INSTITUTE for REPLICATION

No. 90
I4R DISCUSSION PAPER SERIES

Mortality, Temperature, and Public Health provision: A Comment on Cohen and Dechezlepretre (2022)

Grant Benjamin

Ben Couillard

Jonathan D. Hall

December 2023

I4R DISCUSSION PAPER SERIES

I4R DP No. 90

Mortality, Temperature, and Public Health Provision: A Comment on Cohen and Dechezlepretre (2022)

Grant Benjamin[1], Ben Couillard[1], Jonathan D. Hall[2]

[1]*University of Toronto/Canada*
[2]*University of Alabama, Tuscaloosa/USA*

DECEMBER 2023

ISSN: 2752-1931

Mortality, temperature, and public health provision: A comment on Cohen and Dechezlepretre (2022)*

Grant Benjamin, Ben Couillard, and Jonathan D. Hall

November 29, 2023

Abstract

Cohen and Dechezleprêtre (2022) investigate the heterogeneous impact of temperature on mortality across Mexico, and how affordable healthcare services that target the low-income population attenuate the mortality effects of weather events. They find that while extreme temperatures are more dangerous than less extreme temperatures, the increased frequency of non-extreme temperatures mean these temperatures cause more deaths. First, we reproduce the paper's main findings, uncovering a minor coding error that has a trivial effect on the main results. Second, we test the robustness of the results to clustering at the state level, omitting precipitation, and using a different weighting scheme. The original results are robust to all of these changes.

KEYWORDS: Mortality, Weather Events, Affordable Health Care, Poverty

JEL CODES: I12, I13, I14, O13, O14, Q54.

*Authors: Benjamin: University of Toronto. E-mail: grant.benjamin@mail.utoronto.ca. Couillard: University of Toronto. E-mail: ben.couillard@mail.utoronto.ca. Hall: University of Alabama. E-mail: jonathan.hall@ua.edu. We have no conflicts of interest. Hall is the corresponding author.

1 Introduction

Cohen and Dechezleprêtre (2022) investigate the heterogeneous impact of temperature shocks on mortality across Mexico, and how affordable healthcare services that target the low-income population attenuate the mortal effects of weather events. Using an administrative dataset of over 14 million daily mortality rates from 1998 to 2017 in 2,297 municipalities, Cohen and Dechezleprêtre (2022) combines temperature data into 7 bins to estimate how each temperature range contributes to a municipality's mortality rate. The authors estimate a distributed lag model, including 30 lags of each temperature bin, to account for deaths that occur later but result from the weather event. The causal evidence from the regressions suggests that random variation in temperature is responsible for 3.8% of the annual deaths in Mexico (26,000 per year). Although the media and previous studies focus on extreme weather events, 71% of weather-related deaths stem from mildly cold days when the average temperature is between 12°C and 20°C. At the extremes, unusually cold days (<12°C average temperature) trigger 5,705 deaths each year (s.e. 591) and extremely hot days (>32°C) kill around 539 people annually (s.e. 75). Finally, Cohen and Dechezleprêtre (2022) determine vulnerability to weather shocks is strongly correlated with income; death following cold and mildly cold days below 24°C is more than twice as frequent for people living below the national median personal income. Hence, a large majority of cold-related deaths affect the poorest income groups.

Cohen and Dechezleprêtre (2022) next estimate how Seguro Popular, a universal health care policy introduced in 2004 to provide health care services to the low-income population. The main results suggest that the policy saves at least 1,600 lives (s.e. 641) annually from mildly cold weather since 2004, or 6.6% of weather-induced deaths. Understanding how weather shocks impact human health across the income distribution in developing countries and quantifying how public policies may alleviate the impact of weather events is, therefore, of great policy importance.

In the present paper, we investigate whether Cohen and Dechezleprêtre (2022)'s main results are reproducible and perform additional robustness checks. We acknowledge that the original study was successfully reproduced by the data editor's team at the American Economic Association. We also successfully reproduce Cohen and Dechezleprêtre (2022)'s main results (Table 2 and Figure 2) using their code. We uncover one minor coding error where the authors accidentally exclude data from February 29$^{\text{th}}$ in leap years. Correcting the coding error does not significantly alter any of the main results in the analysis.

We expand the initial analysis with multiple robustness tests to evaluate the sensitivity of the paper's main results. Specifically, we (1) cluster standard errors at the state-level as opposed to the municipality-level clustered errors in the original paper, (2) omit precipitation from the main specification, and (3) estimate the unweighted regression coefficients for all specifications to evaluate the importance of the authors' population weights in their main specification.

Our sensitivity analysis shows that when we include state-clustered standard errors, the standard error estimates for all temperature bins significantly increase for the low-temperature bins, while the high-temperature bins remain similar. Despite the increase across the low-temperature bins, the estimates remain statistically significant after controlling for this state-level variation. When we omit precipitation from the main specification, the coefficient estimates are marginally smaller for all temperature bins. Although the magnitudes for all temperature bins are lower, the signs and significance remain the same, so the original conclusions are unchanged: both mildly cold days and extreme weather events are essential to understanding weather's role in mortality. Finally, our unweighted regression model indicates that Cohen and Dechezleprêtre (2022)'s narrative does not depend on using population weights as the lower temperature bin coefficients do not statistically change. One notable change from the unweighted regressions is that the 28–32 °C temperature bin is no longer statistically different from the baseline temperature bin.

2 Reproducibility

For the replication, we reproduce Figure 2 and Table 2 from Cohen and Dechezleprêtre (2022) that estimates the average deaths per year across Mexico using the fixed effects specification defined in Equation 1 below. Cohen and Dechezleprêtre (2022) use temperature bins to estimate how both extreme weather events and mild changes in temperature affect mortality. The 7 temperature bins are defined as follows: <12 °C, 12–16 °C, 16–20 °C, 20–24 °C, 24–28 °C, 28–32 °C, and >32 °C. For all specifications, 24–28 °C is the omitted category.[1]

$$Y_{i,d,m,t} = \sum_{k=0}^{K=30} \sum_{s} \theta_{s,-k} \cdot B_{s,i,d-k,m,t} + \sigma \cdot P_{i,d,m,t} + \mu_{i,d,m} + \mu_{i,t} + \mu_{d,m,t} + \epsilon_{i,d,m,t} \quad (1)$$

The dependent variable in the above specification captures the number of annual deaths in municipality i, on day d of month m in year t. $\mu_{i,d,m}$,$\mu_{i,t}$, and $\mu_{d,m,t}$ are various fixed effects to control for unobserved heterogeneity across time and municipalities. The specification also includes a distributed lag of temperature bins for 30-days to determine if weather events in previous days affect deaths on the current day. The coefficients of interest are $\theta_{s,-k}$, which captures how day $-k$'s temperature affects today's mortality for each temperature bin s. Finally, σ provides insights on the effect of weather events as it estimates how precipitation, $P_{i,d,m,t}$, affects mortality.

We uncover one minor error in the replication code. In the data construction, the authors impute dates for all missing values for all observations and intentionally create February 29$^{\text{th}}$ observation in all years. Although the authors intend to drop all February 29$^{\text{th}}$ observations that occur during non-leap years, a coding error dropped February 29$^{\text{th}}$ during leap years instead. The end result of this coding error is only that the data from true February 29$^{\text{th}}$s are not used, as the February 29$^{\text{th}}$ observations from non-leap years are not used in the regressions since no weather data exists for these observations. We re-run the authors' code to correct the coding error and reproduce the authors' results in Column (1) of Table 2. We determine

[1] There is a minor typo in the published version of Table 2, as the final row ought to read > 32°C rather than >30 °C

that the point estimates do not significantly change, with the sign, magnitude and statistical significance being remarkably similar.

3 Replication

3.1 Sensitivity Analysis

For our sensitivity analysis, we first replicate the main results from Cohen and Dechezleprêtre (2022): Table 2 and Figure 2. Column (1) of our Table 1 presents the exact replication of their Table 2 while our Figure 1 presents the accompanying graph that replicates their Figure 2. For the remainder of this robustness replication, we adapt the base specification as Cohen and Dechezleprêtre (2022) in various manners. Specifically, we (1) cluster standard errors at the state level as opposed to the municipality level, (2) omit precipitation from the main specification, and (3) estimate the unweighted regression coefficients for all specifications to evaluate the importance of the authors' population weight. The specification changes are shown in Appendix I. We decide on these robustness checks after reading Cohen and Dechezleprêtre (2022), but prior to observing the original code and programs.

3.2 State-Clustered Standard errors

For our first adaptation, we investigate whether the results in Cohen and Dechezleprêtre (2022) are robust to changes in the standard error definition. We cluster our standard errors at the state level rather than the municipality level because mortality for municipalities within a state may be correlated because of state policies, demographic similarities, income, or other idiosyncratic components the main specification does not capture. Column (2) in Table 1 and Figure 2 show the results with state-clustered standard errors. With state-level clustered standard errors, the point estimates are identical to Figure 1 and Column (1) of Table 1 because we only change the level of clustering; the only changes occur in the standard error estimates. However, the confidence intervals for our estimates using the new level of clustering become significantly larger for all temperature bins compared to the baseline model. The decrease in precision in the estimates suggests there is an unobserved correlation across municipalities within the same state that the authors do not account for. However, even with the lesser statistical precision with the new estimates, the statistical significance remains consistent for all point estimates as all temperature bins except for the 20–24 °C bin are statistically significant; the 20–24 °C bin is not statistically different from 0 in both specifications. Therefore, Cohen and Dechezleprêtre (2022)'s main results are robust in both significance and magnitude to the change in clustering level.

3.3 Omitting Precipitation

For our next robustness check, we analyze the interdependency between precipitation and temperature and their effects on mortality. We omit the precipitation variable from the main regression specification and estimate how the causal effect of each temperature bin changes without this control. Column (3) of Table 2 and Figure 3 present the coefficient estimates from the modified regression, with standard errors clustered at the municipality level as in the original specification. After we exclude the precipitation variable, the point estimates for almost all temperature bins are lower in magnitude. The lone exception is the >32 °C temperature bin, which is marginally greater relative to the replication results. The downward bias in the regression estimates that omit precipitation suggests that high levels of precipitation increase mortality; the total mortality each temperature bin causes understates the true mortal effects of temperature deviations.[2] Despite this induced bias, the estimates in mortality changes do not statistically differ for any of the individual temperature bins nor the total estimate of deaths from the original specification.

[2]If we assume temperature and precipitation are negatively correlated in Mexico.

3.4 Population Weights

For our final modification, we place equal weight on all municipalities rather than weighting by population. When the authors construct their final dataset, they impute missing population values using each municipality's latest population trend. For instance, if the population value of a municipality is missing in 2007 but the population increased 3% from 2005 to 2006, the authors assume it will also grow 3% from 2006 to 2007. However, this imputation procedure assigns negative populations to municipalities when the latest trend is a significant decrease. This negative population estimate uses non-positive regression weights in the original regression model. To avoid this non-positive population weight problem, we estimate the main specification without population weights and report the results in Column (4) of Table 1 and Figure 4, with standard errors clustered at the municipality level. Although the results in Column (4) do not significantly differ from the main specification, three notable changes occur. First, the 28–32 °C bin is no longer statistically different from the baseline bin of 24–28 °C. Second, the point estimate for the 20–24 °C bin becomes positive, but remains statistically insignificant. Finally, all standard error estimates are significantly larger. Accordingly, Figure 4 maintains the u-shaped trend relative to the baseline case of 20–24 °C, with increasing death rates as temperatures deviate towards either end of the temperature range, with a larger confidence interval for all temperature bins. Although the standard errors are larger without population weights, the results do not statistically differ from the main specification.

3.5 Leap-Year Correction

As discussed in Section 2, we uncover one minor error in the replication code where the authors inadvertently dropped February 29$^{\text{th}}$ in leap years. We re-run the authors' code to correct the coding error and reproduce the authors' results in Column (1) of Table 2. After this correction, we determine that the point estimates do not significantly differ from the main results, with the same sign, and the magnitude and statistical significance remain remarkably similar. We extend our robustness replication and investigate how the robustness checks change after correcting the leap year error. Columns (2)–(4) of Table 2 present the results for the three specification changes using the corrected data. For all columns and all temperature bins, the results do not significantly differ from the original estimates. From this, we conclude that the error in the code does not alter the conclusions of Cohen and Dechezleprêtre (2022).

4 Conclusion

From our sensitivity analysis, we determine that the results from Cohen and Dechezleprêtre (2022) are robust to meaningful changes in the regression specification and do not significantly change with state-clustered standard errors, omitting precipitation, or omitting the regression weights. Further robustness tests that could be conducted in the future include investigating how the unweighted results interact with the low-income population, including distributed lags of precipitation, or using changes in temperature as treatment as opposed to temperature levels.

References

Cohen, F. and Dechezleprêtre, A.: 2022, Mortality, temperature, and public health provision: Evidence from Mexico, *American Economic Journal: Economic Policy* **14**(2), 161–192.

5 Figures

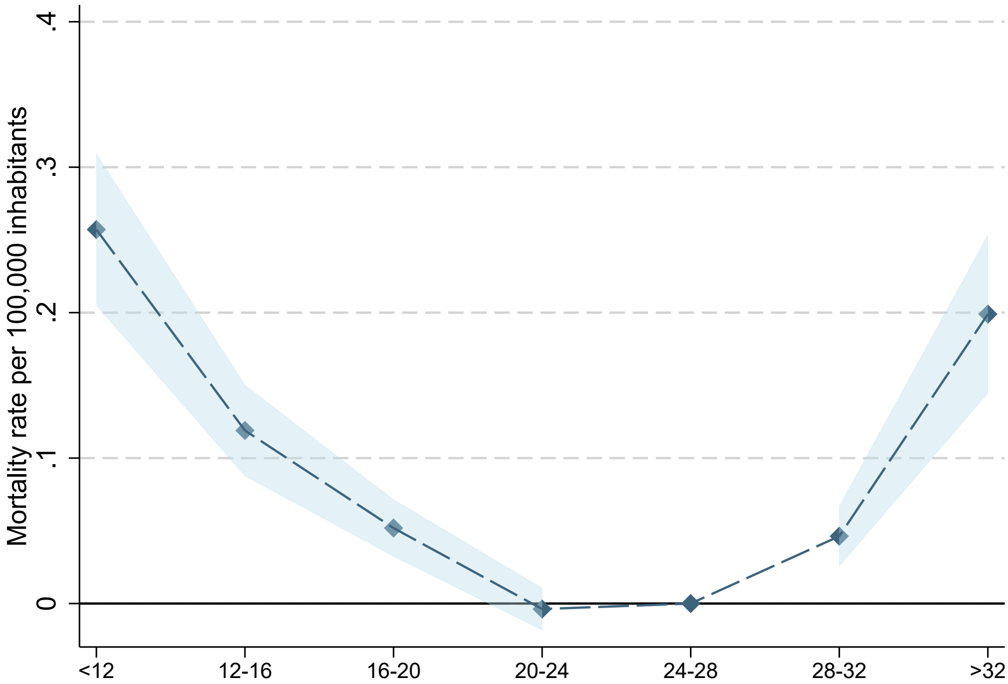

Figure 1: Reproduced Figure 2: Impact of temperature bins on 31-day cumulative mortality, in deaths per 100,000 inhabitants

Notes: This graph reproduces Figure 2 from Cohen and Dechezleprêtre (2022). It shows "the cumulative effect of a day with a temperature within each bin (relative to the 24 °C–28 °C category) obtained from a dynamic model with 30 lags. The diamonds show the sum of the coefficients on these thirty lags in each category. The shaded area corresponds to the 95 percent confidence interval (clustered at the municipality level). The dependent variable is the daily mortality rate at the municipality level."

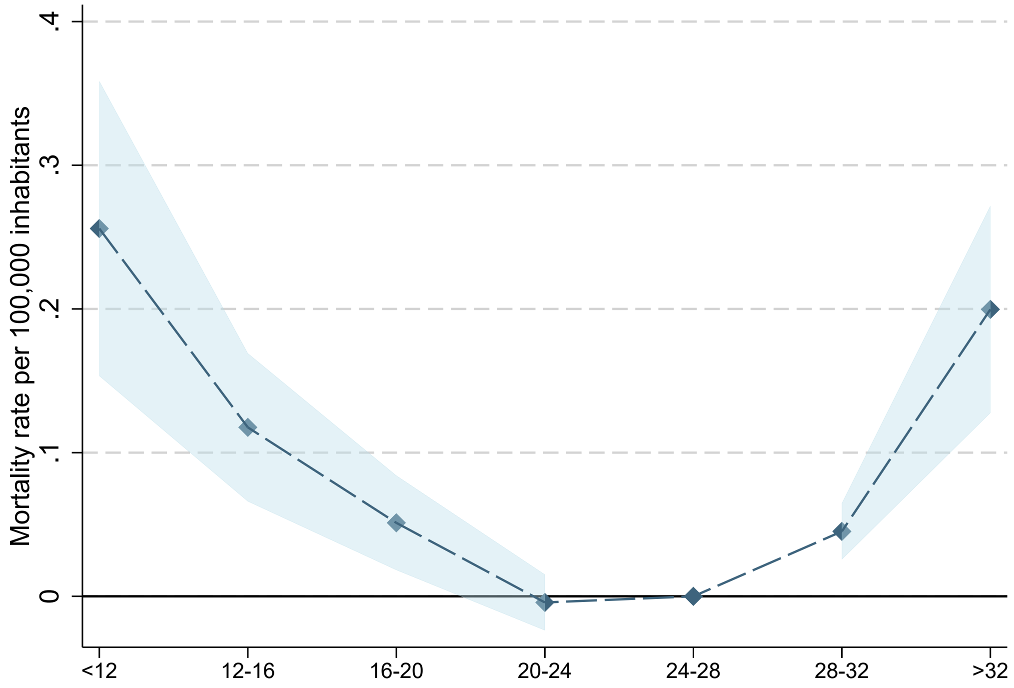

Figure 2: Impact of temperature bins on 31-day cumulative mortality: Clustering at the state level

Notes: This graph reproduces Figure 2 from Cohen and Dechezleprêtre (2022), when the clustering is done at the state, rather than municipality, level. See the notes to Figure 1 for more detail.

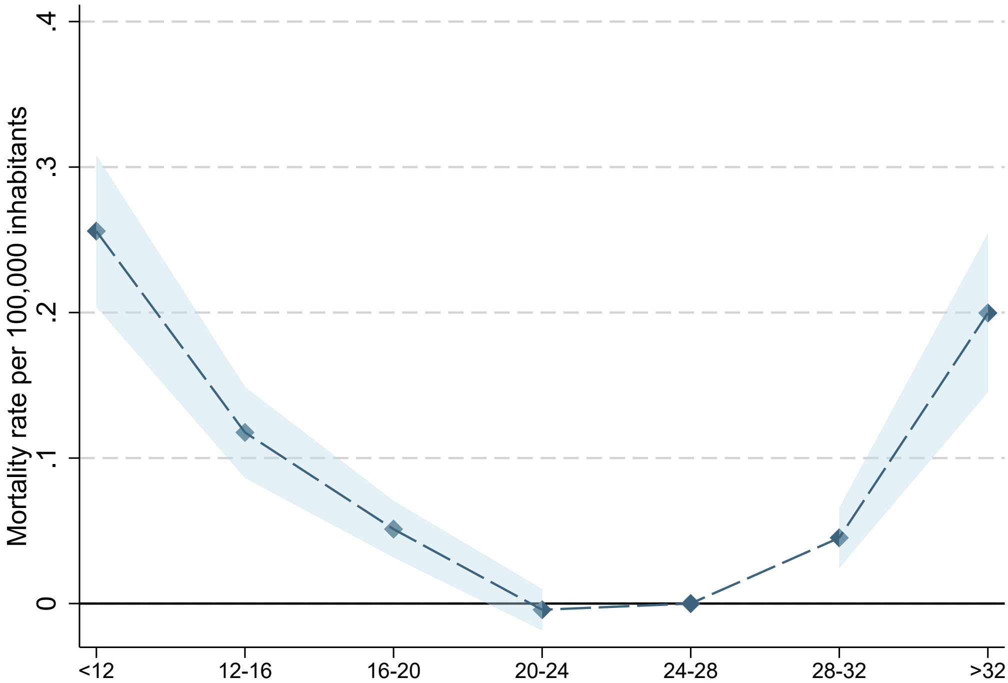

Figure 3: Impact of temperature bins on 31-day cumulative mortality: Not controlling for precipitation

Notes: This graph reproduces Figure 2 from Cohen and Dechezleprêtre (2022), without controlling for precipitation. See the notes to Figure 1 for more detail.

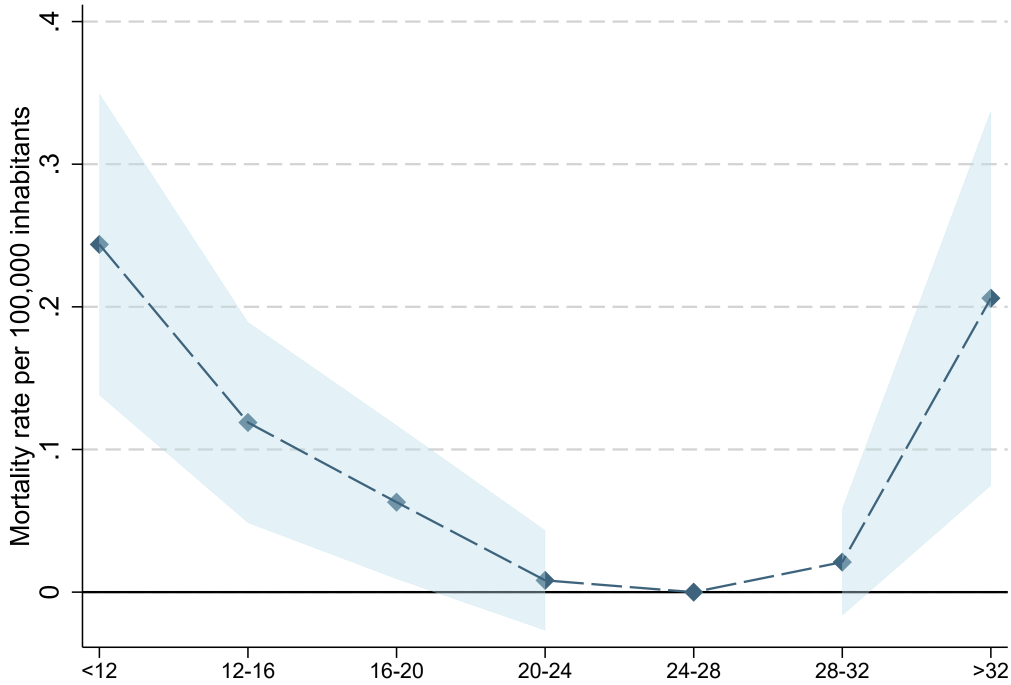

Figure 4: Impact of temperature bins on 31-day cumulative mortality: Weighting municipalities equally

Notes: This graph reproduces Figure 2 from Cohen and Dechezleprêtre (2022), without weighting municipalities by their population. See the notes to Figure 1 for more detail.

6 Tables

Table 1: Cohen and Dechezlepretre (2022) Replication

Specification	(1) Replication	(2) State Clustered Errors	(3) Omit Precipitation	(4) Equal Weights
<12 °C	5,705***	5,705***	5,680***	5,409***
	(591)	(1,168)	(587)	(1,196)
	p < 0.001	p < 0.001	p < 0.001	p < 0.001
12–16 °C	11,142***	11,142***	11,015***	11,139***
	(1,494)	(2,473)	(1,488)	(3,360)
	p < 0.001	p < 0.001	p < 0.001	0.001
16–20 °C	7,536***	7,536***	7,426***	9,155**
	(1,449)	(2,440)	(1,430)	(3,984)
	p < 0.001	0.004	p < 0.001	0.022
20–24 °C	-367	-367	-413	803
	(718)	(981)	(702)	(1,738)
	0.609	0.711	0.556	0.644
24–28 °C	—	—	—	—
28–32 °C	1,769***	1,769***	1,726***	804
	(403)	(366)	(399)	(728)
	p < 0.001	p < 0.001	p < 0.001	0.269
>32 °C	539***	539***	541***	558***
	(75)	(97)	(75)	(182)
	p < 0.001	p < 0.001	p < 0.001	0.002
Total	26,324***	26,324***	25,973***	27,869***
	(3,609)	(6,214)	(3,573)	(8,886)
	p < 0.001	p < 0.001	p < 0.001	0.002

Notes: p-values are presented below the standard errors, which are in parentheses. Significant at the ***[1%] **[5%] *[10%] level.

Table 2: Robustness Checks: Corrected Leap Year

Specification	(1) Replication	(2) State Clustered Errors	(3) Omit Precipitation	(4) Equal Weights
<12 °C	5,686***	5,686***	5,652***	5,439***
	(588)	(1,169)	(585)	(1,195)
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$
12–16 °C	11,089***	11,089***	10,948***	11,223***
	(1,490)	(2,491)	(1,484)	(3,352)
	$p < 0.001$	$p < 0.001$	$p < 0.001$	0.001
16–20°C	7,494***	7,494***	7,380***	9,318**
	(1,449)	(2,456)	(1,429)	(3,980)
	$p < 0.001$	0.005	$p < 0.001$	0.019
20–24 °C	-388	-388	-437	815
	(717)	(983)	(701)	(1,737)
	0.588	0.696	0.533	0.639
24–28°C	—	—	—	—
28–32°C	1,769***	1,769***	1,726***	808
	(402)	(364)	(399)	(728)
	$p < 0.001$	$p < 0.001$	$p < 0.001$	0.267
>32°C	540***	540***	542***	553***
	(75)	(97)	(75)	(182)
	$p < 0.001$	$p < 0.001$	$p < 0.001$	0.002
Total	26,190***	26,190***	25,811***	28,155***
	(3,604)	(6,263)	(3,568)	(8,871)
	$p < 0.001$	$p < 0.001$	$p < 0.001$	0.002

Notes: p-values are presented below the standard errors, which are in parentheses. Significant at the ***[1%] **[5%] *[10%] level.

7 Appendix

7.1 Specification Changes

Cohen and Dechezleprêtre (2022)'s main regression specification is:

$$Y_{i,d,m,t} = \sum_{k=0}^{K=30} \sum_{s} \theta_{s,-k} \cdot B_{s,i,d-k,m,t} + \sigma \cdot P_{i,d,m,t} + \mu_{i,d,m} + \mu_{i,t} + \mu_{d,m,t} + \epsilon_{i,d,m,t}, \quad (2)$$

where the dependent variable is the number of annual deaths in municipality i, on day d of month m in year t. $B_{s,i,d-k,m,t}$ is an indicator variable if the temperature on day $d-k$ is in temperature bin s. Finally, $P_{i,d,m,t}$ represents the total level of precipitation and $\mu_{i,d,m}$, $\mu_{i,t}$, and $\mu_{d,m,t}$ are fixed effects to control for unobserved heterogeneity. The specification change where we omit precipitation is below:

Omitting Precipitation

$$Y_{i,d,m,t} = \sum_{k=0}^{K=30} \sum_{s} \theta_{s,-k} \cdot B_{s,i,d-k,m,t} + \mu_{i,d,m} + \mu_{i,t} + \mu_{d,m,t} + \epsilon_{i,d,m,t} \quad (3)$$