The Role of Water Availability in Mitigating Heat-Related Mortality: Empirical Evidence from South Africa

Kelly Hyde, University of Pittsburgh[1]

Abstract

I find that higher water availability reduces the slope of the temperature-mortality relationship above the excess heat threshold in South Africa, with heterogeneous effects based on existing water infrastructure, the prevalence of waterborne illness, and race. As rising global surface temperatures threaten to reduce precipitation and evaporate surface freshwater in areas already experiencing water stress, this finding suggests decreasing precipitation will amplify the direct effect of climate change on mortality. I estimate of the economic cost of this interaction as a lower bound of optimal investment in water infrastructure and supply technologies to preempt it.

[1]4518 Wesley W. Posvar Hall, 230 S Bouquet St, Pittsburgh, PA 15260; email kdh50@pitt.edu.

1 Introduction

> *"Cape Town running out of water is like San Diego going dry. Which, if you factor in the looming threat of climate change, may not be that far off... ...we, like many other cities around the globe, are facing a drier future with increasingly unpredictable rains. What is happening to us in Cape Town might not be an outlier. It could happen to you too."*
>
> —Aryn Baker, for TIME Magazine[2]

Perhaps the most striking contrast between developed and developing economies is the accessibility and quality of water. For all but the poorest individuals in OECD countries, potable water is a given, with 99% using at least basic drinking water services and 91% using a "safely managed" one as of 2015. In sub-Saharan Africa, these percentages are 58% and 24%, respectively.[3] Lacking the water to survive is a figment of a post-apocalyptic imagination for the average household in the United States, but an everyday possibility for more than 40% of the global population.[4]

As the world works toward the goal of universal access to safe drinking water articulated in WHO and UNICEF's 2030 Agenda for Sustainable Development, rising global surface temperatures threaten to reduce precipitation and evaporate surface freshwater in areas already experiencing water stress. In South Africa, climate assessment models predict a decrease in average precipitation of up to 10% by 2050 as a consequence of climate change (Nkhonjera et al. 2017). The water supply of South Africa is predominantly groundwater; thus reductions in precipitation, which are already erratic in South Africa's semi-arid climate, and the resulting decrease in groundwater recharge could lead to catastrophic levels of water stress. The opening quote of this paper shows these effects have already been felt in Cape Town, which in 2018 narrowly averted "Day Zero," the day in which the city's reservoirs dried up completely.

The independent effects of water stress and excess heat on mortality, respectively, are well understood. The literature has established a U-shaped relationship between ambient temperature and the mortality rate, with extreme temperatures at both ends of a regional climate's distribution increasing the likelihood of various causes of mortality (Curriero et al. 2002, Hajat and Kosatsky 2010, Basagaña et al. 2011). In this paper, I provide empirical evidence of a significant *interaction* between water availability and the heat-mortality relationship. I find that higher water availability, as measured by volumetric flow rates of key rivers in South Africa, reduces the slope of the temperature-mortality relationship above the excess heat threshold. This suggests that the decrease in precipitation predicted in South Africa (Nkhonjera et al. 2017) and more broadly (Intergovernmental Panel on Climate Change 2014) as a consequence of climate change will amplify the direct effect on mortality of rising temperatures. I provide estimates of the economic cost of this interaction as a lower bound of the optimal amount to invest in water infrastructure and

[2] http://time.com/cape-town-south-africa-water-crisis/

[3] Statistics from WHO/UNICEF Joint Monitoring Programme (JMP) for Water Supply, Sanitation and Hygiene, accessed via World Bank Data Portal.

[4] Millenium Development Goals Report, 2015.

technologies to increase the supply of potable water (e.g. desalination, rainwater harvesting) to preempt it.

To identify the causal effect of water availability on heat-related mortality, I construct a panel comprising daily climate data from the National Oceanic and Atmospheric Administration (NOAA), river flow measures from the Hydrological Services department of South Africa's Department of Water and Sanitation, administrative cause of death records, and General Household Survey data from Statistics South Africa. I construct a weighted-average water availability index based on within-province variation in two distinct flow measures: the water flowing from the river into pipes to reservoirs and sanitation facilities for distribution, hereafter referred to as *infrastructural* availability, and the water flowing downstream past that pipe, hereafter *natural* availability. The index measures variation at the intensive margin of household water supply.

Using unit fixed-effects panel regression, I estimate the effect of the interaction between the water availability index and ambient temperatures above an excess heat threshold on the number of deaths, finding a highly statistically significant negative coefficient. To estimate the effect of variation at the extensive margin, I add an "on/off" indicator of infrastructural availability, finding an independently significant effect for populations with sufficiently high levels of water infrastructure to benefit from it. Finally, I include robustness checks demonstrating the stability of the estimated coefficient at alternative excess heat thresholds. Overall, I find that increased water availability mitigates heat-related mortality, reducing the slope of the heat-mortality relationship and in some cases making the relationship statistically indistinguishable from zero. This strongly suggests that investments to increase the availability of potable water should be a central focus of policies to reduce the adverse effects of climate change.

The paper proceeds as follows. Section 2 reviews the existing literature. Section 3 describes the empirical strategy. Section 4 describes the data used and provides summary statistics. Section 5 presents results and evidence of heterogeneous treatment effects based on existing water infrastructure, water quality, and race. Section 6 measures the economic cost of the projected mortality rates. Section 7 concludes.

2 Literature Review

Global surface temperature is expected to increase more than 1.5° C by the end of the century, increasing the duration, intensity, and frequency of exposure to high temperatures around the world (Allen et al. 2014). Excess heat has been shown to damage welfare via several channels, including decreased cognitive performance (Zivin et al. 2018), increased cardiovascular and respiratory mortality risk (Curriero et al. 2002, Basagaña et al. 2011), increased incidence and severity of injury during physical exertion (Nelson et al. 2011), increased incidence of low birth weight (Deschênes et al. 2009) and infant mortality (Banerjee and Bhomwick 2016), and ultimately, increased overall mortality (Hajat and Kosatsky 2010). Heat increases can also damage welfare indirectly, through increases in hazardous atmospheric pollution (Tressol et al. 2008), exacerbations of humidity-related injuries (Barreca et al. 2015, Wehner et al. 2016), and decreases in precipitation, resulting in drought (Intergovernmental Panel on Climate Change 2014).

Adaptations—i.e., actions or behavior modifications intended to mitigate the aforementioned negative effects of heat—can reduce or prevent the aforementioned impacts. At the household level, potential adaptations include residential air conditioning, spending more time indoors, or moving to a more amenable climate; at the community level, institutions can invest in early warning systems for anticipated heat waves, build local cooling centers, or increase access to quality water (Deschênes et al. 2009). Existing literature has mainly explored the effects of household-level adaptations, with residential air conditioning dominating in efficacy (Barreca et al. 2015). In developed economies such as the United States, where equipment is readily available and most individuals have access to the electricity required to use it, promoting adoption of residential air conditioning is among the best approaches to mitigate the effects of excess heat.

Indeed, Barreca et al. (2016) find that 95 percent of the decline in the temperature-mortality relationship in the United States over the 20$^{\text{th}}$ century is attributable to mass adoption of residential air conditioning. However, in developing economies such as South Africa, there are several barriers to this approach. Air conditioning equipment may be prohibitively costly or simply unavailable, especially in rural areas. Even if equipment is available and attainable, the requisite household electricity may not be, or may be low-quality, hazardous, and subject to frequent service interruptions. The especially vulnerable—children, the elderly, the poor, agricultural workers, and those with preexisting respiratory or cardiovascular conditions, for example—are also unlikely to be able to move to a different climate or simply spend more time indoors. For these individuals, a community-level intervention is necessary. In this paper, I provide evidence for the efficacy of one community-level intervention proposed by Deschênes et al. (2009): investments to increase access to quality water.

This paper also contributes to the literature on water resource economics, which has grown alongside and complementarily to the literature on climate change. Beyond its necessity for human life and strong association with health outcomes, economic significance of water availability and quality has been demonstrated in property rights institutions (Kremer et al. 2011), comparative advantage (Debaere 2014), human capital accumulation (Beach et al 2016), schooling (Ao 2016), cognitive ability (Troesken et al. 2011), and mental health (Devoto et al. 2010). In this paper, I show that mitigation of heat-related mortality is yet another potential benefit of investments in water infrastructure. While the volume of water flowing into rivers is largely determined by natural, exogenous factors, investment in technologies that increase the efficiency of water distribution systems (e.g. improved sanitation facilities, improved piping) or pull in water from other sustainable sources (e.g. rainwater harvesting, desalination) can both broaden access to the benefits of quality water and protect households from the negative shocks to natural availability threatened by climate change.

3 Empirical Strategy

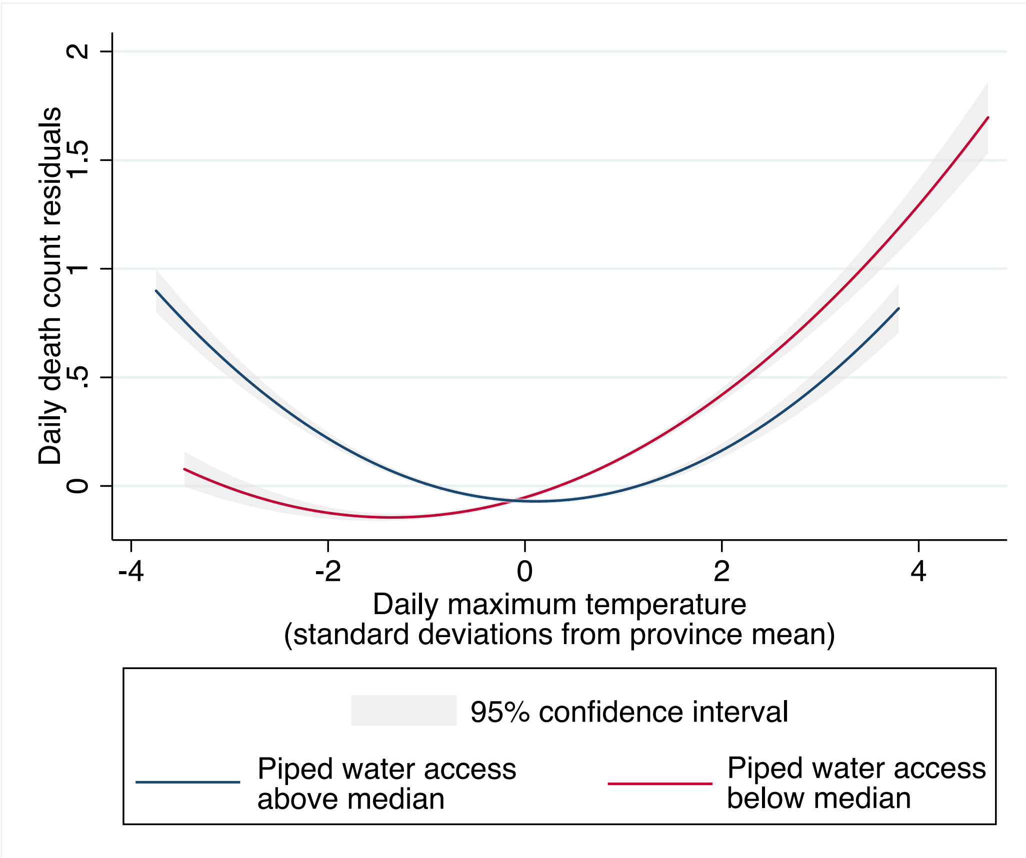


Figure 1: Heat-Mortality Relationship for Selected Provinces Based on Water Access, 2009-2015

Figure 1 presents the temperature-mortality curves for two halves of the population based on likelihood of access to a piped water source. To remove confounding sources of variation in mortality, I regress the monthly mortality rate on a battery of controls[5] and use the residuals. Also, to compare across provinces with different average temperatures, I use standard deviations from the province mean. Using deviations in temperature instead of the level is supported by the heat epidemiology literature, which has established that the definition of "excess heat" is dependent on regional climate because humans acclimatize to the usual conditions of their surroundings.[6]

The curves presented in Figure 1 have two main implications: that temperature affects the mortality rate in South Africa with the expected U-shaped relationship, and that individuals less likely than the median to have access to a piped water source are (ceteris paribus) more likely to die following exposure to heat. Figure 2 makes the same conclusions more stark by comparing the

[5] Province, month, and year fixed effects; household expenditure; electricity access; ownership of a cell phone; educational attainment; presence of a flushing toilet in household; urban-rural indicator; head of household disability; gender, age, and marital status of head of household.

[6] Direct evidence of this in my data is presented in Appendix 8.1.

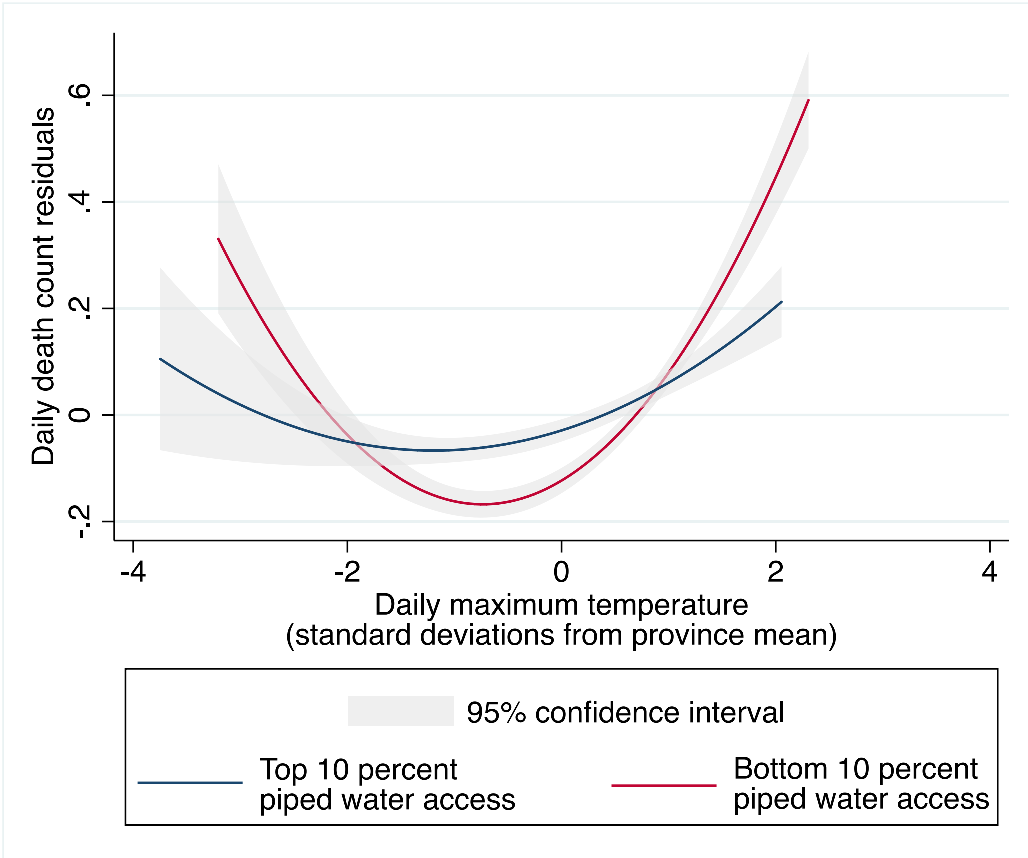


Figure 2: Heat-Mortality Relationship by Water Access Percentile, 2009-2015

temperature-mortality curve for individuals in the top decile of piped water access with the curve for the bottom decile. The curve for the bottom decile is significantly more convex than for the top decile on both sides of the origin, suggesting a greater degree of sensitivity to both unusually hot and unusually cold temperatures; in this paper, I focus only on the former.

While Figures 1 and 2 strongly suggest the proposed mitigating effect of water availability on heat-related mortality, they are not sufficient to demonstrate a causal relationship. The possession of a home with piped water is clearly not exogenously imposed, and is likely to be strongly associated with other factors that may impact heat-related mortality, such as wealth. The "ideal experiment" to identify a causal relationship would randomly assign access to piped water to households and test whether the treatment of piped water reduced household members' likelihood of death during periods of excess heat. However, this experiment would be incredibly difficult, costly, and ethically challenging to implement. As a feasible alternative, I exploit variation in two volumetric flow (m^3/s) measures of water availability. To determine *infrastructural* water availability, I use measurements of the volume of water flowing through pipes from key rivers in each province to sanitation facilities; these pipes feed water into reservoirs from which households' piped water is distributed. To determine *natural* water availability, I use measurements from the downstream component at each of these pipes—i.e., the volume of water flowing past the pipe through the river. This provides a measure of the amount of water available to an individual retrieving their water directly from the river or another natural source of water. I assume that households strictly prefer piped water to natural sources of water, so the only individuals relying on these natural sources are those who do not have access to potable piped water (either they do not have pipes at all, or the water coming through the pipes is unusable due to contamination or pipe failure).

From these two measures of water availability, I construct a weighted average:

$$WaterAvail_{it} = InfraSource_{ijt} \times InfraAvail_{it} + (1 - InfraSource_{ijt}) \times NaturalAvail_{it} \quad (1)$$

In equation (1), $InfraSource_{ijt}$ represents the proportion of population subgroup j in province i whose primary water source is infrastructural (i.e., indoor plumbing, a private source of piped water on their property, or a public tap maintained by the municipality). $InfraAvail_{it}$ represents infrastructural water availability and $NaturalAvail_{it}$ represents natural water availability in province i at time t, respectively. Thus $WaterAvail_{it}$ places greater weight on infrastructural availability in areas where people have access to it, and lower weight where they do not. I use this as an index of overall availability of potable water.

To estimate the effect of water availability on the slope of the heat-mortality curve, I regress the monthly number of deaths in each province on the index of water availability, a dummy for temperatures in excess of a hazardous threshold, and the interaction between water availability and excess heat. This regression model has the following form:

$$\begin{aligned} Deaths_{ijt} = {} & \beta_1 TempThreshold_{it} + \beta_2 WaterAvail_{it} + \gamma_1 \left(TempThreshold_{it} \times WaterAvail_{it} \right) \\ & + \beta_3 InfraOn_{it} + \gamma_2 \left(TempThreshold_{it} \times InfraOn_{it} \right) + \rho_i + \phi_j + \delta_t + \epsilon_{ijt} \end{aligned} \quad (2)$$

where $TempThreshold_{it}$ is a dummy variable which takes a value of 1 when the maximum temperature in geographic region i at time t exceeds a selected threshold (e.g. 100° F), $WaterAvail_{it}$ is the index of water availability in province i at time t, $InfraOn_{it}$ is a dummy variable that equals 1 when infrastructural availability is significantly greater than zero[7], ρ_i represents province fixed effects, ϕ_j represents population subgroup[8] controls, and δ_t represents month-year fixed effects. The coefficients of interest to this paper are γ_1, which captures the mortality effect of a change in water availability at the intensive margin conditional on the maximum temperature exceeding the selected threshold, and γ_2, which captures the effect of a change at the extensive margin of infrastructural availability. I estimate this model using unit fixed-effects panel regression.

For γ_1 and γ_2 to be interpreted as causal, it is necessary to assume that both the maximum temperature and the water availability experienced by population subgroup j in region i at time t are exogenous. For the reasons previously discussed in this section, using the proportion of

[7]at or above the 25th percentile within-province

[8]grouped by age group and gender within each province

each subgroup with access to piped water in the household does not satisfy this assumption. Additionally, to isolate the impact of water availability, it is necessary to find a measure of water availability which is independent of the instantaneous maximum temperature. This eliminates precipitation, since rainfall lowers the temperature when it occurs, so any measured effect of precipitation on the heat-related mortality rate may be explained by precipitation simply reducing the amount of excess heat instead of increasing water availability.

The volumetric flow data I use satisfy both of these assumptions. Both measures are largely determined by precipitation, runoff, and other climatic factors that vary exogenously with weather conditions. Using weather conditions as a source of exogenous variation in heat and humidity has been well-established in the literature (Deschênes and Greenstone 2007, Barreca et al. 2012, Barreca et al. 2016) as a method of forecasting the long-run impacts of climate change. While these measures may be partially endogenous at the province- or country-level, since the amount of water flowing through a river may be influenced by previous water consumption and the amount of water flowing through pipes to sanitation facilities is a function of where the pipes were constructed, they are exogenous at the household-level, since a single household cannot significantly influence either of these measures of water availability. Additionally, since the model is estimated with fixed-effects panel regression, the significant variation is within-province; households cannot select into an area with higher-than-usual water availability without perfect foresight of future weather variation.

4 Data

To estimate equation (1), I collate contemporary data from South Africa on daily climate measures, volumetric flow measures from rivers and sanitation pipelines, household survey data, and administrative cause of death records. I describe each of these components in turn below.

1. **Daily climate measures**: Daily South African province-level data on maximum temperature, average temperature, and precipitation was obtained from the National Oceanic and Atmospheric Administration's National Centers for Environmental Information (NCEI) for all years between 1997 and 2015.

2. **Volumetric flow measures from rivers and sanitation pipelines**: Daily South African province-level data on the average flow rate (measured in m^3/s) was obtained from the South Africa Department of Water and Sanitation's Hydrological Services resource for all years between 1997 and 2015. One reservoir station was selected per province. At each reservoir, two measures of the flow rate were obtained: the flow of the river component at the reservoir (i.e., the river the reservoir is drawn from at the point of extraction), and the flow of a pipeline from the reservoir station to a sanitation facility. These two measures were selected to obtain information on both natural water resources (the river) and infrastructural

water resources (the flow into a sanitation facility), which will be shown to have independent relevance in Section 4.

3. **Household survey data**: Annual data on household characteristics was obtained from Statistics South Africa's General Household Survey for 2002, 2003, and all years between 2005 and 2015. At the time of writing, survey data for 2004 and years prior to 2002 was not available. For the main results reported in this paper, I use data from 2002 for the missing years. All results are replicated without these missing years in Appendix 8.5. A full record of the variables comprising the household survey data is provided in the footnote of Table 1.

4. **Administrative cause of death records**: Death records were obtained from Statistics South Africa for all years between 1997 and 2015.[9] Data from 2009 to 2015 includes the day on which each death occurred; data prior to 2009 only includes the month. Each death record includes the gender, marital status, province of death, age, and causes of death of the decedent. A full record of the variables comprising the cause of death data is provided in the footnote of Table 1.

Table 1 presents summary statistics of the data described above. The average monthly maximum temperature (i.e., the highest daily maximum temperature achieved within a particular month) across all provinces is 86.8° F, with the hottest province, Mpumalanga, having an average monthly maximum temperature of 98.3° F. Across all provinces, the 95th percentile of daily maximum temperature is approximately 91° F, and the 95th percentile of monthly maximum temperature is approximately 100° F; for the remainder of this paper, these temperature levels will be used (for daily and monthly data, respectively) as the "threshold" above which the effect of interest is observed. Figure 3 shows the quadratic relationship between daily (monthly) maximum temperatures and daily (monthly) number of deaths. The selected thresholds are near the inflection point of the associated graph, respectively; i.e., increases beyond the threshold temperature are associated with increases in the mortality rate. This is in line with the heat epidemiology literature, which uses this inflection point as the threshold that defines "heat" for a particular geographic area (Hajat and Kosatsky 2010, Kovats et al. 2004, Curriero et al. 2002).

Consistent with the observation of rising global temperatures over recent decades, Figure 4 shows an increasing time trend for the average maximum temperature, amounting to an increase of approximately 1.2° F from 1997 to 2015. Figure 5 demonstrates opposing trends in the two measures of water availability used—while the pipeline flow rate to sanitation facilities has increased, likely reflecting a response to increased water demand due to population growth, the river flow rate has declined, consistent with the prediction that an increase in temperature will cause depletion of the water table. Alongside these trends, Figure 6 shows that from 1997 to 2015, the proportion of the population relying on public municipal taps as their primary water source has increased,

[9] At the time of writing, no data prior to 1997 or after 2015 was available from Statistics South Africa.

Table 1: Summary Statistics by Province

Province	Western Cape	Eastern Cape	Northern Cape	Free State	KwaZulu-Natal	North West	Gauteng	Mpumalanga	Limpopo	Average
Deaths (monthly total)	194 (92)	319 (153)	56 (28)	180 (105)	484 (269)	178 (96)	435 (218)	165 (92)	193 (93)	245 (196)
Pipe flow rate to water sanitation facility	0.02 (0.01)	0.02 (0.01)	0.03 (0.03)	1.41 (0.26)	8.86 (31.39)	0.37 (0.15)	1.85 (15.86)	3.84 (17.74)	0.43 (0.72)	1.88 (13.45)
River component flow rate at distribution pipeline	7.6 (11.1)	174.6 (210.2)	150.4 (190.5)	41 (81.1)	6.4 (7.3)	10.5 (16.5)	44.9 (93.8)	6.1 (15.9)	8.3 (20.3)	50.2 (121.6)
Monthly maximum temperature (° F)	88 (7.5)	80.5 (10.6)	87.9 (9.7)	82.7 (7.8)	88.5 (4.7)	85.4 (7.1)	80.9 (6.5)	98.3 (7.8)	88.6 (6.1)	86.8 (8.5)
% of deaths attributed to gastroenteritis	0.02 (0.03)	0.04 (0.04)	0.04 (0.05)	0.06 (0.04)	0.06 (0.04)	0.05 (0.04)	0.04 (0.03)	0.07 (0.05)	0.08 (0.05)	0.05 (0.05)
% of population with at least high school diploma	0.28 (0.17)	0.15 (0.12)	0.16 (0.13)	0.18 (0.15)	0.19 (0.16)	0.15 (0.14)	0.27 (0.18)	0.16 (0.16)	0.14 (0.14)	0.19 (0.16)
% of population with access to piped water	0.9 (0.02)	0.42 (0.04)	0.79 (0.04)	0.89 (0.03)	0.54 (0.03)	0.62 (0.02)	0.89 (0.02)	0.68 (0.02)	0.43 (0.03)	0.68 (0.19)
% of population with a cellular telephone	0.73 (0.14)	0.69 (0.16)	0.67 (0.15)	0.75 (0.16)	0.74 (0.20)	0.76 (0.15)	0.83 (0.13)	0.82 (0.15)	0.78 (0.16)	0.75 (0.16)

Means unless otherwise specified; standard deviations in parentheses. Household survey data was collated with the following variables: *PipedWater* (% of population with access to piped water), *PipedWaterInHome* (% of pop. with piped water in their home), *PublicTap* (% of pop. relying on public municipal taps for water), *NaturalSourceWater* (% of pop. relying on a natural water resource, such as a river or borehole), *ElecAccess* (% of population with access to mains electricity in residence), household expenditure, race, marital status, gender, *Metro* (% of population living in a metropolitan area), *Disability* (% of population with a self-reported disability), *ChildHunger* (% of households who reported often lacking enough food to feed their children), *NoChildren* (% of households without any children 17 years or younger), *Diploma* (% of pop. with at least a high school diploma), *NoSchooling* (% of pop. with no formal schooling), *FlushToilet* (% of households with a flushable toilet on their property), *Cellular_telephone* (% of pop. with a cell phone), *SafeWater* (% of households self-reporting their primary water source as safe), *Outdoor_indoor_air_pollution* (% of pop. reported experiencing air pollution either in- or outdoors), and *Water_pollution* (% of pop. reported experiencing water pollution). Cause of death data was collated and collapsed at the province-level to the following variables: *Deaths* (number of deaths occurring in a particular month within a particular province and population subgroup), *GastroRate* (% of deaths in a particular month associated with infectious diarrhea or gastroenteritis), and *UnnaturalRate* (% of deaths in a particular month attributed to unnatural causes).

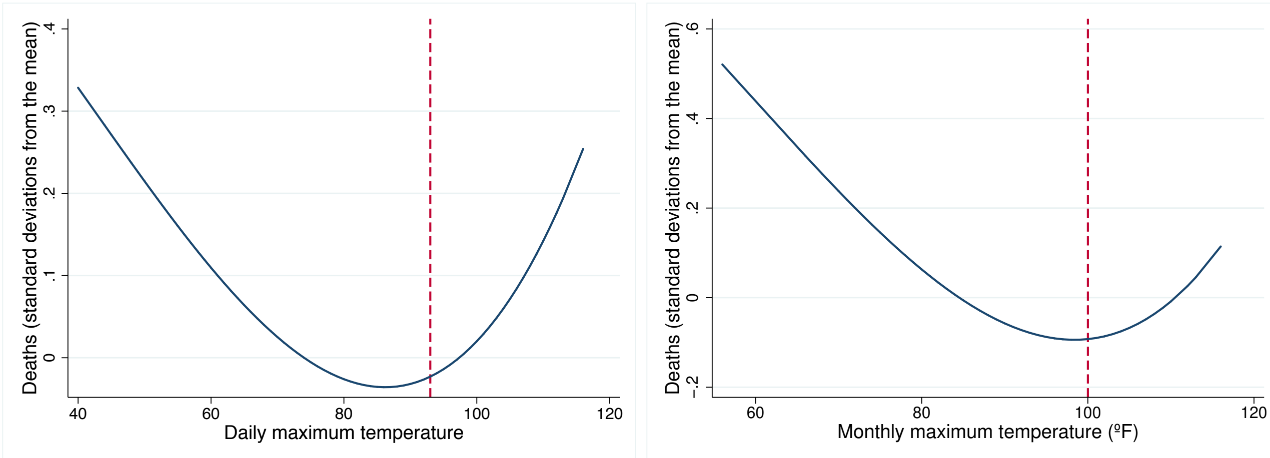


Figure 3: Deaths and maximum temperature (left: daily, 2009-2015; right: monthly, 1997-2015)

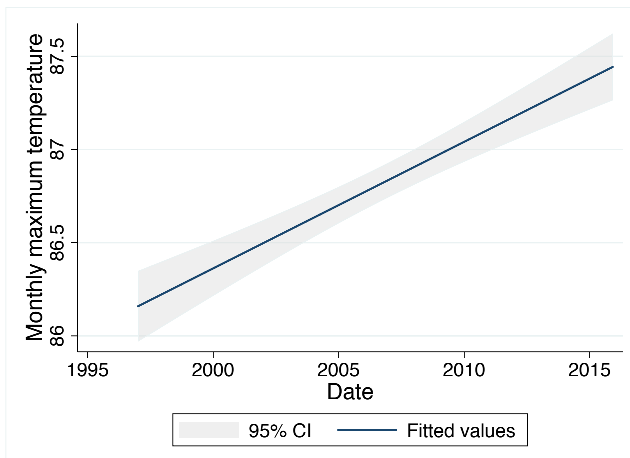


Figure 4: Monthly maximum temperature linear time trend, 1997-2015

while the proportion with piped water (either in their homes or on their property) has decreased. Individuals relying on external sources of water, especially those at a non-negligible distance from their houses, are particularly susceptible to heat-related morbidities such as dehydration because they must go outside once they have depleted their stored water. The trends in Figure 5 have two likely explanations: first, the population is growing faster in areas and among population strata that are less likely to have access to piped water. Second, investment in municipal water resources has improved the reliability of public taps, while household survey data suggests that residential piped water is unreliable with frequent service interruptions; thus households may be voluntarily switching to public taps.

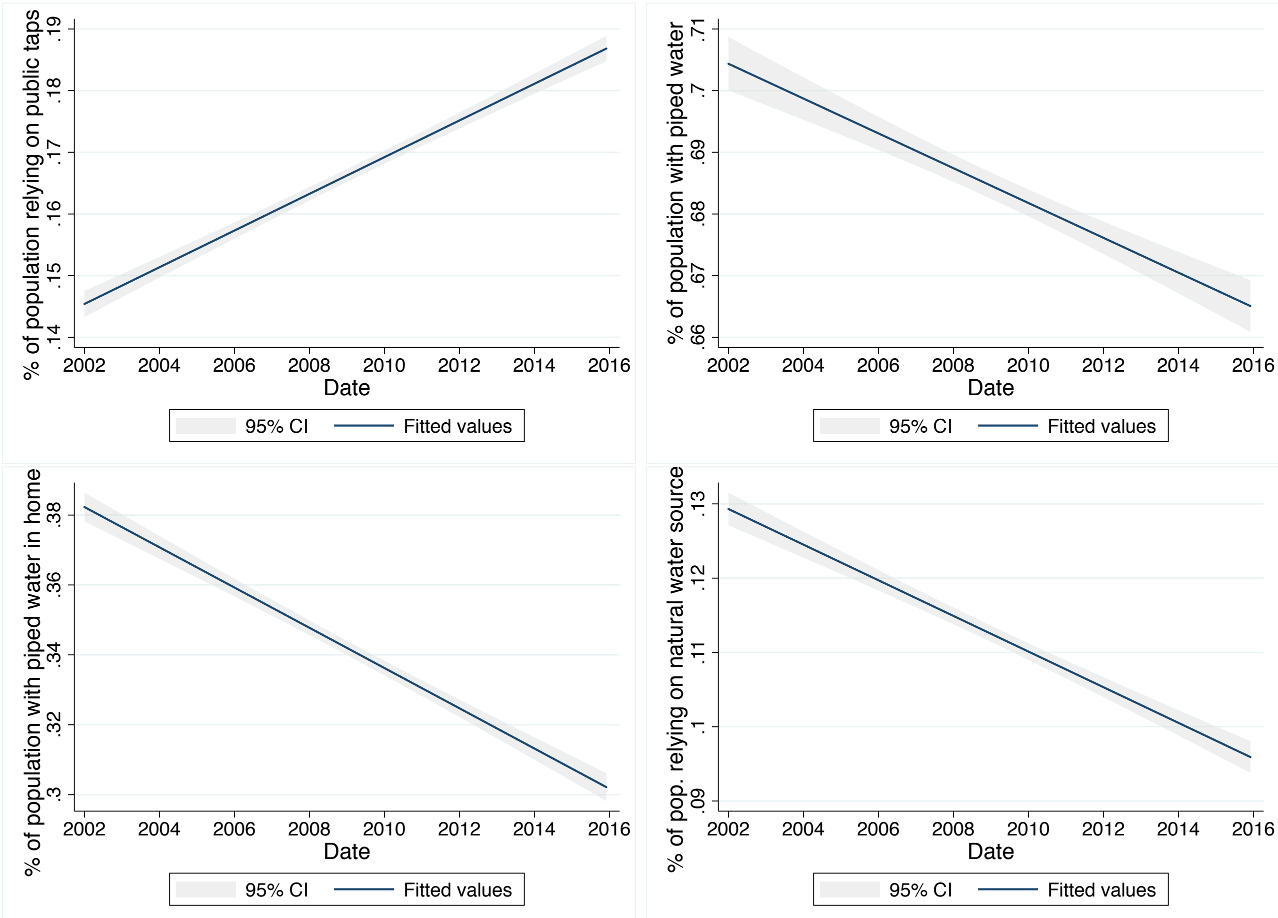


Figure 5: Primary water source population proportion time trends, 2002-2015 (Top left: public taps, top right: piped water, bottom left: piped water in home, bottom right: natural water source)

5 Results

Table 2 reports regression results for the full sample from 1997 to 2015. All results reported were obtained using the fixed-effect model for panel data regression; thus the coefficient estimates are based on the within-effect estimator. Panel groups in the monthly data are divided by province, gender, and age group, with a total of $9 \times 2 \times 10 = 180$ "cells" of population subgroups. Standard errors are clustered by cell. Each cell is observed over 216 months (from 1997 to 2015), yielding a total of $38,880$ possible observations; omitting missing data and outliers[10], the final number of observations is $N = 35,716$ for the full sample.

Columns 1 through 3 estimate the mitigating effect of increased water availability along the extensive and intensive margins on heat-related mortality for the full sample. The estimated coefficients for both regressors are relatively stable as demographic controls and fixed effects are

[10] Volumetric flow measures above the 95th percentile were excluded to avoid capturing the effect of flooding on mortality.

Table 2: Regression Results with Triple Interactions

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	Deaths	Deaths	Deaths	Deaths	Deaths	Deaths
Sample	Full	Full	Full	Full	Full	Full
Water availability index × Maximum temperature > 100 F	-10.99	-11.50	-13.37	-11.40	-27.07	-13.32
	(4.186)	(4.166)	(4.327)	(4.225)	(7.361)	(5.946)
× Higher piped water access				-25.01		
				(19.89)		
× Higher gastroenteritis rate					27.78	
					(9.858)	
× Higher % Black						-2.154
						(9.805)
Pipe flow on × Maximum temperature > 100 F	-7.83	-9.362	-6.202	-5.236	-0.944	-16.38
	(5.027)	(3.887)	(3.282)	(3.045)	(3.664)	(4.252)
× Higher piped water access				-6.953		
				(6.230)		
× Higher gastroenteritis rate					-12.45	
					(3.980)	
× Higher % Black						17.68
						(5.603)
Maximum temperature > 100 F	14.125**	7.162	13.81	12.69	11.27	20.43
	(5.589)	(3.623)	(4.071)	(4.938)	(4.382)	(8.317)
N	35716	35716	35716	35716	35716	35716
Mean of dep. var.	260.9	260.9	260.9	260.9	260.9	260.9
Demographic controls	No	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	No	No	Yes	Yes	Yes	Yes
R-squared	0.05	0.277	0.419	0.419	0.443	0.420

Standard errors in parentheses

"Water availability index" refers to the measure of potable water availability described in equation 1. "Pipe flow on" is a dummy which equals 1 when the volumetric flow into the pipe to a sanitation facility exceeds the 25th within-province percentile. "Maximum temperature > 100 F" is a dummy that equals 1 when the maximum temperature exceeds 100° F during at least one day in a particular month in a particular province.

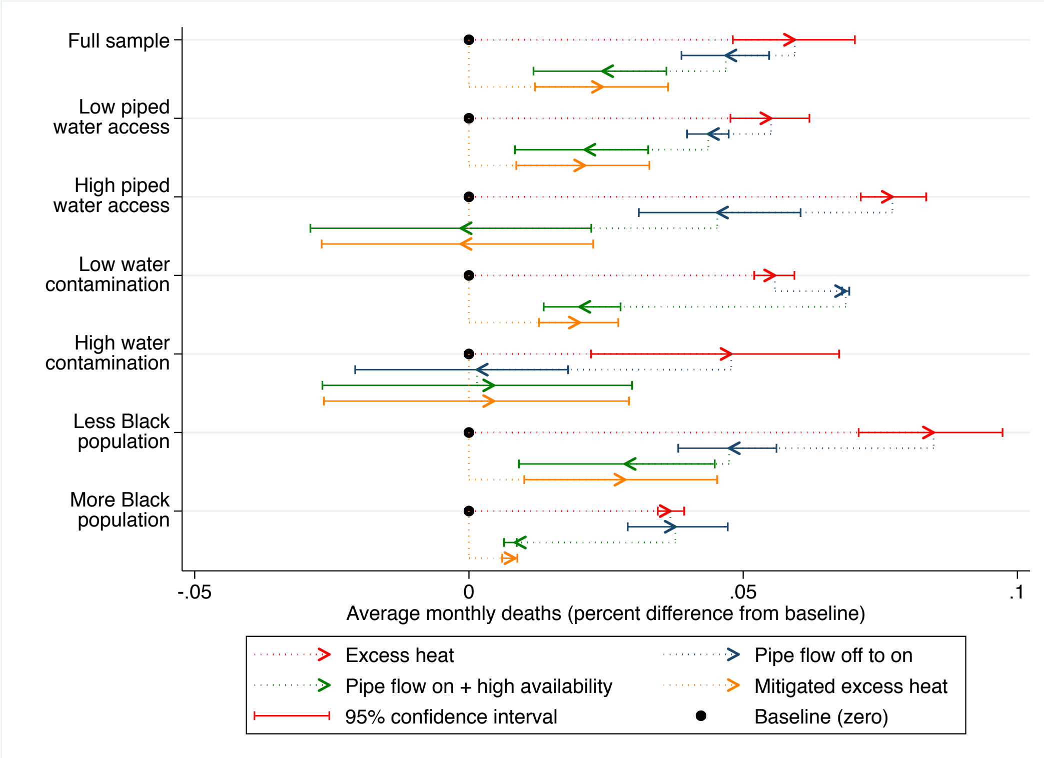
added. As column 3 of Table 2 demonstrates, with full demographic controls and fixed effects in the model, both coefficients of interest are statistically significant at the $\alpha = 0.01$ and the $\alpha = 0.1$ levels, respectively. This suggests mitigating effects of water availability on heat-related mortality at both the extensive and intensive margins. Both coefficients are relatively stable as demographic controls and time fixed effects are included. Using the `psacalc` package in Stata/SE version 15 to assess robustness to possible selection on unobservables (Oster 2013), I find a δ bound estimate of 1.49 for the interaction between the availability index and the temperature threshold and -2.31[11] on the interaction with "Pipe flow on," both well above the recommended standard of 1.0 in absolute value.[12] As explained in Oster (2013), these estimates mean that the degree of selection on unobservables necessary to nullify the estimated effects are 1.49 and 2.31 times the degrees of selection on observables, respectively.

Columns 4 through 6 of Table 2 include triple interactions in the model to assess three possible types of heterogeneity. Since infrastructural water needs to be accessible to have any effect on mortality, it is likely that the effect of infrastructural availability is stronger in areas where households are more likely to rely on infrastructural sources. Put another way, the significance of the coefficient on the "Pipe flow on" interaction may be attenuated by the inclusion of groups with low access in the sample. Also, since infrastructural sources go through sanitation processes that natural sources do not, infrastructural availability is likely to be relatively more significant when natural water sources are highly contaminated. Finally, post-apartheid South Africa remains highly stratified by race, with large differences in average wealth across Black and White individuals and a high degree of spatial segregation. In particular, if water access is highly unequal, it is possible that Black individuals in South Africa would be more directly affected by a decrease in natural water availability, and thus more directly vulnerable to the long-run impacts of climate change.

Figure 6 represents the estimated treatment effects for the full sample and triple-interaction subgroups graphically. The cumulative mitigating effect of increased water availability on heat-related mortality is the difference between the red "excess heat" vector and the orange "mitigated excess heat" vector. In every case, the mitigated effect is significantly smaller than the unmitigated effect, and in some cases, increased water availability makes the effect of excess heat statistically indistinguishable from zero. This cumulative effect is decomposed in the blue "pipe flow off to on" and green "pipe flow on + high availability" vectors, revealing heterogeneity in the type of water availability that is most significant for each subpopulation or circumstance. The most striking of these differences is based on the level of water contamination, as measured by the prevalence of

[11] The negative coefficient is because the coefficient gets larger in absolute value with the inclusion of controls.

[12] To obtain this result for the availability index, it is necessary to control for *InfraSource* in the "uncontrolled" model. This is a side effect of the way the availability index was constructed as a weighted average of natural and infrastructural availability—without this control, the coefficient picks up the higher general mortality rate associated with higher reliance on natural sources of water. To obtain the result for the piped water on dummy variable, in order to isolate the extensive margin, it is necessary to control for the availability index interaction.



"Baseline" in this figure refers to a month in which the maximum temperature did not exceed 100° F and the water availability index is at its mean. "Excess heat" is defined as the maximum temperature exceeding 100° F on at least one day in a particular month within a particular province. "Pipe flow on" is a dummy which equals 1 when the volumetric flow into the pipe to a sanitation facility exceeds the 25th within-province percentile. When this dummy equals 0, the pipe flow is considered "off." "High availability" refers to the water availability index exceeding the mean by one standard deviation. "Mitigated excess heat" is the vector sum of "Excess heat," "Pipe flow off to on," and "Pipe flow on + high availability," representing the effect of excess heat on mortality mitigated by increased water availability. A tabular representation of this figure is available in appendix section 8.1.

Figure 6: Estimated Mitigating Treatment Effects

gastroenteritis and infectious diarrhea as a cause of death in a particular month within a particular province. When natural sources of water are highly contaminated, the entire mitigating effect of water availability is along the extensive margin of infrastructural availability, reflecting the increased importance of sanitation during outbreaks of waterborne disease. In contrast, when the natural sources are relatively safe, the entire effect is along the intensive margin of general availability. In this case, households enjoy the mitigating effects of water availability regardless of their primary source of drinking water.

5.1 Robustness Check: Coefficient Stability at Various Thresholds

Figures 7 through 10 plot the coefficient of key covariates (the interaction term between water availability and excess heat, the excess heat dummy, the non-interaction water availability in-

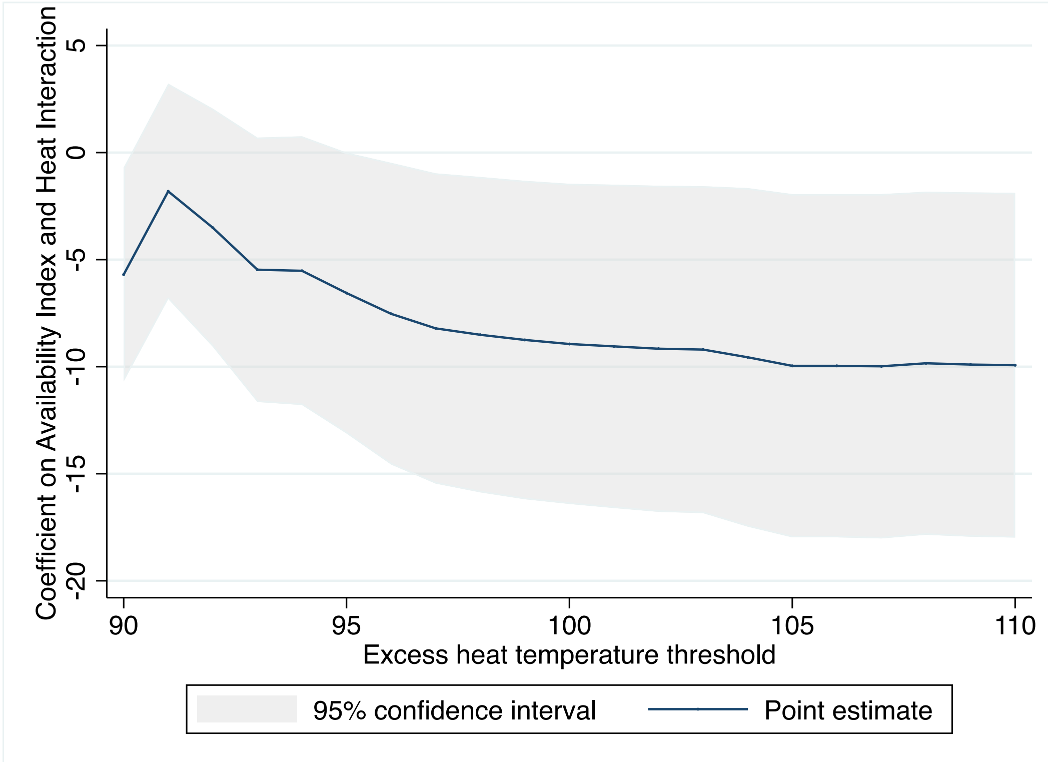


Figure 7: Water Availability Index and Heat Interaction Term Coefficient at Various Temperature Thresholds

dex, and the percentage of the population in each subgroup with access to infrastructural water sources, respectively) over a range of excess heat thresholds from 90° F to 110° F. (Recall that the threshold used throughout the paper was 100° F.) The label above each point in each graph is the statistical significance of the coefficient at that point. As the figures demonstrate, the sign of the point estimate is consistent for every covariate over this entire range. The coefficient on the availability index interaction term is negative and statistically significant at the $\alpha = 0.05$ level at all thresholds above 93° F, which is too low to identify actual excess heat. The coefficient point estimate is remarkably stable above 98° F, which is very close to the inflection point of the monthly temperature-mortality relationship in Figure 2. The coefficient on the "piped water on" interaction term is statistically significant at every threshold tested.

6 Economic Impact

Using the regression model estimated in the previous section, I predict the changes in mortality rate associated with various possible excess heat incidence rates and water availability measures. These predictions should be interpreted as back-of-the-envelope calculations that demonstrate the economic significance of the findings in section 4; they do not take into account the nonlinear relationship between heat and mortality, or the correlation between climate and several other

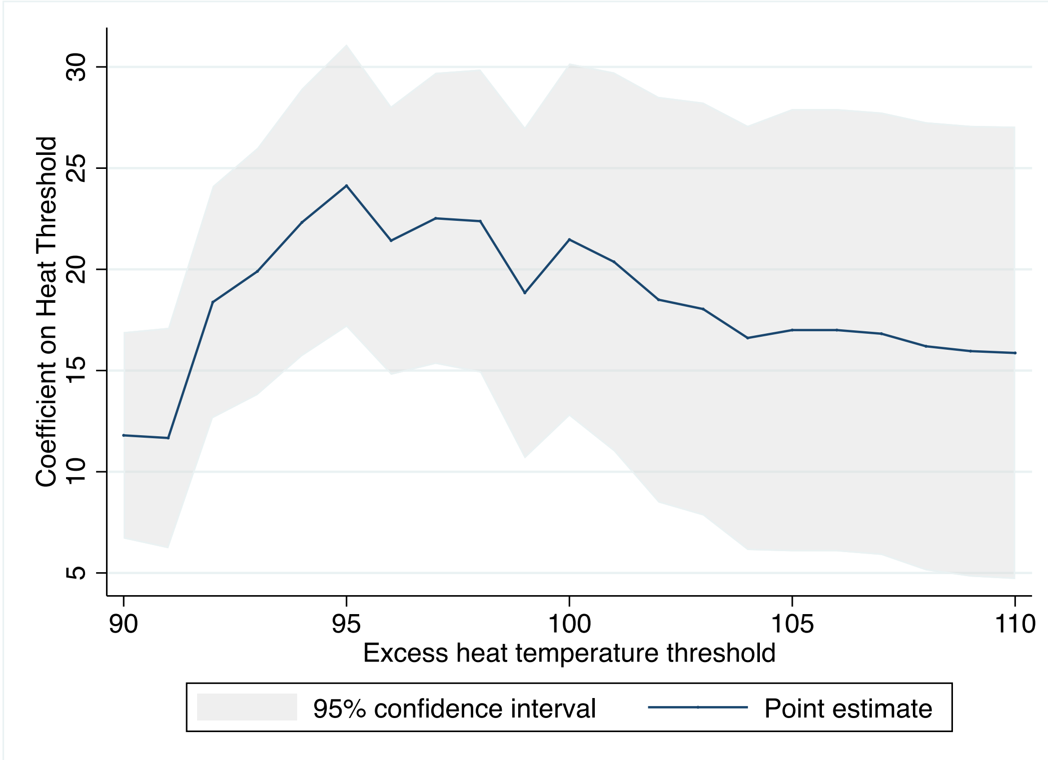


Figure 8: Heat Threshold Coefficient at Various Temperature Thresholds

determinants of mortality (e.g. agricultural production and food scarcity, disease outbreaks, air pollution, etc.). However, these omissions are very likely to bias the predictions down rather than up. Thus these estimates are likely to be lower bounds. Unless otherwise stated, all estimates are conditional on the maximum temperature exceeding 100° F on at least one day in a particular month.

I find that each additional month in which the maximum temperature exceeds 100° F on at least one day in a particular province results in, on average, 5.4 additional deaths. A 10% reduction in both infrastructural and natural water availability, in line with the predictions of Nkhonjera et al. (2017) by 2050, leads to an average of 1.2 additional deaths per month per province. A 50% reduction in infrastructural water availability, in line with the current water consumption limitations mandated by the municipal government of Cape Town to prevent reservoir depletion, leads to 5.5 additional deaths per month per province. A complete collapse in infrastructural water availability, such as the "Day Zero" scenario in which Cape Town's reservoirs are entirely depleted, leads to 10.9 additional deaths per month per province. Using the lower-bound VSL of US$230,000 in South Africa (Laxminarayan et al. 2007), the economic cost of these scenarios respectively range from US$13.66 million (10% reduction) to US$33.74 million (Day Zero) per month in which the maximum temperature exceeds 100° F on at least one day.

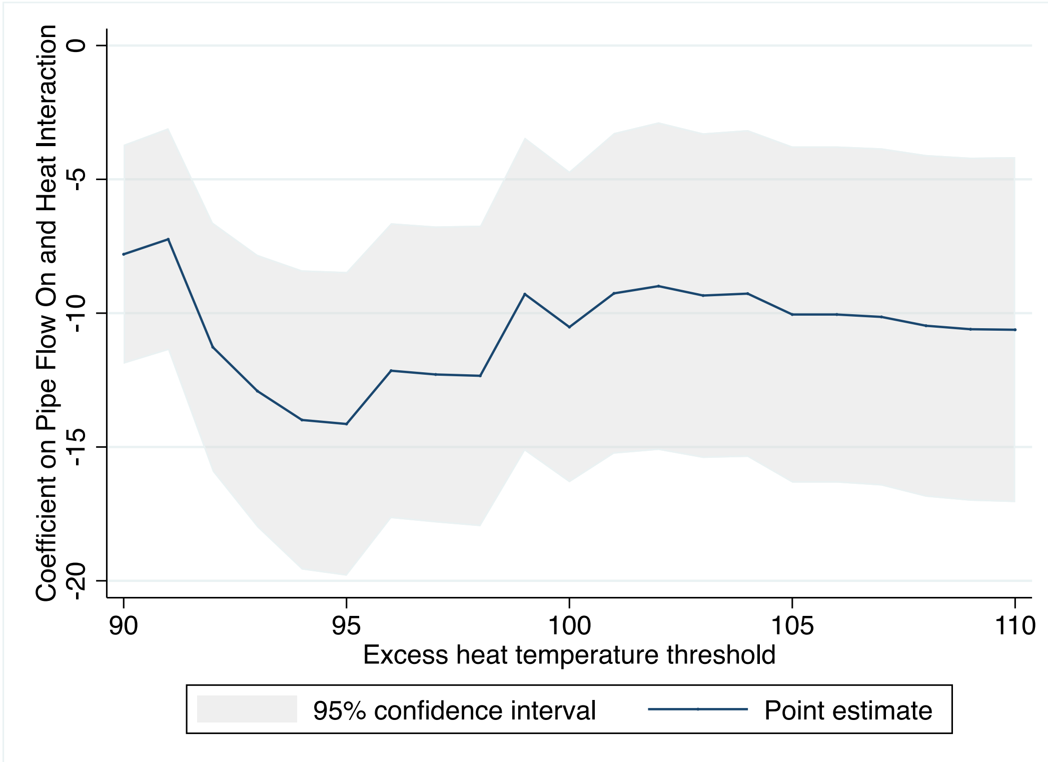


Figure 9: Piped Water On and Heat Interaction Term Coefficient at Various Temperature Thresholds

7 Conclusion

In this paper, I identify a novel, indirect channel by which rising global temperatures may exact a global mortality cost. Alongside the obvious direct effect of higher temperatures on heat-related mortality, the diminishment of ecological water resources associated with climate change in tandem with these higher temperatures will further amplify the risk, especially among vulnerable populations with limited access to high-quality water distribution and sanitation systems. Thus studies of the long-run costs of climate change that only consider the direct effects of climate change may substantially underestimate the total cost by failing to consider this interaction. While much of this literature focuses on projecting the medium- to long-term costs of climate change, South Africa is already in the throes of a water crisis. Cape Town, one of its largest cities, has seen its reservoirs dwindle to critical levels in recent years, with aggressive campaigns for individuals to limit their water consumption to stave off "Day Zero"—the day when the reservoirs run dry. The findings in this paper suggest the mortality consequences of a "Day Zero," should it come, may dramatically sharpen the relationship between extreme heat and mortality rates, especially among the Black population, as well as the young and those more vulnerable to waterborne disease.

These results also suggest that investments in water infrastructure can be an effective policy to defend households against negative shocks to heat and natural water availability, such as those threatened by climate change. As discussed in Section 3, moving from the lowest decile of piped water access to the highest appears to eliminate most of the heat-mortality relationship, even in the absence of widespread residential air conditioning. As the results in Section 5 suggest, investing in water infrastructure to increase the number of households with access to piped water and improve the quality, reliability, and abundance of piped water reduces the significance of negative shocks to natural water availability in reducing the population's dependence on natural water sources. Thus any technology which improves the efficiency of sanitation and distribution systems, or any technology that establishes new sources of potable water could significantly reduce the long-run mortality consequences of climate change. Some such technologies are already being developed, such as rainwater harvesting devices that provide quick and low-cost access to residential water without needing to build pipes by collecting and processing rainwater on-site, and desalination technology that can make salt water, the majority of the planet's water, potable. These results suggest that investment in research and development of these technologies should be a central part of any plan to combat climate change.

The most immediate opportunity for further research is to assess the impact and efficacy of Cape Town's attempts to avoid "Day Zero," and the general efficacy of adaptation initiatives at the municipal level to counteract the local effects of climate change. These attempts include behavioral interventions such as the Two-Minute Shower Songs[13] campaign, which enlisted famous musicians to record two-minute versions of their most popular songs for individuals to listen to in the shower. While the results of the campaign are thought to be positive, with widespread coverage and positive response from individuals, research is needed to identify any causal effect of this campaign. In a similar vein, further research is needed to study the underlying causes of the increasing proportion of the population in South Africa (or elsewhere) relying on public municipal taps as a primary water source, and evaluate the trade-off a local government faces between investing in public water infrastructure and subsidizing the construction of private residential water pipelines to guarantee their constituents' constitutional, and human, right to water.

References

[1] M. R. Allen, V. R. Barros, J. Broome, W. Cramer, R. Christ, J. A. Church, L. Clarke, Q. Dahe, P. Dasgupta, and N. K. Dubash. IPCC fifth assessment synthesis report-climate change 2014 synthesis report. 2014.

[2] C. K. Ao. Essays on the impact of clean water on human capital and productivity. PhD Thesis, 2016.

[3] R. Banerjee and R. Bhowmick. Heat, Infant Mortality and Adaptation: Evidence from India. 2016.

[13] See https://www.npr.org/2018/09/07/644918801/singing-in-the-shower-to-help-save-cape-towns-water

[4] A. Barreca, K. Clay, O. Deschenes, M. Greenstone, and J. S. Shapiro. Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century. Journal of Political Economy, 124(1):105–159, 2016.

[5] A. Barreca, K. Clay, O. Deschênes, M. Greenstone, and J. S. Shapiro. Convergence in Adaptation to Climate Change: Evidence from High Temperatures and Mortality, 1900–2004. American Economic Review, 105(5):247–251, 2015.

[6] A. Barreca and Jay P Shimshack. Absolute Humidity, Temperature, and Influenza Mortality: 30 Years of County-Level Evidence from the United States. American journal of epidemiology, 176(7):S114, Oct. 2012.

[7] A. I. Barreca. Essays on weather, mortality, and economic wellbeing. PhD Thesis, 2008.

[8] A. I. Barreca. Climate change, humidity, and mortality in the United States. Journal of Environmental Economics and Management, 63(1):19–34, Jan. 2012.

[9] X. Basagaña, C. Sartini, J. Barrera-Gómez, P. Dadvand, J. Cunillera, B. Ostro, J. Sunyer, and M. Medina-Ramón. Heat Waves and Cause-specific Mortality at all Ages. Epidemiology, 22(6):765–772, 2011.

[10] B. Beach, J. Ferrie, M. Saavedra, and W. Troesken. Typhoid Fever, Water Quality, and Human Capital Formation. The Journal of Economic History, 76(01):41–75, Mar. 2016.

[11] M. F. Bergeron. Reducing Sports Heat Illness Risk. Pediatrics in Review, 34(6):270–279, June 2013.

[12] A. Comfort. Long-term effect of in utero conditions on maternal survival later in life: evidence from Sub-Saharan Africa | SpringerLink.

[13] F. C. Curriero, K. S. Heiner, J. M. Samet, S. L. Zeger, L. Strug, and J. A. Patz. Temperature and Mortality in 11 Cities of the Eastern United States. American Journal of Epidemiology, 155(1):80–87, Jan. 2002.

[14] P. Debaere. The Global Economics of Water: Is Water a Source of Comparative Advantage? American Economic Journal: Applied Economics, 6(2):32–48, 2014.

[15] T. Deryugina and S. M. Hsiang. Does the Environment Still Matter? Daily Temperature and Income in the United States. Working Paper 20750, National Bureau of Economic Research, Dec. 2014.

[16] O. Deschenes. Temperature, human health, and adaptation: A review of the empirical literature. Energy Economics, 46:606–619, Nov. 2014.

[17] O. Deschenes and M. Greenstone. The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather: Reply. The American Economic Review, 102(7):3761–3773, 2012.

[18] O. Deschênes and M. Greenstone. The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather. The American Economic Review, 97(1):354–385, 2007.

[19] O. Deschênes, M. Greenstone, and J. Guryan. Climate Change and Birth Weight. American Economic Review, 99(2):211–217, May 2009.

[20] J. P. Ferrie and W. Troesken. Water and Chicago's mortality transition, 1850-1925. Explorations in Economic History, 45(1):1–16, 2008.

[21] A. C. Fisher, W. M. Hanemann, M. J. Roberts, and W. Schlenker. The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather: Comment. The American Economic Review, 102(7):3749–3760, 2012.

[22] J. Goodman, M. Hurwitz, J. Park, and J. Smith. Heat and Learning. Working Paper 24639, National Bureau of Economic Research, May 2018.

[23] S. Hajat and T. Kosatky. Heat-related mortality: a review and exploration of heterogeneity. Journal of epidemiology and community health, 64(9):753, 2010.

[24] R. Hanna and P. Oliva. Implications of Climate Change for Children in Developing Countries. The Future of Children, 26(1):115–132, 2016.

[25] A. Isen, M. Rossin-Slater, and W. R. Walker. Every Breath You Take - Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970. Working Paper 19858, National Bureau of Economic Research, Jan. 2014.

[26] J. E. Jackson, M. G. Yost, C. Karr, C. Fitzpatrick, B. K. Lamb, S. H. Chung, J. Chen, J. Avise, R. A. Rosenblatt, and R. A. Fenske. Public health impacts of climate change in Washington State: projected mortality risks due to heat events and air pollution. Climatic Change, 102(1-2):159–186, Sept. 2010.

[27] R. S. Kovats, S. Hajat, and P. Wilkinson. Contrasting patterns of mortality and hospital admissions during hot weather and heat waves in Greater London, UK. Occupational and Environmental Medicine, 61(11):893–898, Nov. 2004.

[28] M. Kremer, J. Leino, E. Miguel, and A. P. Zwane. Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions. The Quarterly Journal of Economics, 126(1):145–205, 2011.

[29] M. Kudamatsu, T. Persson, and D. Strömberg. Weather and infant mortality in Africa. 2012.

[30] J. G. C. Laurent, A. Williams, Y. Oulhote, A. Zanobetti, J. G. Allen, and J. D. Spengler. Reduced cognitive function during a heat wave among residents of non-air-conditioned buildings: An observational study of young adults in the summer of 2016. PLOS Medicine, 15(7):e1002605, July 2018.

[31] R. Laxminarayan, E. Klein, C. Dye, K. Floyd, S. Darley, and O. Adeyi. Economic benefit of tuberculosis control. Policy research working paper, 4295:16–17, 2007.

[32] N. S. Ngo and R. M. Horton. Climate change and fetal health: The impacts of exposure to extreme temperatures in New York City. Environmental Research, 144:158–164, Jan. 2016.

[33] G. K. Nkhonjera. Understanding the impact of climate change on the dwindling water resources of South Africa, focusing mainly on Olifants River basin: A review. Environmental Science & Policy, 71:19–29, 2017.

[34] R. Rocha and R. R. Soares. Water scarcity and birth outcomes in the Brazilian semiarid. Journal of Development Economics, 112:72–91, Jan. 2015.

[35] M. Tressol, C. Ordonez, R. Zbinden, J. Brioude, V. Thouret, C. Mari, P. Nedelec, J.-P. Cammas, H. Smit, H.-W. Patz, and A. Volz-Thomas. Air pollution during the 2003 European heat wave as seen by MOZAIC airliners. Atmospheric Chemistry and Physics, 8(8):2133–2150, Apr. 2008.

[36] W. Troesken, J. P. Ferrie, and K. Rolf. Cognitive Disparities, Lead Plumbing, and Water Chemistry: Intelligence Test Scores and Exposure to Water-Borne Lead Among World War Two U.S. Army Enlistees. Technical report, National Bureau of Economic Research, 2011.

[37] S. Wallsten and K. Kosec. The effects of ownership and benchmark competition: An empirical analysis of U.S. water systems. International Journal of Industrial Organization, 26(1):186–205, 2008.

[38] J. Wilde, B. Apouey, and T. Jung. Heat waves at conception and later life outcomes. University of South Florida Working Paper, 2014.

[39] J. G. Zivin and M. Neidell. Environment, Health, and Human Capital. Journal of Economic Literature, 51(3):689–730, 2013.

[40] J. G. Zivin and J. Shrader. Temperature Extremes, Health, and Human Capital on JSTOR.

[41] J. S. G. Zivin, Y. Song, Q. Tang, and P. Zhang. Temperature and High-Stakes Cognitive Performance: Evidence from the National College Entrance Examination in China. Working Paper 24821, National Bureau of Economic Research, July 2018.

8 Appendix

8.1 Estimated Treatment Effects (Figure 6)

Tables 3 and 4 display the estimated treatment effects depicted in Figure 6 in tabular form, and include a significance test for the proportional differences[14]. All three hypothesized types of heterogeneity are statistically significant on the "Pipe flow on" interaction. As is intuitive, whether or not water is flowing into reservoirs for distribution is more significant for populations more likely to have piped water access; the similar heterogeneity along race lines thus suggests significant racial inequality of access to, or quality of, piped water. The significant difference in coefficients based on gastroenteritis rate demonstrates the effect is driven by safe-to-drink water only. Since gastroenteritis is among the most common waterborne diseases in South Africa, the rate of deaths attributed to it is highly correlated with the average quality of drinking water. As Table 4 shows, when the gastroenteritis rate is high (and thus water quality is low), the effect loads entirely on infrastructural availability, reflecting the necessity of sanitation. When the rate is low, the effect loads entirely on the intensive margin of general availability because the gap in quality between natural and infrastructural sources is smaller. While the point estimates for the predicted effect of an intensive margin increase in water availability suggest the effect is stronger for populations with higher water access and proportionally more Black populations, the differences are not significant at conventional levels. The cumulative difference is only significant across levels of piped water access, with a striking estimated reduction in heat-related deaths of 13.8%, although much smaller reductions cannot be ruled out at the 95% level of confidence.

Table 3: Estimated Overall Treatment Effects on Mortality

	Predicted Change in Number of Deaths per Month	P-value
Pipe flow off $\rightarrow$ on	-3.27	0.20
One standard deviation increase in water availability index, pipe flow on	-5.85	<0.01
Cumulative	-9.12	<0.01
Mean deaths per month	260.9	

[14]Since the average number of deaths per month is different across the population subgroups, the level differences are not necessarily comparable; thus I only present a test of significance for the proportional difference.

Table 4: Heterogeneity of Estimated Treatment Effects on Mortality

Predicted Change in Number of Deaths per Month				
		Below Median	Above Median	Difference p-value
Pipe flow off → on				
	Piped water access	-2.92	-8.59	
		(-1.1%)	(-5.6%)	0.05
	Gastroenteritis rate	+3.14	-12.98	
		(+1.6%)	(-4.5%)	<0.01
	% Black	-10.12	+0.23	
		(-3.3%)	(+0.1%)	0.03
One standard deviation increase in water availability index, pipe flow on				
	Piped water access	-5.76	-12.57	
		(-2.2%)	(-8.2%)	0.22
	Gastroenteritis rate	-11.79	+0.85	
		(-5.9%)	(+0.2%)	<0.01
	% Black	-5.15	-7.56	
		(-1.7%)	(-3.6%)	0.16
Cumulative				
	Piped water access	-8.68	-21.16	
		(-3.3%)	(-13.8%)	0.05
	Gastroenteritis rate	-8.65	-12.13	
		(-4.3%)	(-4.3%)	0.96
	% Black	-15.27	-7.33	
		(-5.0%)	(-3.5%)	0.44

8.2 Graphical Estimated Treatment Effects in Levels

Figure 10 presents the same estimated treatment effects as Figure 6 in levels rather than percentages.

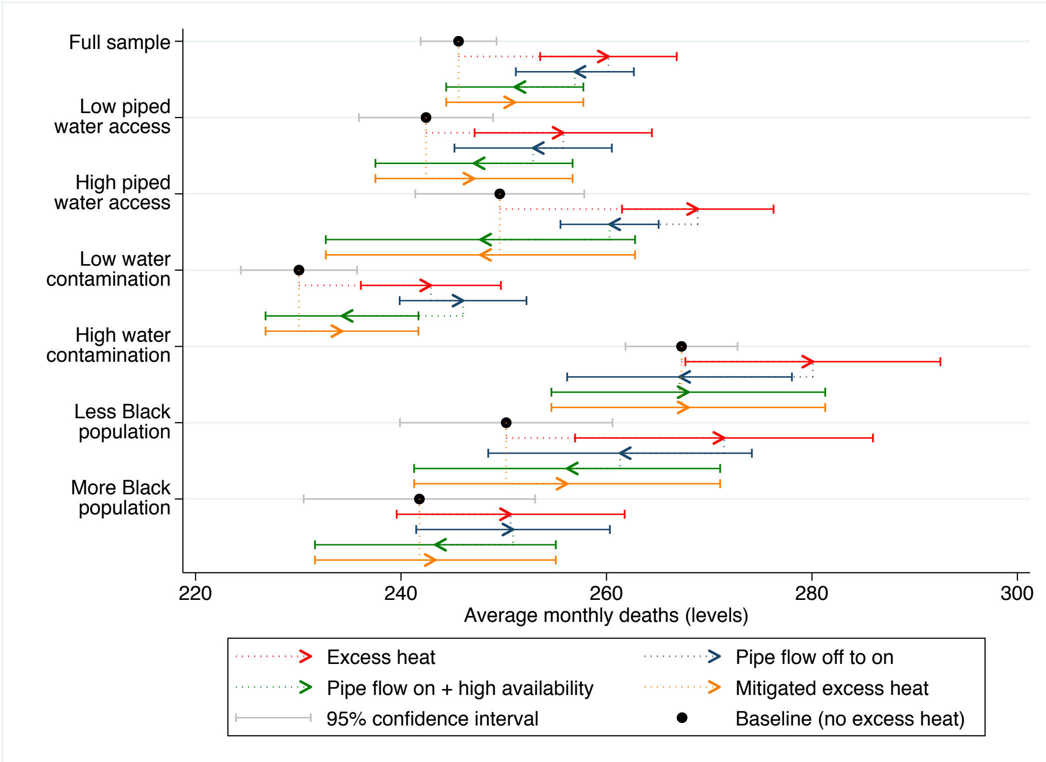


Figure 10: Estimated Treatment Effects in Levels

8.3 Acclimatization

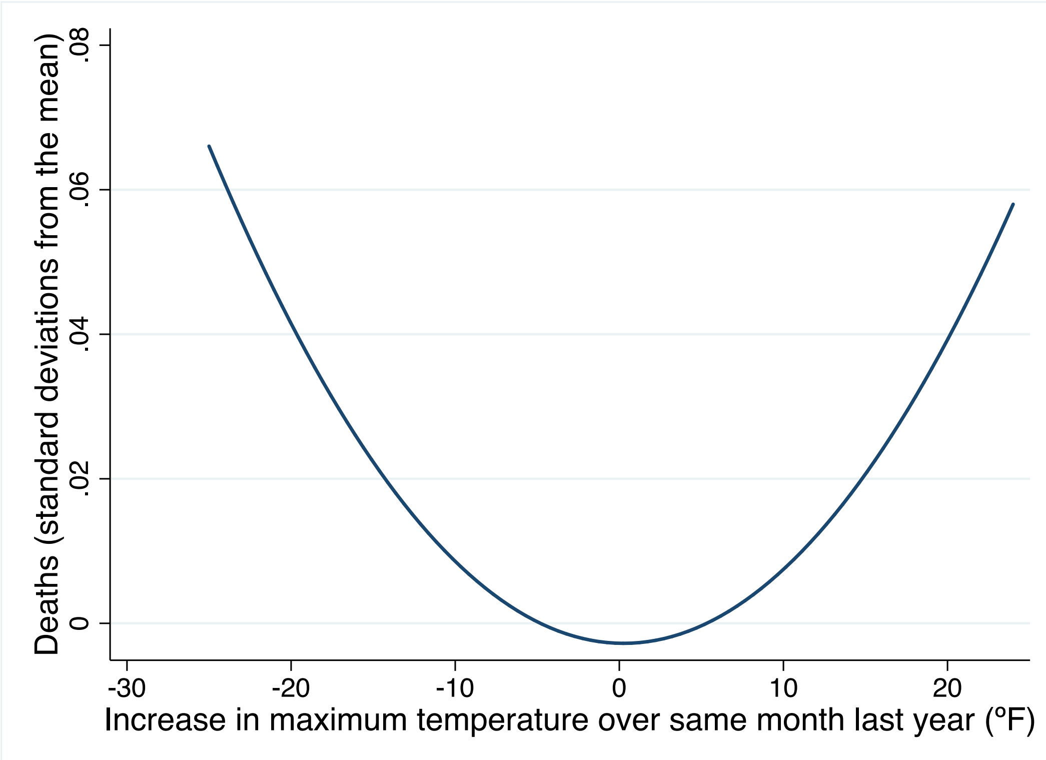


Figure 11: Deaths and Annual Deviations in Temperature in South Africa, 1997-2015

One potential concern with estimating the mortality consequences of rising global temperatures is that, even if the mortality rate increases from already hot days getting hotter, the mortality rate on colder days could *decrease*. Colder temperatures are indeed associated with higher mortality rates (Curriero et al. 2002) and the reduction in mortality from the decrease in number of cold days could attenuate or even reverse the increase from excess heat. To rule this out, Figure 11 plots the relationship between the death rate and the annual change in temperature. The strong U-shape relationship suggests that *deviations* in temperature from the norm drive at least part of the temperature-mortality relationship. This is consistent with the epidemiology literature on ambient temperature, in which the thresholds above (below) which the mortality rate begins to increase from excess heat (cold) are specific to the regular climate of the area.

8.4 Robustness Check: Province-Specific Temperature Thresholds

To address the potential concern that the above results are biased by climate differences across provinces, Table 5 replicates Table 2 using the 97th percentile of each province's maximum temperature as the threshold. All qualitative findings are comparable.

Table 5: Regression Results with Triple Interactions and Province-Specific Temperature Thresholds

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	Deaths	Deaths	Deaths	Deaths	Deaths	Deaths
Sample	Full	Full	Full	Full	Full	Full
Water availability index ×	-10.57	-10.62	-13.70	-11.33	-25.71	-12.59
Maximum temperature > 100 F	(4.245)	(4.092)	(4.370)	(4.141)	(7.235)	(6.332)
× Higher piped water access				-101.9		
				(45.81)		
× Higher gastroenteritis rate					24.94	
					(9.887)	
× Higher % Black						-2.154
						(9.805)
Pipe flow on × Maximum	-11.14	-9.486	-6.819	-7.481	-2.261	-13.86
temperature > 100 F	(4.162)	(3.528)	(3.155)	(3.033)	(3.438)	(3.712)
× Higher piped water access				5.537		
				(4.600)		
× Higher gastroenteritis rate					-9.556	
					(3.739)	
× Higher % Black						12.82
						(5.286)
Maximum temperature > 97th percentile	27.48	4.600	7.896	10.57	4.836	11.09
	(5.357)	(3.380)	(3.788)	(4.876)	(4.565)	(6.839)
N	27418	27418	27418	27418	27418	27418
Mean of dep. var.	260.9	260.9	260.9	260.9	260.9	260.9
Demographic controls	No	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	No	No	Yes	Yes	Yes	Yes
R-squared	0.05	0.277	0.419	0.419	0.442	0.420

Standard errors in parentheses

"Water availability index" refers to the measure of potable water availability described in equation 1. "Pipe flow on" is a dummy which equals 1 when the volumetric flow into the pipe to a sanitation facility exceeds the 25th within-province percentile. "Maximum temperature > 97th percentile" is a dummy that equals 1 when the maximum temperature exceeds the 97th within-province percentile during at least one day in a particular month in a particular province.

8.5 Subgroup Regressions Instead of Triple Interactions

Tables 6 through 8 present regression results analogous to columns 4 through 6 of Table 2, but instead of triple interactions, the domain of each regression is restricted to the associated subsample. All qualitative findings are comparable to the conclusions of Section 5.

Table 6: Regression Results by Water Access and Quality

	(5)	(6)	(7)	(8)
Dep. var.	Deaths	Deaths	Deaths	Deaths
Sample	Low Water Access	High Water Access	Lower % Gastro	Higher % Gastro
Water availability index × Maximum temperature > 100 F	-15.70 (4.461)	-20.06 (22.83)	-13.34 (5.268)	1.709 (4.542)
Pipe flow on × Maximum temperature > 100 F	-19.43 (3.806)	-2.437 (4.023)	-10.85 (3.284)	-8.158 (2.953)
Maximum temperature > 100 F	21.56 (6.417)	18.07 (4.944)	17.23 (3.718)	16.51 (3.587)
Mean of dep. var.	298.9 (215.8)	221.5 (193.4)	206.4 (161.6)	309.8 (232.8)
Pred. Δ deaths from std. dev. ↑ in availability index	-7.89	-8.15	-4.64	+0.22
Pred. Δ deaths from pipe flow off → on	-9.68	-13.01	-12.45	-8.89
N	13964	13454	12969	14449
Demographic controls	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes
R-squared	0.457	0.270	0.277	0.287

Standard errors in parentheses

p<0.1, p<0.05, p<0.01

"Water availability index" refers to the measure of potable water availability described in equation 1. "Maximum temperature > 100 F" is a dummy that equals 1 when the maximum temperature exceeds 100° F during at least one day in a particular month in a particular province. "Pred. Δ deaths from std. dev. ↑ in availability index" is a linear prediction of the dependent variable, the average monthly number of deaths in a province, if the availability index increased by one standard deviation. "Low Water Access" refers to population subgroups in which the proportion of individuals with access to infrastructural water sources is below the sample median. "Lower % Gastro" refers to population subgroups and months during which the percentage of deaths that were attributed to infectious diarrhea or gastroenteritis was below the sample median.

Table 7: Regression Results for Selected Population Subgroups

	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Dep. var.	Deaths	Deaths	Deaths	Deaths	Deaths	Deaths	Deaths
Sample	Male	Female	Less Educated	More Educated	Age < 30	Age 30-60	Age > 60
Water availability index × Maximum temperature > 100 F	-9.586	-11.09	-4.302	-15.87	-11.68	-7.621	-1.043
	(5.692)	(5.574)	(3.331)	(8.367)	(6.534)	(5.404)	(3.123)
Pipe flow on × Maximum temperature > 100 F	-9.562	-9.135	-6.133	-15.68	-7.779	-12.26	-6.306
	(4.428)	(4.458)	(3.147)	(5.123)	(5.823)	(4.347)	(3.411)
Maximum temperature > 100 F	17.72	23.88	16.14	27.54	22.44	24.32	12.46
	(6.263)	(7.115)	(4.138)	(7.923)	(9.518)	(5.846)	(3.519)
Mean of dep. var.	268.6	253.3	194.0	330.3	261.7	321.1	193.3
	(221.1)	(195.4)	(167.4)	(224.1)	(242.3)	(188.0)	(147.7)
Pred. Δ deaths from std. dev. $\uparrow$ in availability index	-5.02	-6.21	-2.11	-8.65	-6.62	-3.97	-0.65
Pred. Δ deaths from pipe flow off $\rightarrow$ on	-11.10	-13.51	-8.79	-18.50	-16.90	-12.68	-6.92
N	13709	13709	13936	13482	11304	8478	7636
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.316	0.394	0.220	0.498	0.517	0.518	0.444

Standard errors in parentheses

$p<0.1$, $p<0.05$, $p<0.01$

"Water availability index" refers to the measure of potable water availability described in equation 1. "Maximum temperature > 100 F" is a dummy that equals 1 when the maximum temperature exceeds 100° F during at least one day in a particular month in a particular province. "Pred. Δ deaths from std. dev. $\uparrow$ in availability index" is a linear prediction of the dependent variable, the average monthly number of deaths in a province, if the availability index increased by one standard deviation. "Less Educated" refers to population subgroups in which the proportion of individuals that have attained at least a high school diploma is below the sample median.

Table 8: Regression Results by Race and Water Access

	(16)	(17)	(18)	(19)
Dep. var.	Deaths	Deaths	Deaths	Deaths
Sample	Higher % Black	Higher % Black, Low Water Access	Lower % Black	Lower % Black, High Water Access
Water availability index × Maximum temperature > 100 F	-18.83	-22.33	-11.50	-43.67
	(8.455)	(9.223)	(4.919)	(52.11)
Water availability index × Maximum temperature > 100 F	-19.20	-16.68	-19.16	-6.14
	(5.280)	(5.427)	(3.349)	(4.730)
Maximum temperature > 100 F	18.84	14.02	5.192	11.52
	(6.120)	(7.588)	(5.919)	(7.310)
Mean of dep. var.	245.5	255.2	276.8	223.3
	(187.3)	(190.2)	(227.7)	(199.4)
Pred. Δ deaths from std. dev. ↑ in availability index	-8.65	-11.05	-6.19	-19.44
Pred. Δ deaths from std. dev. ↑ in availability index	-5.60	-7.61	-20.88	-8.03
N	13871	10442	13547	10025
Demographic controls	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes
R-squared	0.446	0.453	0.329	0.267

Standard errors in parentheses

$p<0.1$, $p<0.05$, $p<0.01$

"Water availability index" refers to the measure of potable water availability described in equation 1. "Maximum temperature > 100 F" is a dummy that equals 1 when the maximum temperature exceeds 100° F during at least one day in a particular month in a particular province. "Pred. Δ deaths from std. dev. ↑ in availability index" is a linear prediction of the dependent variable, the average monthly number of deaths in a province, if the availability index increased by one standard deviation. "Low Water Access" refers to population subgroups in which the proportion of individuals with access to infrastructural water sources is below the sample median. "Lower % Black" refers to population subgroups in which the proportion of Black individuals is lower than the sample median.

8.6 Infrastructural Availability and Natural Availability as Independent Regressors

Tables 9 through 12 use the two volumetric flow measures of water availability described in Section 2 as covariates directly and independently. The comparability of the qualitative findings eliminates the potential concern that the results in section 5 were a byproduct of the way the water availability index was constructed.

Table 9: Baseline Regression Results

Dep. var.	(1) Deaths	(2) Deaths	(3) Deaths
Sample	Full	Full	Full
Monthly pipe flow rate × Maximum temperature > 100 F	2.380 (3.216)	-7.670 (2.311)	-5.540 (2.316)
Monthly river flow rate × Maximum temperature > 100 F	0.138 (0.0323)	-0.0200 (0.0314)	-0.109 (0.0402)
Maximum temperature > 100 F	9.327 (6.407)	14.02 (4.935)	22.30 (5.064)
Mean of dep. var.	260.9 (208.8)	260.9 (208.8)	260.9 (208.8)
Pred. Δ deaths from 0.25 std. dev. ↑ in pipe flow	-0.53	-32.64	-23.98
Pred. Δ deaths from 0.25 std. dev. ↑ in river flow	+1.13	-3.10	-1.59
N	26050	26050	26050
Demographic controls	No	Yes	Yes
Month-year fixed effects	No	No	Yes
R-squared	0.015	0.244	0.350

Standard errors in parentheses

$p<0.1$, $p<0.05$, $p<0.01$

"Monthly pipe flow rate" refers to the average of volumetric flow measures in pipes leading to a sanitation facility in the month when, and within the province where, the death occurred. "Monthly river flow rate" refers to the average of volumetric flow measures in the downstream river component at the same site in the month when, and within the province where, the death occurred. "Maximum temperature > 100 F" is a dummy that equals 1 when the maximum temperature exceeds 100° F during at least one day in a particular month in a particular province. "Pred. Δ deaths from 0.25 std. dev. ↑ in pipe/river flow" is a linear prediction of the dependent variable, the average monthly number of deaths in a province, if the pipe/river flow increased by one-quarter of a standard deviation.

Table 10: Regression Results by Water Access and Quality

	(5)	(6)	(7)	(8)
Dep. var.	Deaths	Deaths	Deaths	Deaths
Sample	Low Water Access	High Water Access	Lower % Gastro	Higher % Gastro
Monthly pipe flow × Maximum temperature > 100 F	-7.232	-2.017	-3.885	-3.124
	(3.007)	(4.059)	(2.065)	(1.679)
Monthly river flow × Maximum temperature > 100 F	-0.821	0.0328	-0.120	-0.0198
	(0.217)	(0.0265)	(0.0285)	(0.0209)
Maximum temperature > 100 F	39.67	18.46	17.92	12.43
	(9.697)	(4.725)	(4.389)	(3.330)
Mean of dep. var.	298.9	221.5	206.4	309.8
	(215.8)	(193.4)	(161.6)	(232.8)
Pred. Δ deaths from std. dev. $\uparrow$ in pipe flow	-30.03	-51.22	-14.27	-12.08
Pred. Δ deaths from std. dev. $\uparrow$ in river flow	-1.87	+4.02	-2.06	+0.60
N	13964	13454	12969	14449
Demographic controls	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes
R-squared	0.476	0.265	0.283	0.291

Standard errors in parentheses

p<0.1, p<0.05, p<0.01

"Monthly pipe flow rate" refers to the average of volumetric flow measures in pipes leading to a sanitation facility in the month when, and within the province where, the death occurred. "Monthly river flow rate" refers to the average of volumetric flow measures in the downstream river component at the same site in the month when, and within the province where, the death occurred. "Maximum temperature > 100 F" is a dummy that equals 1 when the maximum temperature exceeds 100° F during at least one day in a particular month in a particular province. "Pred. Δ deaths from 0.25 std. dev. $\uparrow$ in pipe/river flow" is a linear prediction of the dependent variable, the average monthly number of deaths in a province, if the pipe/river flow increased by one-quarter of a standard deviation. "Low Water Access" refers to population subgroups in which the proportion of individuals with access to infrastructural water sources is below the sample median. "Lower % Gastro" refers to population subgroups and months during which the percentage of deaths that were attributed to infectious diarrhea or gastroenteritis was below the sample median.

Table 11: Regression Results for Selected Population Subgroups

	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Dep. var.	Deaths	Deaths	Deaths	Deaths	Deaths	Deaths	Deaths
Sample	Male	Female	Less Educated	More Educated	Age $<$ 30	Age 30-60	Age $>$ 60
Monthly pipe flow $\times$ Maximum temperature $>$ 100 F	-4.019	-7.022	-3.312	-8.902	-10.50	-1.689	-1.093
	(3.035)	(3.437)	(1.832)	(4.001)	(3.274)	(2.812)	(1.891)
Monthly river flow $\times$ Maximum temperature $>$ 100 F	-0.112	-0.106	-0.0569	-0.113	-0.0866	-0.120	-0.0587
	(0.0547)	(0.0583)	(0.0323)	(0.0649)	(0.0608)	(0.0537)	(0.0234)
Maximum temperature $>$ 100 F	17.85	27.02	16.49	26.88	28.89	20.42	10.61
	(6.668)	(7.516)	(4.076)	(7.629)	(8.683)	(6.884)	
Mean of dep. var.	268.6	253.3	194.0	330.3	261.7	321.1	193.3
	(221.1)	(195.4)	(167.4)	(224.1)	(242.3)	(188.0)	(147.7)
Pred. Δ deaths from 0.25 std. dev. $\uparrow$ in pipe flow	-18.21	-29.68	-13.38	-38.04	-44.75	-8.59	-3.72
Pred. Δ deaths from 0.25 std. dev. $\uparrow$ in river flow	-1.62	-1.54	-0.93	-1.62	-1.59	-1.67	+0.05
N	13709	13709	13936	13482	11304	8478	7636
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.320	0.395	0.225	0.497	0.516	0.522	0.451

Standard errors in parentheses

$p<0.1$, $p<0.05$, $p<0.01$

"Monthly pipe flow rate" refers to the average of volumetric flow measures in pipes leading to a sanitation facility in the month when, and within the province where, the death occurred. "Monthly river flow rate" refers to the average of volumetric flow measures in the downstream river component at the same site in the month when, and within the province where, the death occurred. "Maximum temperature > 100 F" is a dummy that equals 1 when the maximum temperature exceeds 100° F during at least one day in a particular month in a particular province. "Pred. Δ deaths from 0.25 std. dev. $\uparrow$ in pipe/river flow" is a linear prediction of the dependent variable, the average monthly number of deaths in a province, if the pipe/river flow increased by one-quarter of a standard deviation. "Less Educated" refers to population subgroups in which the proportion of individuals that have attained at least a high school diploma is below the sample median.

Table 12: Regression Results by Race and Water Access

	(16)	(17)	(18)	(19)
Dep. var.	Deaths	Deaths	Deaths	Deaths
Sample	Higher % Black	Higher % Black, Low Water Access	Lower % Black	Lower % Black, High Water Access
Monthly pipe flow × Maximum temperature > 100 F	-6.951	-9.014	-10.88	-7.194
	(2.577)	(3.249)	(4.039)	(5.838)
Monthly river flow × Maximum temperature > 100 F	-0.338	-0.852	-0.0869	0.0161
	(0.140)	(0.273)	(0.0239)	(0.0274)
Maximum temperature > 100 F	35.03	49.00	16.88	17.96
	(7.614)	(11.60)	(6.099)	(5.873)
Mean of dep. var.	245.5	255.2	276.8	223.3
	(187.3)	(190.2)	(227.7)	(199.4)
Pred. Δ deaths from 0.25 std. dev. $\uparrow$ in pipe flow	-27.11	-37.42	-46.86	-66.30
Pred. Δ deaths from 0.25 std. dev. $\uparrow$ in river flow	-3.64	-2.24	+0.16	+2.86
N	13871	10442	13547	10025
Demographic controls	Yes	Yes	Yes	Yes
Month-year fixed effects	Yes	Yes	Yes	Yes
R-squared	0.449	0.473	0.336	0.275

Standard errors in parentheses

$p<0.1$, $p<0.05$, $p<0.01$

"Monthly pipe flow rate" refers to the average of volumetric flow measures in pipes leading to a sanitation facility in the month when, and within the province where, the death occurred. "Monthly river flow rate" refers to the average of volumetric flow measures in the downstream river component at the same site in the month when, and within the province where, the death occurred. "Maximum temperature > 100 F" is a dummy that equals 1 when the maximum temperature exceeds 100° F during at least one day in a particular month in a particular province. "Pred. Δ deaths from 0.25 std. dev. $\uparrow$ in pipe/river flow" is a linear prediction of the dependent variable, the average monthly number of deaths in a province, if the pipe/river flow increased by one-quarter of a standard deviation. "Low Water Access" refers to population subgroups in which the proportion of individuals with access to infrastructural water sources is below the sample median. "Lower % Black" refers to population subgroups in which the proportion of Black individuals is lower than the sample median.

8.7 Robustness Check: Results Without Imputation for Missing Household Survey Years

As discussed in section 4, there is no publicly available General Household Survey data from Statistics South Africa for years before 2002, nor the year 2004. The results reported in the main paper use averages from 2002 for these missing years for each control variable from the household survey. To eliminate the concern that this imputation affected the results, I present replications of Tables 2, 3, and 4 from the main paper excluding the years with missing survey data. Nearly all qualitative findings are the same, and some effects are estimated to be even stronger. This includes the number of deaths prevented by an extensive margin increase in infrastructural water availability, which more than triples the prediction reported in the main paper.

Table 13: Regression Results with Triple Interactions, No Imputation

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	Deaths	Deaths	Deaths	Deaths	Deaths	Deaths
Sample	Full	Full	Full	Full	Full	Full
Water availability index ×	27.81	-12.61	-8.761	-6.353	-15.81	-11.98
Maximum temperature > 100 F	(5.564)	(4.873)	(4.660)	(4.552)	(5.976)	(5.397)
× Higher water access				-13.94		
				(21.61)		
× Higher gastroenteritis rate					9.733	
					(7.266)	
× Higher % Black						3.283
						(9.217)
Pipe flow on × Maximum	-9.013	-13.35	-6.858	-15.01	-4.768	-15.27
temperature > 100 F	(3.803)	(3.486)	(3.006)	(3.527)	(3.755)	(3.817)
× Higher water access				-6.953		
				(6.230)		
× Higher gastroenteritis rate					-5.359	
					(3.700)	
× Higher % Black						9.109
						(5.490)
Maximum temperature > 100 F	24.52	16.04	17.67	26.05	19.27	19.10
	(6.939)	(4.943)	(4.442)	(6.157)	(4.931)	(7.185)
N	25474	25474	25474	25474	25474	25474
Mean of dep. var.	260.9	260.9	260.9	260.9	260.9	260.9
Demographic controls	No	Yes	Yes	Yes	Yes	Yes
Month-year fixed effects	No	No	Yes	Yes	Yes	Yes
R-squared	0.017	0.266	0.362	0.364	0.388	0.365

Standard errors in parentheses

"Water availability index" refers to the measure of potable water availability described in equation 1. "Maximum temperature > 100 F" is a dummy that equals 1 when the maximum temperature exceeds 100° F during at least one day in a particular month in a particular province. "Pred. Δ deaths from std. dev. ↑ in availability index" is a linear prediction of the dependent variable, the average monthly number of deaths in a province, if the availability index increased by one standard deviation. "Pipe flow on" is a dummy that equals 1 when the pipe flow measure described in Section 2 is above the sample 25th percentile (i.e., significantly greater than zero). "Pred. Δ deaths from pipe flow off → on" is the difference between the predicted number of deaths when pipe flow is "on" and when it is "off."

Table 14: Estimated Overall Treatment Effects on Mortality

	Predicted Change in Number of Deaths per Month	P-value
Pipe flow off → on	-10.89	<0.01
One standard deviation increase in water availability index, pipe flow on	-4.04	0.05
Cumulative	-14.93	<0.01
Mean deaths per month	260.9	

Table 15: Heterogeneity of Estimated Treatment Effects on Mortality

		Predicted Change in Number of Deaths per Month		
		Below Median	Above Median	Difference p-value
Pipe flow off → on				
	Water access	-9.80 (-3.4%)	-18.15 (-11.1%)	<0.01
	Gastroenteritis rate	-9.18 (-4.3%)	-13.37 (-4.7%)	0.80
	% Black	-23.20 (-6.8%)	-4.98 (-2.2%)	<0.01
One standard deviation increase in water availability index, pipe flow on				
	Water access	-3.83 (-1.6%)	-6.66 (-4.8%)	0.57
	Gastroenteritis rate	-7.47 (-4.1%)	-1.43 (-0.5%)	<0.01
	% Black	-4.82 (-1.7%)	-4.53 (-2.4%)	0.62
Cumulative				
	Water access	-13.63 (-5.0%)	-24.81 (-15.9%)	0.06
	Gastroenteritis rate	-16.65 (-8.5%)	-14.80 (-5.3%)	0.11
	% Black	-28.02 (-8.6%)	-9.51 (-4.7%)	0.04