highly correlated with the detrimental effect accumulated rainfall and flooding have on cholera prevalence.

5.3 Spatial Spillovers: Neighbours Contagion.

Next, we focus on spatial precipitation spillovers as in equation (5). Precipitation recorded in adjacent wards might exacerbate pressure on water infrastructures. They might also contaminate common water sources, particularly if wards are at different levels of elevation. In this case, uphill rainfall may also wash down contaminated waste or soil material, harming wards below. Table 8 contains our baseline results, while Tables A6 and A13 in Appendix I contain the usual alternative specifications.

We first estimate the average effect of weekly accumulated rain in neighbouring wards on the cholera incidence of a given ward. We do not find evidence of an effect here. The estimated coefficients are almost no different from zero and insignificant across econometric specifications. In Table A11 of Appendix I we further look at the effect of one and two-weeks lags of neighbouring rainfall. None of these spatial lagged variables seem to matter.

We then differentiate between rainfall accumulation in uphill and downhill neighbouring wards. Results here are more nuanced. We register a small but significant effect from accumulated weekly precipitation in adjacent downhill wards. That is, a 10 mm increase in weekly accumulated rainfall in downhill neighbouring wards increases cholera cases in the ward by 0.01% to 0.03%. This finding is consistent with water source contamination from relatively lower wards. The size is negligible suggesting almost no spatial spillovers from rainfall in contiguous areas. Further, precipitation in one's own ward is almost always significant at 1% level, retaining the size of baseline estimates. Again allowing for time lags yields no significant effect (Table A12).

Failing to detect any spatial spillovers of precipitation in contiguous wards is unexpected. It suggests only local contamination prevails. This is consistent with findings in Ambrus et al. (2015) on the 1854 cholera epidemic in London's Soho neighbourhood. Their identification strategy and results suggest cholera is contained within a very specific area. Implications concerning channels of transmission are many but go beyond the scope of this paper.

5.4 Time Dynamic Effects.

So far, we have not taken into account the possibility of a dynamic relationship between rainfall and cholera incidence. If cholera responds to precipitation with a delay, that is, if precipitation in previous weeks or days also impacts cholera in the current week, the estimates of Table 3 could underestimate the true effect. While cholera symptoms can manifest 12 hours after an individual being in contact with the bacteria, they can also take up to 5 days. These might not coincide with our weekly definition. Further, rainfall is easily stored and stagnation from previous weeks may contribute heavily to contagion.

To test for dynamic effects we estimate distributed lag models (equation 6) and allow rainfall to affect health up to five weeks later. The sixth lag of rain (not shown) is not statistically significant. We also report the contemporaneous coefficient and the sum of the six week period. Table 9 displays our point estimates. We gradually introduce additional rainfall lags in our model, which includes municipality-week fixed effects.

First, this exercise allows us to confirm that including time-lags does not change our conclusions. The contemporaneous rainfall effect on cholera remains in the same order of magnitude as in Table 3, close to 3% and statistically significant at 1% level. Second, the results in the table clearly show that past rainfall up to five preceding weeks impact cholera incidence in the current week. All lags decrease in size the farther in time, suggesting the contemporaneous effect matters most. The total effect of precipitation is obtained by summing the coefficients on the contemporaneous and lagged precipitation variables. The total cumulated impact amounts to 0.12 points (last row). That is, six-week cumulated rainfall increases current cholera incidence in a ward by up to 12%. The key message of this table is that weekly cumulated rainfall promotes cholera occurrence immediately and with a lag of up to 5 weeks.

We repeat the exercise with our more flexible non-parametric specification in Table 10 (quartiles). The table is as Table 9 except that we only include lags up to two weeks later. The third lags (not shown) are not statistically significant. Again, the predominant effect is that of extreme rainfall or flooding captured in the upper-quartile, and up to two prior weeks. As before, all lags decrease in size the farther in time, suggesting the contemporaneous effect matters most. Further, we confirm that the inclusion of the lags does not affect our conclusions as the size and statistical significance are unchanged for contemporaneous coefficients. The total cumulative effect of each quartile is also computed in the last row. The total cumulated impact amounts to 0.37 points for the upper-quartile. It is interesting to highlight that the size of the two-week lag of the first quartile (no rain) remains positive but increases in size. It is even statistically significant at 10% level for one specification. This supports theories according to which dryness also matters for cholera incidence by increasing the risk of drinking unsafe water. The time lag is consistent with this type of behavioural changes.

6 Conclusion

Rapid urbanization in developing countries has often led to unplanned cities, particularly in sub-Saharan Africa, with large shares of the urban population living in informal settlements, with poor transport infrastructure and limited access to water and sanitation. Under these conditions, developing countries' city dwellers have become more vulnerable to disease transmission and epidemics. Global warming is expected to exacerbate these health-related risks. The World Bank (2016) estimates that climate change may push up to 77 million more urban residents into poverty by 2030. As extreme weather events become more frequent, understanding the relationship between disease transmission, infrastructure quality and weather shocks in urban areas is important. We make significant advances on this issue.

The key contribution of this paper has been to show that heavy rainfall has a strong positive effect on weekly cholera incidence within wards. We assemble a panel dataset defined at the ward level containing weekly information on cholera incidence, precipitation, and infrastructure quality from various sources. On average, we find that a 10 mm increase in weekly accumulated precipitation leads to an increase of up to 3.5% of weekly recorded cholera cases. Extreme rainfall has a larger impact: a single additional week of rainfall falling above the 75th percentile of the total rainfall distribution increases the number of effective cholera cases by up to 20.3% relative to a week with very light rain. The impact is even higher in wards at greater risk of flooding.

Results in the paper also emphasize the key role of local infrastructure. We find the effect of weekly rainfall on cholera cases to be consistently higher in wards with larger shares of informal housing and a higher density of footways (i.e. informal roads). These results are in line with the mechanisms outlined. Neighbourhoods with low-quality infrastructure are likely to be more exposed to the cholera bacteria when surfaces are washed and drains are overflown by severe precipitation. Vulnerable populations in these wards are also more likely to suffer from negative income shocks during extreme weather events.

Findings here have important policy implications. Cities in developing countries need to address infrastructure gaps to contain the risk of recurrent epidemic outbreaks in fragile environments. Policies that improve the quality of local infrastructures and housing conditions should mitigate the negative impact of rainfall on health. Given the transmission channels of cholera, the proper servicing of informal areas, including sewerages connections and the pavement of informal roads, as well as the regulation of waste-dumping, may prove to be more beneficial in the long-term than the use of short-term palliative measures during outbreaks. Interventions improving access to drinking water as well as access to sanitation should also greatly reduce cholera risk. In the theoretical framework considered, governments can also reduce the adverse effect of weather on health outcomes by supporting households in periods of health-shocks through subsidized health goods or direct transfers. These policies also need to be taken into account given the large room for increasing social safety nets in urban areas. Evidence on large-scale policy interventions in urban areas are limited and more is needed to understand priority-investments that increase resilience and prevent contagion of treatable diseases in developing cities if these are to become engines of growth (Glaeser 2011).

References

Ambrus, A., Field, E. and Gonzalez, R., 2015. *Loss in the Time of Cholera: Long-run Impact of a Disease Epidemic on the Urban Landscape*. Working Paper.

Banerjee, A., Deaton, A. and Duflo, E., 2004. Health, health care, and economic development: Wealth, health, and health services in rural Rajasthan. *The American Economic Review*, *94*(2), p.326.

Barrios, S., Bertinelli, L. and Strobl, E., 2010. Trends in rainfall and economic growth in Africa: A neglected cause of the African growth tragedy. *The Review of Economics and Statistics*, *92*(2), pp.350-366.

Barrios, T., Diamond, R., Imbens, G.W. and Kolesár, M., 2012. Clustering, spatial correlations, and randomization inference. *Journal of the American Statistical Association*, *107*(498), pp.578-591.

Burgess, R., Deschenes, O., Donaldson, D. and Greenstone, M., 2011. Weather and death in India. *Cambridge, United States: Massachusetts Institute of Technology, Department of Economics. Manuscript*.

Burgess, R., Deschenes, O., Donaldson, D., and Greenstone, M. (2017). *Weather , Climate Change and Death in India*.

Chen, J.J., Mueller, V., Jia, Y. and Tseng, S.K.H., 2017. Validating Migration Responses to Flooding Using Satellite and Vital Registration Data. *American Economic Review*, *107*(5), pp.441-445.

Chen, F. and Xi L., 2016 "Evaluation of IMERG and TRMM 3B43 Monthly Precipitation Products over Mainland China," Remote Sensing.

Clasen, T., L. Haller, D. Walker, J. Bartram, and S. Cairncross. 2007, Cost-effectiveness of water quality interventions for preventing diarrhoeal disease in developing countries. Journal of Water and Health 05.4, 599 – 608

Conley, T.G., 1999. GMM estimation with cross sectional dependence. *Journal of econometrics*, *92*(1), pp.1-45.

Deschenes, O. and Greenstone, M., 2007. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *The American Economic Review*, *97*(1), pp.354-385.

Deschenes, O. and Moretti, E., 2009. Extreme weather events, mortality, and migration. *The Review of Economics and Statistics*, *91*(4), pp.659-681.

Deschenes, O., 2011 "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US," American Economic Journal: Applied Economics,2011, 3 (4), 142–85

Dasgupta,P., 2010 (5). "The place of nature in economic development," Handbook of Development Economics.

Dell, M., Jones, B. F., and Olken, B. A., 2012. Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, *4*(3), 66–95.

Dell, M., Jones, B. F., and Olken, B. A., 2014. What Do We Learn from the Weather ? The New Climate–Economy Literature. *Journal of Economic Literature*, *52*(3), 740–798.

Devoto, F., Duflo, E., Dupas, P., Parienté, W., and Pons, V., 2012. Happiness on tap: Piped water adoption in urban Morocco. *American Economic Journal: Economic Policy*, *4*(4), 68–99.

Dunkle, S.E., Mba-Jonas, A., Loharikar, A., Fouché, B., Peck, M., Ayers, T., Archer, W.R., De Rochars, V.M.B., Bender, T., Moffett, D.B. and Tappero, J.W., 2011. Epidemic cholera in a crowded urban environment, Port-au-Prince, Haiti. *Emerging infectious diseases*, *17*(11), p.2143.

Fetzer, T., 2014. Can Workfare Programs Moderate Violence? Evidence from India, STICERD Working Paper.

Glaeser, E., 2011. *Triumph of the city: How our greatest invention makes us richer, smarter, greener, healthier, and happier*. Penguin.

Glaeser, E., and Sims, H., 2015. *Contagion, crime and congestion: overcoming the downsides of density.* (International Growth Center, Policy Brief).

Glaeser, E. and Henderson, J.V., 2017. Urban economics for the developing World: An introduction. *Journal of Urban Economics*, *98*, pp.1-5.

Goddard Earth Sciences Data and Information Services Center, "TRMM (TMPA) Precipitation L3 1 day 0.25 degree x 0.25 degree V7," Goddard Earth Sciences Data and Information Services Center (GES DISC) 2016. Accessed: May 11 2017.

Harrington, L.J., Frame, D.J., Fischer, E.M., Hawkins, E., Joshi, M. and Jones, C.D., 2016. Poorest countries experience earlier anthropogenic emergence of daily temperature extremes. *Environmental Research Letters*, *11*(5), p.055007.

Henderson, J.V., Storeygard, A. and Deichmann, U., 2017. Has climate change driven urbanization in Africa?. *Journal of Development Economics*, *124*, pp.60-82.

Hsiang, S. M., 2010. Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. PNAS, 107(35).

Huffman, G., 2016"GPM IMERG Late Precipitation L3 1 day 0.1 degree x 0.1 degree V03," Goddard Earth Sciences Data and Information Services Center (GES DISC). Accessed: April 20 2017.

Kebede, A.S., Nicholls, R.J., Hanson, S. and Mokrech, M., 2010. Impacts of climate change and sea-level rise: a preliminary case study of Mombasa, Kenya. *Journal of Coastal Research*, *28*(1A), pp.8-19.

Kesztenbaum, L. and Rosenthal, J.L., 2017. Sewers' diffusion and the decline of mortality: The case of Paris, 1880–1914. *Journal of Urban Economics*, *98*, pp.174-186.

Kocornik-Mina, A., McDermott, Th. K.J., Michaels, G. and Rauch, F., 2015. *Flooded cities*. CEP Discussion Paper, 1398. London School of Economics and Political Science, CEP, London, UK.

Lipp, E.K., Huq, A. and Colwell, R.R., 2002. Effects of global climate on infectious disease: the cholera model. *Clinical microbiology reviews*, *15*(4), pp.757-770.

McNally, A., 2016. "FLDAS Noah Land Surface Model L4 monthly 0.1 x 0.1 degree for Eastern Africa (GDAS and RFE2) V001," Goddard Earth Sciences Data and Information Services Center (GES DISC). Accessed: June 18 2017.

Miguel, E. and Kremer, M., 2004. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, *72*(1), pp.159-217.

Miguel, E., Satyanath, S. and Sergenti, E., 2004. Economic shocks and civil conflict: An instrumental variables approach. *Journal of political Economy*, *112*(4), pp.725-753.

Munshi, K., 2003. Networks in the modern economy: Mexican migrants in the US labor market. *The Quarterly Journal of Economics*, *118*(2), pp.549-599.

NASA Precipitation Measurement Missions, "What is the difference between "Realtime" (RT) and "Production" (Prod) Data?," https://pmm.nasa.gov/data-access 2016. Accessed: April 20 2017.

Natty, M., 2013 "Introduction: Key Urban Characteristics of Dar es Salaam Challenges and Opportunities for Resilient Development in the Times of Climate Change," Technical Report, GIZ and Dar es Salaam City Council 2013.

Osei, F.B. and Duker, A.A., 2008. Spatial dependency of V. cholera prevalence on open space refuse dumps in Kumasi, Ghana: a spatial statistical modelling. *International Journal of Health Geographics*, *7*(1), p.62.

Osei, F.B., Duker, A.A., Augustijn, E.W. and Stein, A., 2010. Spatial dependency of cholera prevalence on potential cholera reservoirs in an urban area, Kumasi, Ghana. *International Journal of Applied Earth Observation and Geoinformation*, *12*(5), pp.331-339.

Pan-African START, 2011 "Urban Poverty & Climate Change in Dar es Salaam, Tanzania: A Case Study," Technical Report, Pan-African START Secretariat & International START Secretariat & Tanzania Meteorological Agency & Ardhi University.

Penrose, K., De Castro, M. C., Werema, J., and Ryan, E. T. 2010. Informal urban settlements and cholera risk in Dar es Salaam, Tanzania. *PLoS Neglected Tropical Diseases*, *4*(3).

Sharifi, E., Reinhold S., and Bahram S., 2016. "Assessment of GPM-IMERG and Other Precipitation Products against Gauge Data under Different Topographic and Climatic Conditions in Iran: Preliminary Results," Remote Sensing.

Sasaki, S., Suzuki, H., Igarashi, K., Tambatamba, B. and Mulenga, P., 2008. Spatial analysis of risk factor of cholera outbreak for 2003–2004 in a peri-urban area of Lusaka, Zambia. *The American Journal of Tropical Medicine and Hygiene*, *79*(3), pp.414-421.

Sur D., Manna B., Deb A.K., Deen J.L., Danovaro-Holliday M.C., et al., 2004. Factors associated with reported diarrhoae episodes and treatment-seeking in an urban slum in Kolkata, India. Journal of Health, Population and Nutrition 22: 130–138.

Taylor, D.L., Kahawita, T.M., Cairncross, S. and Ensink, J.H., 2015. The impact of water, sanitation and hygiene interventions to control cholera: a systematic review. *PLoS one*, *10*(8), p.e0135676.

UN-HABITAT 2010, "Informal Settlements and Finance in Dar es Salaam, Tanzania". The Human Settlements Financing Tools and Best Practices Series.

UNDP 2015, Economic Transformation for Human Development, Human Development Report 2014 Tanzania.

Wang, R., Jianyao C. and Xianwei W., 2017 "Comparison of IMERG Level-3 and TMPA 3B42V7 in

Estimating Typhoon-Related Heavy Rain," Water.
Well, David N., 2007. Accounting for the effect of health on economic growth. *The Quarterly Journal of Economics* 122, no. 3 (2007): 1265-1306.
Wenban-Smith, H., 2014. Population Growth, Internal Migration and Urbanisation in Tanzania, 1967-2012: A Census Based Regional Analysis," Technical Report, International Growth Centre 2014.
WHO, 2008 "Cholera Country Profile: United Republic of Tanzania," WHO Global Task Force on Cholera Control 2008.
World Bank, 2015. Measuring Living Standards within Cities. Households Surveys: Dar es Salaam and Durban. Washington, DC: World Bank
World Bank, 2016. Investing in Urban Resilience. GFDRR, World Bank Group Publications.
World Bank, 2017, Eds. Lall, S.V., Henderson, J.V. and Venables, A.J. Africa's Cities: Opening Doors to the World. World Bank Publications.

Main Tables & Figures

Cholera in times of floods.

July 31st, 2017

1. Figures

Figure 1. Distribution of cholera effective cholera cases (epidemic weeks)

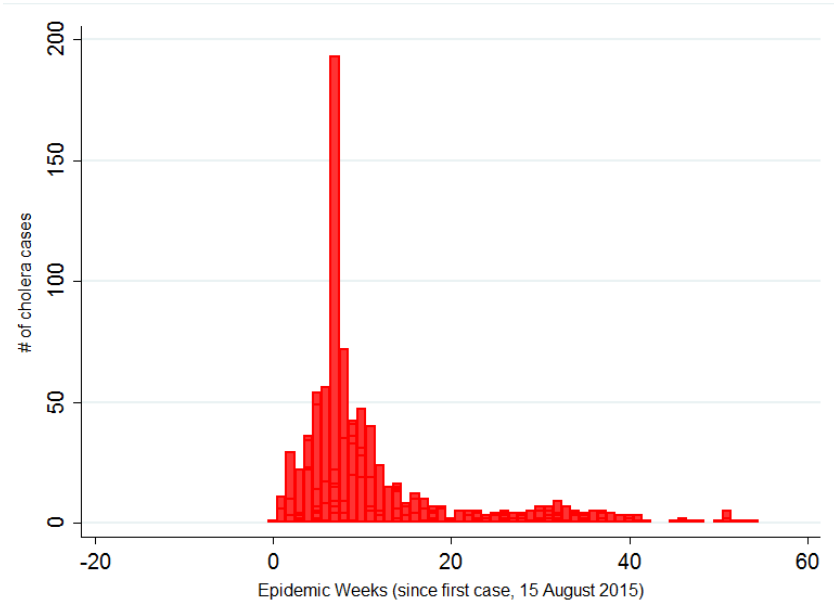


Figure 2. Distribution of effective cholera cases, by age (epidemic weeks)

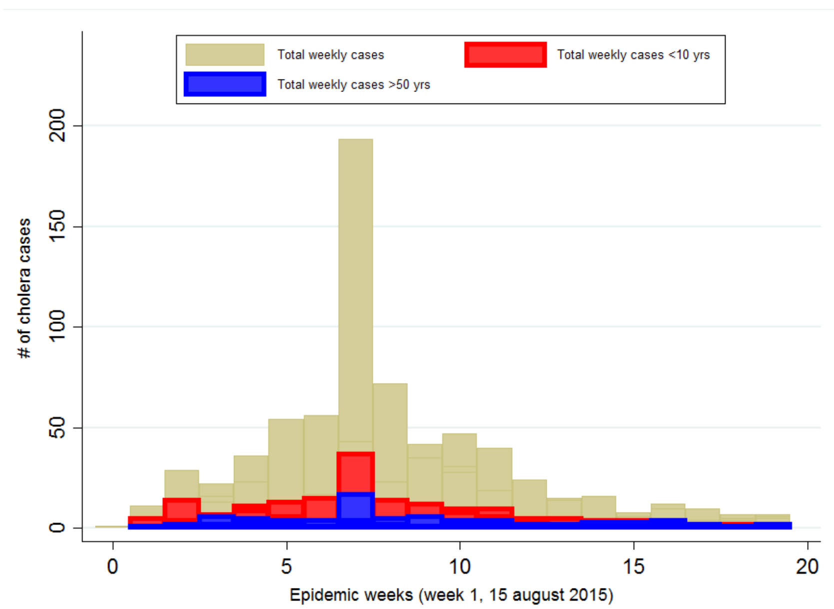
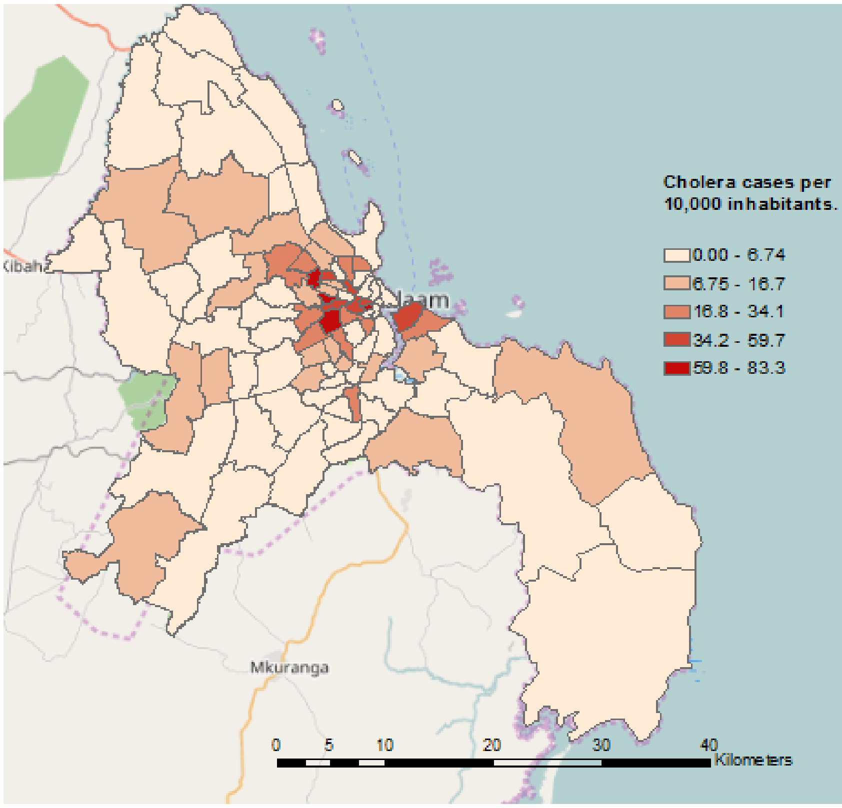


Figure 3. Spatial distribution of cholera indicidence, total outbreak (cases per 10,000 inhabitants)



2. Tables

Table 1: Summary Statistics: Ward characteristics

	Mean	Std. Dev.	Min.	Max.	N
Area (km^2)	18.12	31.15	0.414	209.55	90
Pop (c2012)	48,495	26,064	6,411	106,946	90
HH size (c2012)	4.00	0.21	3.60	4.40	90
Density (km^2)	11.53	11.12	0.05	46.74	90
Improved sanitation	0.08	0.10	0.00	0.48	90
Electricity	0.06	0.08	0.00	0.37	90
Drinking water	0.07	0.08	0.00	0.42	90
Hospital per 10tho.	0.19	0.43	0.00	2.97	90
Density of roads*	4.00	3.84	0.00	18	88
Density of footways*	8.83	15.76	0.01	75.99	63
Density of waterways*	5.00	5.61	0.00	24.81	78
Density of drains*	3.28	2.66	0.02	11.04	45
# water wells	1.25	2.99	0.00	21	76
% area informal	34	28	0.00	88	23

Notes: c2012 referes to data from census 2012, all of the infrastructure density measures (*) are measured in km per square km

Table 2: Summary Statistics: Weather and Cholera

	Mean	Std. Dev.	Min.	Max.	N
Weather:					
% flooded area	10.00	16.00	0.00	73.00	90
Average temperature (C)	26.73	0.37	25.43	27.17	90
Total rainfall (10mm)	162.36	18.15	134.27	216.42	90
Average weekly temperature (C)	26.73	1.57	22.92	29.99	6930
Weekly rainfall accumulation (10mm), GPM	2.11	4.05	0.00	40.86	6930
Weekly rainfall accumulation (10mm), TRMM	2.67	5.53	0.00	44.62	6930
Cholera:					
Total cases 2015-2016	63.32	94.01	0.00	588	90
Total weekly cases per ward (excl. neg)	0.72	3.84	0.00	192	6930
Total weekly cases female	0.36	1.85	0.00	90	6930
Total weekly cases below 5 yrs	0.08	0.49	0.00	14	6930
Total weekly cases tested neg	0.11	0.56	0.00	20	6930
Total effective cases epiweek10	3.16	9.40	0.00	192	890
Total effective cases epiweek20	2.54	7.25	0.00	192	1780
Total effective cases epiweek30	1.77	6.03	0.00	192	2670

Notes: Temperature are degrees celsius; all measures of rainfall are accumulated rainfall (units: 10mm), cholera cases are total numbers. 88 of 90 wards where affected throughout the outbreak.

Table 3: Impact of Weekly Precipitation on Cholera Incidence

	Cholera cases (log)		
	(1)	(2)	(3)
Precipitation	0.0198***	0.0208***	0.0344***
	(0.0072)	(0.0073)	(0.0077)
N	6930	6930	6930
R^2	0.4491	0.4502	0.5254
Ward FE	Yes	Yes	Yes
Week FE	Yes	Yes	
Municipal time trend		Yes	
Municipality × week FE			Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature; they are weighted by the population of the ward (census 2012). The period coverered is from the first week of March 2015 to the first week of September 2016. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table 4: Impact of Weekly Precipitation on Cholera Incidence (Instrumental Variable Estimates)

	Cholera cases (log)		
	(1)	(2)	(3)
Precipitation	0.0689**	0.0779**	0.0451
	(0.0341)	(0.0356)	(0.0476)
N	6930	6930	6930
First Stage F-test	36.812	35.354	24.202
Ward FE	Yes	Yes	Yes
Week FE	Yes	Yes	
Municipal time trend		Yes	
Municipality × week FE			Yes

Notes: Robust standard errors in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016; they are weighted by the population of the ward (census 2012). Nasa GPM v3 precipitation measurement is instrumented with NASA TRMM measurement. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table 5: Impact of Weekly Quartiles of Precipitation on Cholera Incidence

	Cholera cases (log)		
	(1)	(2)	(3)
Quartile 1	0.0010	0.0040	0.0049
	(0.0197)	(0.0193)	(0.0177)
Quartile 3	-0.0360	-0.0286	-0.0536*
	(0.0324)	(0.0329)	(0.0317)
Quartile 4	0.1867***	0.2032***	0.1525**
	(0.0594)	(0.0610)	(0.0611)
N	6930	6930	6930
R^2	0.4510	0.4522	0.5254
Ward FE	Yes	Yes	Yes
Week FE	Yes	Yes	
Municipal time trend		Yes	
Municipality $\times$ week FE			Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature; they are weighted by the population of the ward (census 2012). The period coverered is from the first week of March 2015 to the first week of September 2016. The quartiles of the rainfall distribution are defined as follows: Q1 (0mm), Q2(0-0.29mm), Q3(0.29-2.69mm), Q4(2.69-40.86mm). *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table 6: Impact of Flooding on Cholera Incidence

	Cholera cases (log)		
	(1)	(2)	(3)
Panel A:			
Precipitation	0.0192***	0.0204***	0.0340***
	(0.0071)	(0.0073)	(0.0077)
Precipitation × % Flood-prone area	0.0091*	0.0076	0.0043
	(0.0048)	(0.0046)	(0.0047)
Panel B:			
Flooded (precipitation $\geq$75th p)	0.2189***	0.2280***	0.2029***
	(0.0485)	(0.0487)	(0.0498)
Panel C:			
Flooded (precipitation $\geq$75th p)	0.2007***	0.2103***	0.1970***
	(0.0470)	(0.0472)	(0.0517)
Flooded × % Flood-prone area	0.2262**	0.2176**	0.2076**
	(0.1019)	(0.1007)	(0.1028)
N	6930	6930	6930
Ward FE	Yes	Yes	Yes
Week FE	Yes	Yes	
Municipal time trend		Yes	
Municipality × week FE			Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. All panels are independent regressions. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. Flooded is a dummy variable for weekly precipitation falling above the 75th percentile of the total rainfall distribution. Flood-prone area is the total area of the ward that is prone to flooding. All regressions control for weekly ward air temperature;they are weighted by the population of the ward (census 2012). The period coverered is from the first week of March 2015 to the first week of September 2016. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table 7: Impact of Weekly Precipitation on Cholera Incidence: Infrastructure & Ward Characteristics

	Cholera cases (log)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Precipitation	0.0364***	0.0355***	0.0439***	0.0354***	0.0668**	0.0023	0.0525***	0.0518***	0.0696**	0.0058
	(0.0087)	(0.0080)	(0.0116)	(0.0079)	(0.0296)	(0.0135)	(0.0179)	(0.0126)	(0.0314)	(0.0244)
Precipitation × Pop. density	0.0002**							0.0000	-0.0002	-0.0007
	(0.0001)							(0.0002)	(0.0003)	(0.0004)
Precipitation × Roads density		0.0017*						0.0007	0.0019	-0.0054
		(0.0009)						(0.0016)	(0.0020)	(0.0066)
Precipitation × Footways density			0.0020*					0.0023	0.0033	0.0054
			(0.0011)					(0.0021)	(0.0028)	(0.0054)
Precipitation × # Water wells				0.0009				-0.0002	-0.0000	-0.0014
				(0.0007)				(0.0012)	(0.0013)	(0.0023)
Precipitation × Drains density					0.0018				0.0002	0.0039
					(0.0016)				(0.0021)	(0.0032)
Precipitation × % Informal housing						0.0168*				0.0264**
						(0.0083)				(0.0092)
Precipitation × % Formal housing							0.0032			
							(0.0038)			
N	6930	6776	4851	5852	3465	1771	2618	4004	2926	1463
R^2	0.5229	0.5335	0.5703	0.5370	0.6182	0.6606	0.5383	0.5860	0.6398	0.6768
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE										
Municipality × week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Separate regressions in columns (1-7). Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature; they are weighted by the population of the ward (census 2012). The period coverered is from the first week of March 2015 to the first week of September 2016. Population density is the number of inhabitants per square km (census 2012), Roads density, footway density, and drains density are the meters of roads, footways and drains per km (OSM), % informal and formal houses in the ward are obtained from surveyed plots (not all plots are surveyed). *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table 8: Impact of Neighbours' Weekly Precipitation on Cholera Incidence: Spatial Spillovers

	Cholera cases (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Precipitation	0.0183**	0.0192***	0.0196***	0.0204***	0.0316***	0.0320***
	(0.0071)	(0.0070)	(0.0073)	(0.0072)	(0.0077)	(0.0076)
Neighbours precipitation	0.0004		0.0004		0.0009	
	(0.0007)		(0.0007)		(0.0008)	
Uphill neighbours precipitation		-0.0007		-0.0007		0.0001
		(0.0010)		(0.0009)		(0.0009)
Downhill neighbours precipitation		0.0011		0.0010		0.0013*
		(0.0007)		(0.0007)		(0.0007)
N	6930	6930	6930	6930	6930	6930
R^2	0.4491	0.4494	0.4502	0.4505	0.5255	0.5256
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes		
Municipal time trend			Yes	Yes		
Municipality × week FE					Yes	Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature; they are weighted by the population of the ward (census 2012). The period coverered is from the first week of March 2015 to the first week of September 2016. Neighbours' precipitation measures weekly accumulated rainfall in a neighbouring ward. Uphill and downhill measures are for neighbouring wards at a higher or lower elevation than the given ward. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table 9: Dynamic Effects: Lags of Weekly Precipitation on Cholera Incidence

	Cholera cases (log)				
	(1)	(2)	(3)	(4)	(5)
Precipitation	0.0337***	0.0313***	0.0300***	0.0305***	0.0298**
	(0.0075)	(0.0071)	(0.0073)	(0.0073)	(0.0072)
Precipitation $_{(w-1)}$	0.0249***	0.0243***	0.0214***	0.0200**	0.0205**
	(0.0086)	(0.0085)	(0.0081)	(0.0082)	(0.0083)
Precipitation $_{(w-2)}$		0.0222***	0.0215***	0.0195***	0.0185**
		(0.0071)	(0.0071)	(0.0068)	(0.0070)
Precipitation $_{(w-3)}$			0.0256***	0.0252***	0.0240***
			(0.0077)	(0.0076)	(0.0073)
Precipitation $_{(w-4)}$				0.0182**	0.0180**
				(0.0076)	(0.0075)
Precipitation $_{(w-5)}$					0.0113*
					(0.0061)
Cumulative (6 weeks)					0.1221***
					(0.0288)
N	6840	6750	6660	6570	6480
R^2	0.5263	0.5267	0.5274	0.5275	0.5270
Ward FE	Yes	Yes	Yes	Yes	Yes
Municipality $\times$ week FE	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature; they are weighted by the population of the ward (census 2012). The period coverered is from the first week of March 2015 to the first week of September 2016. Precipitation $_{w-n}$ are the lags of weekly precipitation up to n weeks. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table 10: Dynamic Effects: Lags of Quartiles of Weekly Precipitation on Cholera Incidence

	Cholera cases (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Q1	0.0013	-0.0020	0.0039	0.0005	0.0061	0.0058
	(0.0194)	(0.0190)	(0.0190)	(0.0187)	(0.0175)	(0.0171)
Q3	-0.0335	-0.0372	-0.0262	-0.0301	-0.0519*	-0.0551*
	(0.0309)	(0.0300)	(0.0315)	(0.0307)	(0.0300)	(0.0296)
Q4	0.1825***	0.1753***	0.1982***	0.1908***	0.1504**	0.1442**
	(0.0559)	(0.0538)	(0.0578)	(0.0556)	(0.0574)	(0.0554)
$Q1_{w-1}$	-0.0034	-0.0066	0.0004	-0.0030	-0.0049	-0.0071
	(0.0218)	(0.0222)	(0.0220)	(0.0224)	(0.0198)	(0.0202)
$Q3_{w-1}$	-0.0198	-0.0215	-0.0137	-0.0154	-0.0227	-0.0271
	(0.0352)	(0.0346)	(0.0350)	(0.0344)	(0.0314)	(0.0302)
$Q4_{w-1}$	0.1212**	0.1139**	0.1355**	0.1277**	0.1368**	0.1256**
	(0.0590)	(0.0565)	(0.0592)	(0.0569)	(0.0605)	(0.0566)
$Q1_{w-2}$		0.0312		0.0347*		0.0158
		(0.0192)		(0.0192)		(0.0206)
$Q3_{w-2}$		0.0320		0.0372		0.0495
		(0.0356)		(0.0357)		(0.0352)
$Q4_{w-2}$		0.0903*		0.1029*		0.1052*
		(0.0543)		(0.0549)		(0.0568)
Q1 (Cumulative 3 weeks)						0.0145
						(0.350)
Q3 (Cumulative 3 weeks)						-0.0327
						(-0.490)
Q4 (Cumulative 3 weeks)						0.3750***
						(0.1348)
N	6930	6930	6930	6930	6930	6930
R^2	0.4522	0.4526	0.4535	0.4540	0.5267	0.5271
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes		
Municipal time trend			Yes	Yes		
Municipality $\times$ week FE					Yes	Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature; they are weighted by the population of the ward (census 2012). The period coverered is from the first week of March 2015 to the first week of September 2016. The quartiles are defined as in Table 4. Quartile_{w-n} are the lags of the quartiles of weekly precipitation up to n weeks. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

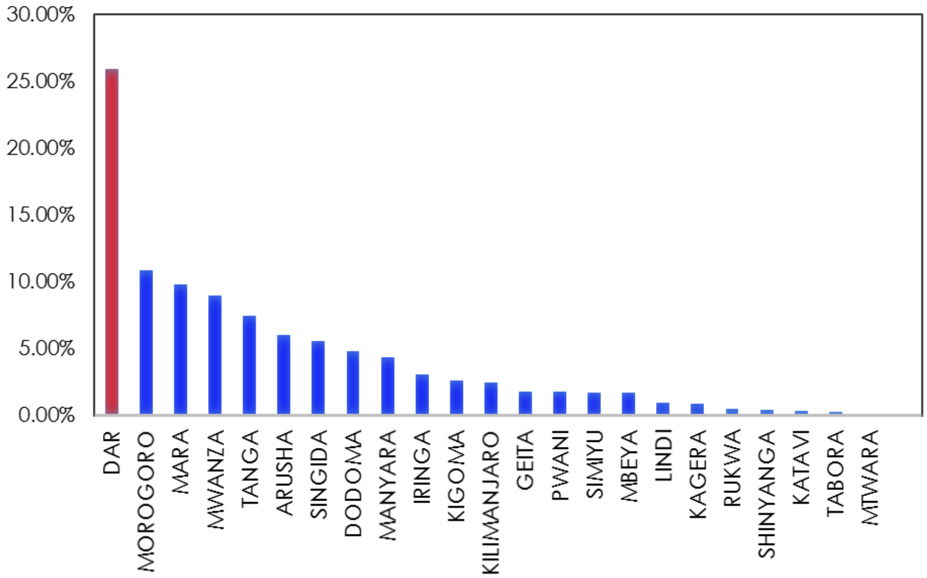
Appendix I

Cholera in times of floods.

July 31st, 2017

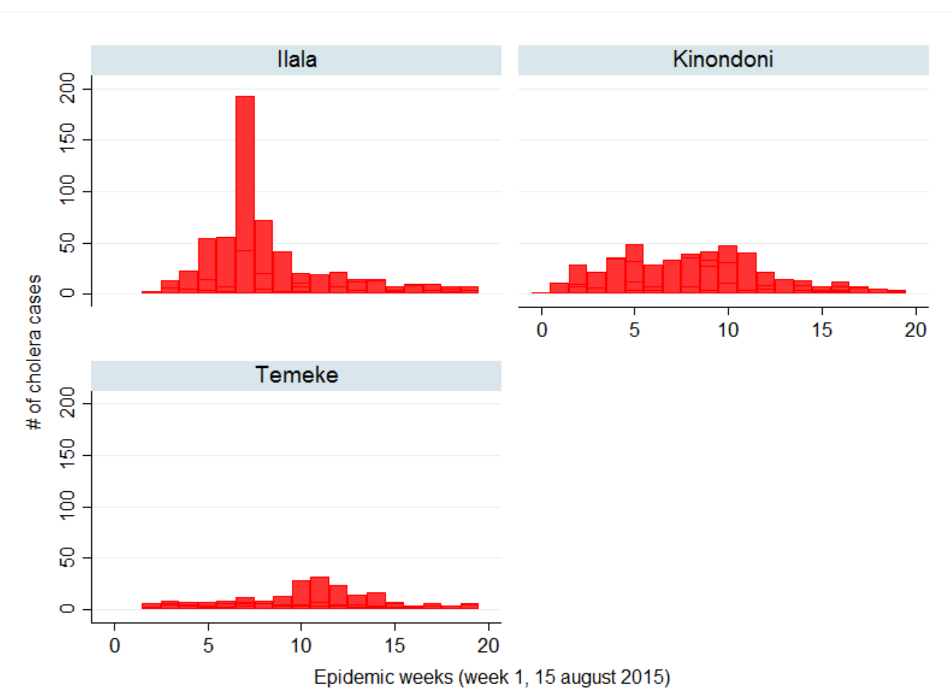
I. Appendix Figures

Figure A1. Cholera cases during 2015-2016 outbreak, by region



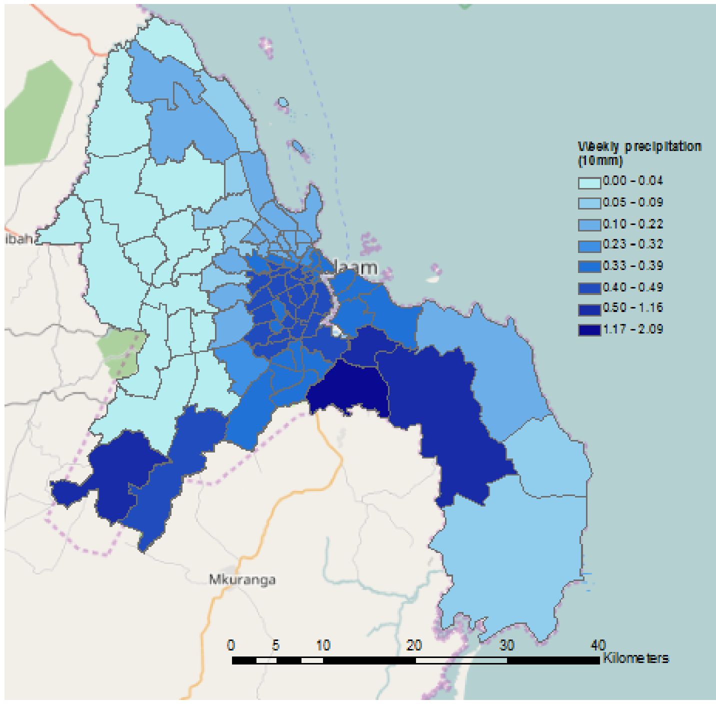
Notes: Data obtained from the Red Cross. Total cases (vs. effective in analysis) up until April 2016.

Figure A2. Distribution of effective cholera cases (epidemic week), by district municipality



Notes: Therea are currently 5 municipal districts in Dar es Salaam. Here we use the three that existed when the cholera outbreak started and at the levels at which the data was collected.

Figure A3. Ward-level weekly rainfall accumulation (area-weighted)



II. Appendix Tables

II.1 Unweighed Main Regressions

Table A1: Impact of Weekly Precipitation on Cholera Incidence (unweighted)

	Cholera cases (log)		
	(1)	(2)	(3)
Precipitation	0.0157**	0.0165***	0.0231***
	(0.0060)	(0.0061)	(0.0063)
N	6930	6930	6930
R^2	0.4049	0.4057	0.4638
Ward FE	Yes	Yes	Yes
Week FE	Yes	Yes	
Municipal time trend		Yes	
Municipality × week FE			Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A2: Impact of Weekly Precipitation on Cholera Incidence (IV Estimates, unweighted)

	Cholera cases (log)		
	(1)	(2)	(3)
Precipitation	0.0431**	0.0477**	0.0394
	(0.0188)	(0.0194)	(0.0266)
N	6930	6930	6930
First Stage F-test	40.822	40.21	27.985
Ward FE	Yes	Yes	Yes
Week FE	Yes	Yes	
Municipal time trend		Yes	
Municipality × week FE			Yes

Notes: Robust standard errors in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016. Nasa GPM v3 precipitation measurement is instrumented with NASA TRMM measurement. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A3: Impact of Weekly Quartiles of Precipitation on Cholera Incidence (unweighted)

	Cholera cases (log)		
	(1)	(2)	(3)
Q1	-0.0089	-0.0058	-0.0030
	(0.0160)	(0.0157)	(0.0155)
Q3	-0.0301	-0.0250	-0.0359
	(0.0264)	(0.0265)	(0.0269)
Q4	0.1592***	0.1706***	0.1521***
	(0.0521)	(0.0528)	(0.0565)
N	6930	6930	6930
R^2	0.4065	0.4074	0.4645
Ward FE	Yes	Yes	Yes
Week FE	Yes	Yes	
Municipal time trend		Yes	
Municipality × week FE			Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016. The quartiles of the rainfall distribution are defined as follows: Q1 (0mm), Q2(0-0.29mm), Q3(0.29-2.69mm), Q4(2.69-40.86mm)*p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A4: Impact of Flooding on Cholera Incidence (unweighted)

	Cholera cases (log)		
	(1)	(2)	(3)
Panel A:			
Precipitation	0.0154**	0.0162***	0.0227***
	(0.0060)	(0.0061)	(0.0062)
Precipitation × % Flood-prone area	0.0079	0.0066	0.0059
	(0.0051)	(0.0050)	(0.0052)
Panel B:			
Flooded (precipitation ≥75th p)	0.1865***	0.1930***	0.1856***
	(0.0424)	(0.0425)	(0.0456)
Panel C:			
Flooded (precipitation ≥75th p)	0.1688***	0.1757***	0.1658***
	(0.0420)	(0.0421)	(0.0451)
Flooded × %Flood-prone area	0.2181***	0.2113***	0.2136**
	(0.0795)	(0.0786)	(0.0879)
N	6930	6930	6930
Ward FE	Yes	Yes	Yes
Week FE	Yes	Yes	
Municipal time trend		Yes	
Municipality × week FE			Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. All panels are independent regressions. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. Flooded is a dummy variable for weekly precipitation falling above the 75th percentile of the total rainfall distribution. Flood-prone area is the total area of the ward that is prone to flooding. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016.*p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A5: Impact of Weekly Precipitation on Cholera Incidence: Infrastructure & Ward Characteristics (unweighted)

	Cholera cases (log)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Precipitation	0.0238***	0.0243***	0.0246**	0.0221***	0.0245	-0.0008	0.0509***	0.0242*	0.0258	0.0063
	(0.0069)	(0.0067)	(0.0121)	(0.0064)	(0.0225)	(0.0128)	(0.0173)	(0.0138)	(0.0253)	(0.0221)
Precipitation × Pop. density	0.0002***							0.0001	0.0000	-0.0008
	(0.0001)							(0.0002)	(0.0003)	(0.0005)
Precipitation × Roads density		0.0014						0.0007	0.0024	-0.0055
		(0.0009)						(0.0016)	(0.0020)	(0.0064)
Precipitation × Footways density			0.0027**					0.0023	0.0024	0.0063
			(0.0012)					(0.0019)	(0.0023)	(0.0055)
Precipitation × # Water wells				0.0013*				-0.0004	-0.0001	-0.0015
				(0.0007)				(0.0012)	(0.0013)	(0.0023)
Precipitation × Drains density					0.0029*				0.0006	0.0046
					(0.0015)				(0.0019)	(0.0038)
Precipitation × % Informal housing						0.0178**				0.0233**
						(0.0072)				(0.0083)
Precipitation × % Formal housing							0.0023			
							(0.0038)			
N	6930	6776	4851	5852	3465	1771	2618	4004	2926	1463
R^2	0.4633	0.4728	0.5160	0.4668	0.5534	0.6435	0.5398	0.5303	0.5858	0.6693
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE										
Municipality × week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Separate regressions in columns (1-7). Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016. Population density is the number of inhabitants per square km (census 2012), Roads density, footway density, and drains density are the meters of roads, footways and drains per km (OSM), % informal and formal houses in the ware are obtained from surveyed plots (not all plots are surveyed).*p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A6: Impact of Neighbours' Weekly Precipitation on Cholera Incidence: Spillovers (unweighted)

	Cholera cases (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Precipitation	0.0158**	0.0162***	0.0166***	0.0170***	0.0226***	0.0227***
	(0.0061)	(0.0061)	(0.0062)	(0.0062)	(0.0063)	(0.0062)
Neighbours' precipitation	-0.0000		-0.0001		0.0002	
	(0.0007)		(0.0007)		(0.0008)	
Uphill neighbours' precipitation		-0.0011		-0.0011		-0.0007
		(0.0008)		(0.0008)		(0.0009)
Downhill neighbours' precipitation		0.0008		0.0007		0.0008
		(0.0007)		(0.0007)		(0.0007)
N	6930	6930	6930	6930	6930	6930
R^2	0.4049	0.4052	0.4057	0.4061	0.4638	0.4640
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes		
Municipal time trend			Yes	Yes		
Municipality × week FE					Yes	Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016. Neighbours' precipitation measures weekly accumulated rainfall in a neighbouring ward. Uphill and downhill measures are for neighbouring wards at a higher or lower elevation than the given ward. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

II.2 Main Regressions, Spatial Auto-correlation of Standard Errors (Conley HAC SE)

Table A7: Impact of Weekly Precipitation on Cholera Incidence (HAC SE)

	Cholera cases (log)	
	(1)	(2)
Precipitation	0.0157***	0.0271***
	(0.0051)	(0.0098)
N	6930	6930
R^2	0.0015	0.0088
Ward FE	Yes	Yes
Week FE	Yes	Yes
Municipality × week FE		Yes

Notes: Conley HAC standard errors in parenthesis (Conley 1999, 2008). Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period covered is from the first week of March 2015 to the first week of September 2016. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A8: Impact of Weekly Quartiles of Precipitation on Cholera Incidence (HAC SE)

	Cholera cases (log)	
	(1)	(2)
Q1	-0.0089	0.0254
	(0.0327)	(0.0276)
Q3	-0.0301	0.0013
	(0.0211)	(0.0194)
Q4	0.1592***	0.2165***
	(0.0508)	(0.0662)
N	6930	6930
R^2	0.0043	0.0098
Ward FE	Yes	Yes
Week FE	Yes	Yes
Municipality × week FE		Yes

Notes: Conley HAC standard errors in parenthesis (Conley 1999, 2008). Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016. The quartiles of the rainfall distribution are defined as follows: Q1 (0mm), Q2(0-0.29mm), Q3(0.29-2.69mm), Q4(2.69-40.86mm). *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A9: Impact of Flooding on Cholera Incidence (HAC SE)

	Cholera cases (log)	
	(1)	(2)
Panel A:		
Precipitation	0.0154***	0.0225***
	(0.0051)	(0.0073)
Precipitation × % Flood-prone area	0.0079	0.0486**
	(0.0223)	(0.0219)
Panel B:		
Flooded (precipitation ≥75th p)	0.1865***	0.2128***
	(0.0455)	(0.0665)
Panel C:		
Flooded (precipitation ≥75th p)	0.1688***	0.1511***
	(0.0430)	(0.0585)
Flooded × % Flood-prone area	0.2181	0.5879***
	(0.2266)	(0.1720)
N	6930	6930
Ward FE	Yes	Yes
Week FE	Yes	Yes
Municipality × week FE		Yes

Notes: Conley HAC standard errors in parenthesis (Conley 1999, 2008). All panels are independent regressions. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. Flooded is a dummy variable for weekly precipitation falling above the 75th percentile of the total rainfall distribution. Flood-prone area is the total area of the ward that is prone to flooding. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016. *p ≤ 0.10 ** p≤0.05 *** p≤0.01

Table A10: Impact of Weekly Precipitation on Cholera Incidence: Infrastructure & Ward Characteristics (HAC SE)

	Cholera cases (log)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Precipitation	0.0221***	0.0319***	0.0297***	0.0256**	0.0371	0.0120	0.0504***	0.0281**	0.0256	0.0272
	(0.0069)	(0.0087)	(0.0089)	(0.0104)	(0.0242)	(0.0200)	(0.0156)	(0.0128)	(0.0306)	(0.0297)
Precipitation × Pop. density	0.0010**							0.0005**	0.0000	-0.0023*
	(0.0004)							(0.0002)	(0.0001)	(0.0014)
Precipitation × Roads density		0.0063***						0.0004	0.0050***	-0.0132
		(0.0020)						(0.0000)	(0.0015)	(0.0091)
Precipitation × Footways density			0.0125*					0.0106*	0.0084**	0.0213*
			(0.0064)					(0.0057)	(0.0040)	(0.0123)
Precipitation × # Water wells				0.0073*				-0.0023	-0.0007	-0.0050*
				(0.0040)				(0.0023)	(0.0023)	(0.0028)
Precipitation × drains density					0.0159**				0.0098**	0.0190
					(0.0078)				(0.0048)	(0.0116)
Precipitation × % Informal housing						0.0455*				0.0571**
						(0.0246)				(0.0267)
Precipitation × % Formal housing							0.0182**			
							(0.0075)			
N	6930	6776	4851	5852	3465	1771	2618	4004	2926	1463
R^2	0.0226	0.0122	0.0197	0.0161	0.0293	0.0172	0.0424	0.0295	0.0371	0.0442
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week FE										
Municipality × week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes:Conley HAC standard errors in parenthesis (Conley 1999, 2008). Separate regressions in columns (1-7). Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016. Roads density, footway density, and drains density are the meters of roads, footways and drains per km (OSM), % informal and formal houses in the ward are obtained from surveyed plots (not all plots are surveyed). *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

II.3 Extensions

Table A11: Impact of Neighbours' Weekly Lagged Precipitation on Cholera Incidence (1)

	Cholera cases (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Precipitation	0.0285**	0.0288**	0.0292**	0.0293**	0.0346***	0.0335***
	(0.0115)	(0.0113)	(0.0117)	(0.0115)	(0.0111)	(0.0110)
Neighbours precipitation	-0.0046	-0.0049	-0.0044	-0.0045	-0.0007	0.0008
	(0.0036)	(0.0047)	(0.0036)	(0.0047)	(0.0034)	(0.0047)
Neighbours precipitation $_{w-1}$	0.0005	0.0007	0.0005	0.0005	0.0002	-0.0007
	(0.0003)	(0.0013)	(0.0003)	(0.0013)	(0.0003)	(0.0013)
Neighbours precipitation $_{w-2}$		-0.0001		-0.0001		0.0007
		(0.0011)		(0.0011)		(0.0011)
N	6929	6924	6929	6924	6929	6924
R^2	0.4493	0.4493	0.4504	0.4504	0.5255	0.5256
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes		
Municipal time trend			Yes	Yes		
Municipality $\times$ week FE					Yes	Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature; they are weighted by the population of the ward (census 2012). The period coverered is from the first week of March 2015 to the first week of September 2016. Neighbours' precipitation measures weekly accumulated rainfall in a neighbouring ward. Neighbour's Precipitation$_{w-n}$ are the lags of weekly accumulated rainfall in neighbouring wards up to n weeks. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A12: Impact of Neighbours' Weekly Lagged Precipitation by Elevation on Cholera Incidence (2)

	Cholera cases (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Precipitation	0.0218**	0.0214**	0.0233**	0.0230**	0.0355***	0.0354***
	(0.0097)	(0.0105)	(0.0100)	(0.0108)	(0.0098)	(0.0105)
Uphill neighbours precipitation	-0.0042	-0.0044	-0.0047	-0.0047	-0.0052	-0.0049
	(0.0051)	(0.0048)	(0.0052)	(0.0049)	(0.0051)	(0.0050)
Downhill neighbours precipitation	-0.0032	-0.0033	-0.0036	-0.0037	-0.0047	-0.0045
	(0.0052)	(0.0051)	(0.0053)	(0.0051)	(0.0051)	(0.0051)
Uphill N's precipitation $_{w-1}$	0.0037	0.0042	0.0040	0.0044	0.0055	0.0046
	(0.0055)	(0.0069)	(0.0055)	(0.0069)	(0.0053)	(0.0070)
Downhill N's precipitation$_{w-1}$	0.0043	0.0062	0.0047	0.0063	0.0062	0.0070
	(0.0054)	(0.0073)	(0.0054)	(0.0072)	(0.0052)	(0.0073)
Uphill N's precipitation$_{w-2}$		-0.0004		-0.0002		0.0007
		(0.0042)		(0.0043)		(0.0042)
Downhill N's precipitation$_{w-2}$		-0.0018		-0.0016		-0.0012
		(0.0043)		(0.0043)		(0.0043)
N	6929	6924	6929	6924	6929	6924
R^2	0.4494	0.4495	0.4505	0.4506	0.5257	0.5258
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes		
Municipal time trend			Yes	Yes		
Municipality × week FE					Yes	Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature; they are weighted by the population of the ward (census 2012). The period coverered is from the first week of March 2015 to the first week of September 2016. Neighbours' precipitation measures weekly accumulated rainfall in a neighbouring ward. Uphill and downhill measures are for neighbouring wards at a higher or lower elevation than the given ward. N's Precipitation Uphill/Downhill $_{w-n}$ are the lags of weekly accumulated rainfall in neighbouring wards up to n weeks according to their elevation with respect to the given ward. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A13: Impact of Neighbours' Weekly Precipitation on Cholera Incidence: Spillovers (HAC SE) (1)

	Cholera cases (log)			
	(1)	(2)	(3)	(4)
Precipitation	0.0158***	0.0162***	0.0196**	0.0196**
	(0.0047)	(0.0047)	(0.0084)	(0.0085)
Neighbours' precipitation	-0.0000		0.0028***	
	(0.0007)		(0.0010)	
Uphill neighbours' precipitation		-0.0011		0.0017**
		(0.0009)		(0.0009)
Downhill neighbours' precipitation		0.0008		0.0035***
		(0.0009)		(0.0012)
N	6930	6930	6930	6930
R^2	0.0015	0.0021	0.0108	0.0114
Ward FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Municipality × week FE			Yes	Yes

Notes:Conley HAC standard errors in parenthesis (Conley 1999, 2008). Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016. Neighbours' precipitation measures weekly accumulated rainfall in a neighbouring ward. Uphill and downhill measures are for neighbouring wards at a higher or lower elevation than the given ward. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A14: Impact of Neighbours' Weekly Lagged Precipitation on Cholera Incidence: Spillovers (HAC SE) (2)

	Cholera cases (log)			
	(1)	(2)	(3)	(4)
Precipitation	0.0189***	0.0187***	0.0247***	0.0251***
	(0.0051)	(0.0060)	(0.0087)	(0.0091)
Uphill neighbours precipitation	-0.0043	-0.0044	-0.0053	-0.0047
	(0.0037)	(0.0042)	(0.0042)	(0.0063)
Downhill neighbours precipitation	-0.0032	-0.0033	-0.0046	-0.0040
	(0.0032)	(0.0051)	(0.0056)	(0.0078)
Uphill N's precipitation $_{w-1}$	0.0033	0.0033	0.0072*	0.0050
	(0.0036)	(0.0108)	(0.0041)	(0.0134)
Downhill N's precipitation$_{w-1}$	0.0041	0.0052	0.0082	0.0072
	(0.0031)	(0.0117)	(0.0056)	(0.0144)
Uphill N's precipitation $_{w-2}$		0.0000		0.0016
		(0.0079)		(0.0079)
Downhill N's precipitation$_{w-2}$		-0.0010		0.0004
		(0.0076)		(0.0074)
N	6929	6924	6929	6924
R^2	0.0022	0.0023	0.0119	0.0119
Ward FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Municipality × week FE			Yes	Yes

Notes:Conley HAC standard errors in parenthesis (Conley 1999, 2008). Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature. The period coverered is from the first week of March 2015 to the first week of September 2016. Neighbours' precipitation measures weekly accumulated rainfall in a neighbouring ward. Uphill and downhill measures are for neighbouring wards at a higher or lower elevation than the given ward. N's Precipitation Uphill/Downhill $_{w-n}$ are the lags of weekly accumulated rainfall in neighbouring wards up to n weeks according to their elevation with respect to the given ward. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Table A15: Lagged Dependent Variable Model

	Cholera cases (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Precipitation	0.0199***	0.0155***	0.0119**	0.0111**	0.0103**	0.0103**
	(0.0054)	(0.0051)	(0.0051)	(0.0051)	(0.0051)	(0.0051)
Cholera cases $_{w-1}$	0.6111***	0.4430***	0.4037***	0.3943***	0.3913***	0.3914***
	(0.0457)	(0.0336)	(0.0348)	(0.0360)	(0.0368)	(0.0371)
Cholera cases $_{w-2}$		0.2747***	0.2112***	0.1975***	0.1922***	0.1923***
		(0.0230)	(0.0285)	(0.0268)	(0.0281)	(0.0283)
Cholera cases $_{w-3}$			0.1435***	0.1172***	0.1085***	0.1087***
			(0.0206)	(0.0245)	(0.0255)	(0.0257)
Cholera cases $_{w-4}$				0.0652***	0.0476**	0.0479**
				(0.0208)	(0.0203)	(0.0197)
Cholera cases $_{w-5}$					0.0447**	0.0455**
					(0.0214)	(0.0217)
Cholera cases $_{w-6}$						-0.0019
						(0.0185)
N	6930	6930	6930	6930	6930	6930
R^2	0.6837	0.7076	0.7136	0.7148	0.7153	0.7153
Ward FE						
Municipality × week FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors clustered at the ward level in parenthesis. Precipitation is measured as the weekly accumulated rainfall in a given ward (10mm units), cholera cases are the log of effective (tested positive) weekly cholera cases in a given ward. All regressions control for weekly ward air temperature; they are weighted by the population of the ward (census 2012). The period coverered is from the first week of March 2015 to the first week of September 2016. Cholera cases$_{w-n}$ are lagged effective cholera cases up week n. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

Appendix II

Cholera in times of floods.

July 25th, 2017

Tanzania 2015-16 DHS Data Analysis

This section presents the results of the analysis of the 2015-16 Tanzania Demographic and Health Survey and Malaria Indicator Survey (DHS). Its objective is to shed light on the relationship, if any, between income, wealth, and incidence of diarrhoeal diseases in Dar-es-Salaam. The 2015-16 DHS collected reliable information on several demographic and health indicators, including infant and child mortality, nutritional status of mothers and children, and childhood immunizations and diseases. It was implemented by several government agencies with financial support from various bilateral and multilateral donors. It is representative at the national, urban and rural area levels.

We restrict the sample to all the households living in the Dar-es-Salaam region. The DHS only asked children under the age of five questions related to diarrhoea. This leaves us with a sample of 367 children living in 97 distinct households. The DHS calculates for each household a wealth index using a battery of socio-economic variables. The construction of this index goes as follows. Respondent households are given scores based on the number and kinds of consumer goods they own, such as a television or a fridge. Housing characteristics, such as access to drinking water, toilet facilities, and flooring materials are also taken into account. These scores are derived using principal component analysis. Each household is assigned a household wealth score index. National wealth quintiles are then compiled by dividing the index distribution into five equal categories, containing 20% of the population each. The primary objective of this section is to examine the relationship between this wealth index and the incidence of diarrhoea among children aged less than five years.

We conduct a simple regression analysis and estimate the linear probability model in Equation (1) below with least squares:

$$D_{ih} = \beta_0 + \beta_1.W_{ih} + \beta_2.X_i + \beta_3.Z_h + \epsilon_{ih} \tag{1}$$

where D_{ih} indicates whether child i in household h has had diarrhoea in the last two weeks. W_{ih} is the household wealth index either measured as five quantile dummies or the continuous index value. X_i is a vector of individual covariates and includes age and a female gender dummy. Mother educational attainment is controlled for in vector Z_h. ϵ_{ih} is the error and β_k are the parameters to be estimated. Standard errors are clustered at the household level.

Table AII.1 presents summary statistics of the variables included in the regression analysis. 16.3% of the children in the sample report a diarrhoea episode in the two weeks prior to interview. The average child is 1.8 years old. 48% of the sample is comprised of young girls. Regression estimates of Equation 1 are shown in Table AII.2. The continuous wealth index score is introduced in column 1. The wealth index quintile dummies are included in the second column. Overall the table presents very weak evidence in favour of a wealth bias regarding diarrheal risk. The continuous wealth index score has a negative but insignificant association with the probability of a child getting diarrhoea. The point estimates of the second column indicate that only children of the third quintiles are more likely to get sick than the poorest children. All other quintile coefficients are insignificant at conventional levels of significance.

Table AII.1: Descriptive Characteristics DHS Analysis

	Mean	**Std. Dev.**	**Min.**	**Max.**	**N**
Diarrhoea in last two weeks	367	0.163	0.37	0	1
Urban wealth index (/1000) - continuous score	367	56.985	66.163	-211.621	186.827
Urban wealth index Q1	367	0.014	0.116	0	1
Urban wealth index Q2	367	0.153	0.36	0	1
Urban wealth index Q3	367	0.259	0.439	0	1
Urban wealth index Q4	367	0.297	0.458	0	1
Urban wealth index Q5	367	0.278	0.449	0	1
Age	367	1.845	1.357	0	4
Female	367	0.48	0.5	0	1
Mother education: no education	367	0.065	0.248	0	1
Mother education: incomplete primary	367	0.054	0.227	0	1
Mother education: complete primary	367	0.534	0.5	0	1
Mother education: incomplete secondary	367	0.057	0.233	0	1
Mother education: complete secondary	367	0.243	0.429	0	1
Mother education: higher education	367	0.046	0.21	0	1

Notes:DHS 2015-16 data for Dar es Salaam region. Under five years old children in the sample.

Table AII.2: DHS Regression Analysis

	Had Diarrhea =1	
Dependent variable	(1)	(2)
Urban wealth index - continuous score	-0.000295	
	-0.000328	
Urban wealth index Q2		0.0657
		-0.0546
Urban wealth index Q3		0.134**
		-0.0647
Urban wealth index Q4		0.0694
		-0.0656
Urban wealth index Q5		0.0363
		-0.0694
Age	-0.0501***	-0.0500***
	-0.0122	-0.0124
Female	0.026	0.0259
	-0.0383	-0.0378
Mother education: incomplete primary	0.138	0.0957
	-0.085	-0.091
Mother education: complete primary	0.155***	0.112***
	-0.0415	-0.0391
Mother education: incomplete secondary	0.618***	0.565***
	-0.105	-0.105
Mother education: complete secondary	0.100*	0.0677
	-0.0596	-0.0579
Mother education: higher education	0.13	0.0959
	-0.0949	-0.0879
R-squared	0.142	0.15
Observations	367	367

Notes: Robust standard errors clustered at the household level in parenthesis. Linear probability model regressions. DHS 2015-16 data for Dar es Salaam region. Under five years old children in the sample. *p $\leq$ 0.10 ** p$\leq$0.05 *** p$\leq$0.01

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about
our work on our website
www.theigc.org

For media or communications
enquiries, please contact
mail@theigc.org

Subscribe to our newsletter
and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic
and Political Science,
Houghton Street,
London WC2A 2AE

International
Growth Centre

DIRECTED BY

FUNDED BY

Designed by soapbox.co.uk