Adapting to High Heat, Failing at the Extreme: Newborn Heat Mortality over Four Decades in Bangladesh*

Mohammad Islamul Haque[†1] and Anoushka Chopra[‡2]

[1]Yale University

[2]Department of Economics, Massachusetts Institute of Technology

October 13, 2026

Abstract

Damage functions in climate economics typically assume that rising income and cooling technology steadily reduce the mortality cost of extreme heat. We test this with a birth-cohort decomposition of neonatal heat mortality for 276,031 Bangladeshi births from nine Demographic and Health Survey rounds merged with ERA5 daily-maximum temperature. At 30–40°C, only the 1980 cohort shows a significant response, about 4.5 additional neonatal deaths per 1,000 live births per 10 additional days of exposure. Each later cohort's response is statistically undetectable, close to zero and significantly below the 1980 response: the signature of adaptation. A heterogeneity-robust difference-in-differences estimator links this flattening to the staggered rollout of the Rural Electrification Board's program, implying roughly 718,000 neonatal deaths averted over the 1990s–2010s. Above 40°C the effect is large and statistically significant for the 2010 cohort, about 43 additional deaths per 1,000 births per 10 additional

*We thank Matthew Kahn and Mushfiq Mobarak for valuable feedback, and seminar participants at Yale for helpful comments. All remaining errors are our own.

[†]Corresponding author. Postdoctoral Associate, Yale Research Initiative on Innovation and Scale (Y-RISE), Yale University. islamul.haque@yale.edu.

[‡]Predoctoral Researcher, Department of Economics, Massachusetts Institute of Technology. anous022@mit.edu.

days, and no test detects a decline from earlier cohorts, whose tail estimates are imprecise. We lack direct evidence on why no adaptation is detected above 40°C; cooling-technology penetration is one candidate: in 2022, 96 % of households owned an electric fan, while air conditioning, which cools even in extreme heat, was found in only 1.4 %. Realized exposure implies about 5,340 such deaths over the 2010s. With bias-corrected projections from 13 climate models, the tail effect implies 657 to 1,920 excess neonatal deaths per year by mid-century, with air-conditioning penetration moving the projection, averaged across scenarios, about two and a half times as much as the emissions pathway.

Keywords: extreme heat; neonatal mortality; climate adaptation; rural electrification; temperature extremes; Bangladesh

JEL codes: Q54; I15; J13; O13; H54

1 Introduction

Heat-mortality elasticities have fallen sharply in the United States as residential air conditioning has diffused (Barreca et al., 2016). Air conditioning (AC) is largely absent in low-income tropical settings, but it is not the only route to adaptation: electric fans, expanded postnatal and prenatal health services, and traditional behavioral and housing responses to heat may each blunt the mortality cost of hot days even where AC penetration is negligible. Whether these lower-cost margins can flatten the heat-mortality relationship for newborns in a poor tropical population is an open empirical question with large stakes. This paper asks how far heat-related neonatal mortality can fall in a low-income, low-AC tropical population facing rising heat, and where in the temperature distribution any reductions have occurred—which matters for how climate damages are priced.

Bangladesh offers empirical leverage on the question. Daily-maximum heat exposure has risen steadily over our study period, though, as Section 2.3 documents, the increase is concentrated in the 30–40°C range rather than at the physiological extreme, which remains rare and regionally concentrated. Air-conditioning penetration sits at 1.4 % of households, while electric fans, present in 96 %, are the dominant cooling technology, according to the 2022 round of the Bangladesh Demographic and Health Survey (BDHS).

We estimate a cohort decomposition of neonatal heat mortality on 276,031 live births from nine rounds of the BDHS (fielded 1993–94 to 2022), merged with inverse-distance-weighted

daily-maximum temperature from the ERA5 reanalysis of the European Centre for Medium-Range Weather Forecasts. Heat exposure is the count of days in each of six 5°C bins during the postnatal at-risk window (28 days for neonatal mortality, 365 for infant, 1,825 for under-five), with 25–30°C as the reference bin, and we let the heat-mortality slope vary by birth decade (1980, 1990, 2000, 2010). The baseline specification accounts for district × birth calendar-month and birth-year fixed effects together with birth-order group dummies and child, household, and maternal controls, and additionally absorbs 63 district-specific linear birth-year trends and full-gestation (in-utero) heat exposure. Identification comes from year-to-year deviations in the temperature distribution within a district-month cell—heat that is unusual for that district in that calendar month—which we assume are orthogonal to other determinants of newborn survival given the fixed effects, district-specific trends, and controls. Standard errors are clustered at the district level.

High-heat adaptation. A positive high-heat (30–40°C) neonatal mortality response appears only in the earliest (1980) birth cohort, where it is significant at the 5 % level ($\widehat{\beta}_{1980}^{30\text{–}40} = +0.0045$, $p = 0.015$)—about 4.5 additional neonatal deaths per 1,000 live births per +10 days of exposure, against a 1980-cohort base rate of 73 per 1,000. Adaptation means that the 1980 slope exceeds the later ones, so against each later cohort c we test $H_1 : \beta_{1980} > \beta_c$ one-sided, and every test rejects at the 5 % level ($p = 0.012$, 0.027 and 0.013 for the 1990, 2000 and 2010 cohorts). The later cohorts sit about 0.0026–0.0036 per +10 days lower, so the 1980 slope of +0.0045 falls to essentially zero by the 1990 cohort and stays there (Section 5). This comparison holds the weather shock fixed—year-to-year heat within a district and calendar month—and asks only whether a given shock kills fewer newborns in later cohorts. Everything else that changed over three decades, from mortality levels and the composition of births to how each survey was fielded, is absorbed by birth-year fixed effects and district-specific trends, and the slopes are essentially unchanged with survey-round fixed effects, each round being a nationally representative sample. A decline in the slope under these controls is what adaptation means: the weakening of the link between heat and newborn survival as society and infrastructure change, which is the object that damage functions need. If the early data were noisier, attenuation would bias the 1980 slope toward zero, working against this finding.

We probe what drove this adaptation using the staggered rollout of Bangladesh's Rural Electrification Board (REB) program across districts (connection years 1980–2001, dated by the first energization of the rural electric cooperative serving each district). The high-heat neonatal slope flattens at connection by 0.0039 per +10 days under conventional two-

way fixed effects (TWFE; $p = 0.004$) and by 0.0058 ($p < 0.001$) under a cohort-pooled interaction-weighted (IW) estimator, in the spirit of Sun and Abraham (2021), that avoids comparisons with already-connected districts. The flattening holds under six definitions of the connection dates (-0.0050 to -0.0058) and when the heat slopes may change across birth decades for other reasons (-0.0048). Births in the large cities, which had grid power before the program, show no flattening at connection, while rural births and births in towns, many inside the program's service areas, do. Neither the event-study pre-period coefficients nor the pre-connection test of the imputation estimator of Borusyak et al. (2024), reported as a robustness check in the Online Appendix ($p = 0.183$), is significant. We find no evidence that connection coincided with a rise in rural wealth or maternal education, as it would if rollout order proxied a broad "development arrived" bundle.

The stakes are large. Applied to realized exposure over the 1990s–2010s, the cohort-pooled IW estimate (-0.0058, which uses all birth years and so matches the accounting horizon) implies about 718,000 neonatal deaths averted by the REB rollout (95 % confidence interval (CI) roughly 323,000 to 1.11 million), or Int$143.5 billion in 2015 international dollars at a central value of a statistical life (VSL) of Int$200,000 for Bangladesh, obtained with the income-elasticity transfer that Robinson et al. (2019) prefer (Section 8).

No detectable adaptation at the extreme tail. We detect no such adaptation at the extreme tail of the distribution. In the pooled (all-cohort) level regression without a recall cap, the ≥ 40°C coefficient is positive and significant ($+0.0187$ per $+10$ days, $p = 0.013$), and for the most recent (2010) birth cohort it is large, positive, and statistically significant ($\widehat{\beta}^{\geq 40}_{2010} = +0.0432$, $p = 0.009$): ten additional days of extreme-heat exposure in the 28-day neonatal window are associated with about 43 additional neonatal deaths per 1,000 live births, against a 2010-cohort base rate of 30 per 1,000. No one-sided test for a decline from an earlier cohort to the 2010 cohort comes close to rejecting (smallest $p = 0.489$), although the earlier cohorts' tail estimates are too imprecise to rule one out. The 2010-cohort tail slope becomes large between 38 and 40°C (Section 5.4). The same cohort pattern appears on heat-wave days, but we do not attribute the tail effect to heat waves specifically (Section 5.2). At the tail the REB estimates are very imprecise (cohort-pooled IW estimate $+0.0107$, standard error (SE) 0.0259); because every program district was connected by 2001, the large 2010-cohort tail slope shows that electrification did not eliminate extreme-heat mortality, though not whether it reduced it from a higher level.

Current and projected mortality costs. The realized toll is modest today. Applied to realized 2010s exposure against a zero-effect counterfactual, the 2010 effect implies roughly 5,340 neonatal deaths attributable to extreme heat over the decade (∼534/year, about 0.6 % of the ∼90,000 annual neonatal deaths implied by 3 million annual live births, United Nations, 2022, at the 2010-cohort base rate of 30 per 1,000; Int$1.07 billion at the same VSL), or roughly a third fewer if part of the effect is displacement from later infancy (Section 6). The tail effect's importance is prospective. Combined with mid-century ≥ 40°C day counts that we compute from the bias-corrected Coupled Model Intercomparison Project Phase 6 (CMIP6) projections of Mishra et al. (2020), the same tail slope implies on the order of 657 to 1,920 excess neonatal deaths a year by mid-century (2041–2060), depending on air-conditioning penetration and the emissions pathway. This is a range across scenarios rather than a confidence interval, and, averaged across scenarios, air-conditioning penetration moves it about two and a half times as much as the emissions pathway does (Section 8).

Robustness. Section 6 and the Online Appendix test each result against the specific concerns it raises; Table 5 collects the estimates. We summarize the main checks here, stating in each case the concern, the test, and the outcome.

Who is observed. Because births are reported retrospectively, by surviving, ever-married mothers, through rounds that differ by cohort (Section 3.4), the cohorts might differ in how they were sampled rather than in how they respond to heat. (The ever-married frame is the BDHS's own design and excludes essentially no births: marriage is nearly universal in Bangladesh and childbearing outside it rare, Section 3.) We add survey-round fixed effects, control for the mother's age at birth, and weight by the Demographic and Health Surveys (DHS) sampling weights. The 1980 high-heat slope is +0.0044 to +0.0046 in each case, against +0.0045 in the baseline, every one-sided test for a decline still rejects at the 5 % level, and the 2010 tail slope is unchanged or larger (the weights enlarge it to +0.0535).

Recall. The 1980 cohort is reported at the longest recall, so its slope could reflect recall error in dates of birth and death. Classical error of that kind would bias the slope toward zero and so work against the decline we find, but we also tighten the recall cap directly. Under a 15-year cap the 1980 slope is +0.0038 ($p = 0.054$) and the decline tests still reject against the 1990 and 2010 cohorts at the 5 % level and against the 2000 cohort at the 10 % level (Section 6).

Exposure timing. Because the exposure window starts at birth, heat-induced early delivery would place preterm babies, who face higher mortality, in the hot weather that triggered their delivery, so postnatal heat could partly proxy for preterm birth. We address this

in two ways. First, we anchor the window at the expected delivery date (conception +9 months), which heat cannot move: the 1980 high-heat slope is +0.0043 and the 2010 tail slope +0.0524, both significant at the 5 % level. Second, we test the mechanism directly on the births with recorded gestation: heat in the expected third trimester does not raise the probability of a preterm birth (+0.0005 per +10 days, SE 0.0088). A related concern is that the cohort decline reflects a changing response to heat *in utero* rather than after birth; letting the in-utero heat bins vary by birth decade leaves the 1980 postnatal slope at +0.0051, with every decline test rejecting at the 5 % level. Conception-month fixed effects, a control for heat in the month before birth, and restricting to nine-month gestations leave both results in place, with every decline test significant at the 10 % level or better (Section 6).

Specification. The cohort contrast might pick up district-specific secular change beyond the baseline's linear trends, or depend on the two cutoffs that define the margins. On the first, adding district-specific quadratic trends leaves the 1980 high-heat slope at +0.0045 with every decline test rejecting at the 5 % level, and letting every temperature bin, not only the two margins, vary by birth decade changes neither result, the added cold-bin × cohort interactions being jointly insignificant ($F = 1.41$, $p = 0.224$). On the cutoffs, the tail result holds at thresholds of 40 to 42°C and the 1980 high-heat slope stays positive when the 10°C band is shifted (Section 5.4), and both results survive 5-year cohort bins (Section 6). Because neonatal death is a rare outcome, we also re-estimate by Poisson pseudo-maximum likelihood: the tail estimates are close to the linear ones, while the high-heat pattern has the same direction but is smaller and less precise (Online Appendix Section F.5).

Inference. District-clustered standard errors might overstate precision given 64 districts, spatially correlated weather and a tail effect identified from few district-years. We therefore also use Conley spatial heteroskedasticity- and autocorrelation-consistent (HAC) standard errors (Conley, 1999), two-way clustering by district and birth year (Cameron et al., 2011), and a restricted wild cluster bootstrap (Cameron et al., 2008; Webb, 2023). The high-heat decline tests reject at the 5 % level under all three in every case but one (the 2000 contrast under two-way clustering, $p = 0.067$); the 2010 tail slope is significant at the 1 % level under two-way clustering and at the 10 % level under Conley standard errors and the bootstrap (bootstrap $p = 0.094$).

Selection into live birth. If extreme heat ends pregnancies, the live births we observe are selected. In the BDHS-2022 pregnancy recode, ten additional third-trimester ≥ 40°C days raise the stillbirth probability by about 1.7 percentage points, against a 2.6 % stillbirth rate ($p = 0.087$), while full-gestation exposure is null; if so, the observed births are positively selected and our neonatal estimates, including $\widehat{\beta}^{\geq 40}_{2010}$, are if anything lower bounds. Tests of

conception timing, of the number of sampled births and of the sex ratio at birth show no fall in births and no female-selective loss, so none indicates selection in the direction that would produce our results (Section 6).

Other weather and the districts at risk. Humidity or rainfall, rather than heat itself, might be what makes hot days deadly, and because the districts with extreme heat change over time, the cohort tail slopes might be identified from different places. The first is not the case: ≥ 40°C days average 49 % relative humidity against roughly 80 % in the 25–35°C range, and controlling for postnatal humidity leaves the 2010 slope at +0.0458 (Section 5.3); rainfall controls change neither result. Restricting to the 21 districts with ≥ 40°C days in every decade, which identifies every tail slope from the same places, leaves the 2010 tail slope at +0.0437 (Section 6).

Related literature and contribution. This paper speaks to three strands of work.

Adaptation of the heat–mortality relationship. The first documents that the mortality cost of heat falls as populations adapt. In the United States the effect of a hot day on mortality fell by about three-quarters over the twentieth century, a decline that Barreca et al. (2016) attribute essentially entirely to the post-1960 diffusion of residential air conditioning; Deschênes and Greenstone (2011) model residential energy use as self-protection against heat, and Heutel et al. (2021) find that hot days are less deadly in warmer places, consistent with greater air-conditioning prevalence there. In the US decline, residential electrification itself played no detectable role (Barreca et al., 2016). This evidence comes from a rich country in which air conditioning was the margin of adaptation; it does not say whether, where in the temperature distribution, or through what channel adaptation occurs in a poor and hot country where air conditioning is essentially absent. We provide that evidence: adaptation at high heat, none detected at the extreme, and rural electrification, which in Bangladesh brought fans rather than air conditioning, as a candidate channel.

Heat, mortality and public programs in poor populations. The second shows that hot days raise mortality in developing countries, including among infants, and that public programs can blunt the effect. In India hot days raise mortality in rural but not urban areas, through agricultural incomes (Burgess et al., 2017); across 53 developing countries, Bangladesh among them, hot and humid days raise neonatal and infant mortality (Geruso and Spears, 2018); in rural India a community health-worker program, but not the rural employment guarantee, attenuated the effect of in-utero heat on infant mortality (Banerjee and Maharaj, 2020); and in Mexico, where weather-related deaths are mostly cold-related, a universal health-care program reduced cold-related mortality (Cohen and Dechezleprêtre, 2022). Bangladesh

enters this strand only through pooled multi-country and South Asian estimates, and the strand measures the level of the heat–mortality relationship at a point in time rather than its change. We provide within-country, cohort-level estimates for Bangladesh, the eighth-most populous country (United Nations, 2022) and one of those for which Carleton et al. (2022) project the largest mortality costs of warming, and we test infrastructure, rather than a health program, as the adaptation channel.

Projecting mortality damages. The third projects the mortality damages of warming by letting the temperature response attenuate as incomes rise and places adapt to warmer climates (Carleton et al., 2022; Auffhammer, 2018). The response functions behind these projections are estimated where subnational, age-specific mortality records exist, in 40 countries, none of them in South Asia (India's all-age data serve only for cross-validation), and are extrapolated to the rest of the world, Bangladesh included, as a smooth function of income and climate. Missing is a direct estimate of the response, and of its change over time, in a poor and hot setting. Ours shows adaptation that is uneven across the temperature distribution, real at high heat and undetected at the extreme, which attenuation with income and climate need not capture (Section 9).

To a literature built largely on level effects, we therefore add three things: causal estimates of the neonatal heat–mortality relationship in Bangladesh, cohort-level evidence on *where in the temperature distribution* adaptation has and has not occurred, and a test of one candidate mechanism, rural electrification, for the range where it has.

2 Background

Three features of the Bangladeshi setting motivate the analysis: a neonatal mortality rate that fell sharply in the 1980s and 1990s and has since declined more slowly from a high level, a newborn physiology that is unusually exposed to ambient heat, and a warming climate in which the increase is concentrated in the 30–40°C range rather than at the extreme tail above 40°C.

2.1 Neonatal Mortality: A Slowing Decline at a High Level

In our BDHS birth histories the neonatal mortality rate falls from 73 deaths per 1,000 live births for the 1980 birth cohort to 51 for the 1990 cohort, 37 for the 2000 cohort, and 30 for the 2010 cohort; the BDHS-2022 round puts the current rate near 20 per 1,000 (NIPORT and ICF, 2024). The decline has slowed—the 37-to-30 drop between the 2000 and 2010 cohorts

is a fraction of the 73-to-51 fall two decades earlier—in line with a global deceleration since 2010 that leaves nearly half (47 %) of all deaths of children under five in the first 28 days of life (World Health Organization, 2024a). The remaining level is high. The UN Inter-agency Group for Child Mortality Estimation (UN IGME) puts Bangladesh's rate at 18 per 1,000 in 2022 (World Health Organization, 2024b), slightly below the Central- and Southern-Asian rate of 21, against 0.7 in the lowest-mortality country (World Health Organization, 2024a). From about 96 per 1,000 in the early 1970s (roughly 15 in the United States, 14 in Germany, and 8 in Japan; World Health Organization, 2024b) it has fallen by about four-fifths, yet remains *above* the rate the United States had already reached by 1970. Neonatal deaths are now the binding margin of child survival in Bangladesh, so any factor that raises neonatal risk—heat among them—acts on a high and persistent base.

2.2 Newborns Are Especially Vulnerable to Heat

Newborns are among the age groups most physiologically exposed to ambient heat. A high surface-area-to-mass ratio helps an infant shed heat when the air is cooler than the skin, but when the air is hotter than the skin the same geometry loads heat *onto* the infant, although how far this raises core temperature is still uncertain (van de Kamp and Daanen, 2025). Infants also sweat from birth, but with a lower output per gland than adults (van de Kamp and Daanen, 2025), and cannot thermoregulate behaviorally, so their defense against heat runs through a caregiver and the household's cooling technology. These vulnerabilities plausibly weigh most on preterm and low-birth-weight infants, whom the World Health Organization singles out for additional care, including closer temperature management; premature birth is among the leading causes of neonatal death (World Health Organization, 2024a).

2.3 A Warming Climate

In the ERA5 inverse-distance-weighted district-day panel we construct (Section 3), the mean daily-maximum temperature rose from 29.8°C in the 1980s to 30.4°C in the 2010s, about +0.2°C per decade. Days at ≥ 30°C rose steadily from 189 per district-year in the 1980s to 218 in the 2010s and 229 in the 2020s. Days at ≥ 35°C were flat or slightly lower through the 2000s ($20 \rightarrow 18$) before rising to 24 in the 2010s and 34 in the 2020s. This warming is national: days in the 35–40°C bin trend upward in all eight divisions, fastest in proportional terms in those that started coolest (the middle panel of Online Appendix Figure A.2).

The extreme tail behaves differently. Days at ≥ 40°C are rare in every decade and have *not* trended upward in our panel: the mean is 0.84 days per district-year in the 1980s,

0.81 in the 1990s, 0.59 in the 2000s, and 0.54 in the 2010s, and the share of district-years recording any $\geq$ 40°C day falls from 20 % to 11 %. Extreme heat is also concentrated in space. In the 2010s, 40 of the 64 districts recorded no $\geq$ 40°C day at all, and exposure clusters in the northwest, led by Chapainawabganj (7.0 days per year), with a secondary cluster in the southwest (Online Appendix Section A.5). The flat national trend masks a redistribution: the decline is concentrated in the historically hottest northwest, while Khulna (southwest) and Chattogram (southeast) divisions have edged up slightly (Online Appendix Table A.2). Rajshahi and Rangpur divisions still hold about four-fifths of the 2010-cohort extreme-heat exposure. The identifying variation at the tail is therefore thin and regionally concentrated in every cohort, which is why we document where the extreme-tail effect is identified (Section 5.5).

3 Data

3.1 Bangladesh Demographic and Health Survey

We pool live-birth records from nine rounds of the BDHS: the 1994, 1997, 2000, 2004, 2007, 2011, 2014, 2018, and 2022 rounds (labeled by the year fieldwork ended; fieldwork for the 1994, 1997, 2000 and 2018 rounds spanned two calendar years). Each round is an independent, nationally representative probability sample of households, drawn anew for that round; the rounds are not a panel, and no woman is followed across rounds. In each round the ever-married women of reproductive age (nominally 15–49; 69 births in the pooled data are to mothers younger than 15 at interview) in the sampled households are eligible for the individual interview, and each interviewed woman reports a complete retrospective birth history: the year and month of every live birth she has ever had, each child's survival status at the interview, and age at death where applicable. Variable codes are harmonized across rounds following the DHS Standard Recode conventions. The pooled sample contains 343,059 live births spanning birth years 1976–2022, with the round-level composition reported in Panel A of Table 1. We do not reach further back because both data sources become less reliable the earlier the year. Births before 1976 that our 20-year recall window, measured from the interview year, would admit—those of 1973–75—are reported only in the first round, at 18 to 20 years' recall, whereas every birth year from 1976 on is reported in at least two rounds. ERA5, although available from 1940, assimilates progressively fewer observations in earlier years, so its quality declines the further back it reaches (Soci et al., 2024); it is markedly weaker before satellite temperature soundings first enter the assimilation in November 1972 (Bell et al., 2021). The three years would add only

1,983 births, 0.6 % of the sample.

Table 1: **BDHS rounds, pooled-sample composition, and sample-size waterfall.**

Panel A. BDHS rounds and pooled-sample composition

Round	Fieldwork year	Birth years captured	Births contributed
BDHS-1994	1993–94	1976–1994	25,408
BDHS-1997	1996–97	1976–1997	24,935
BDHS-2000	1999–00	1976–2000	28,504
BDHS-2004	2004	1976–2004	32,139
BDHS-2007	2007	1976–2007	29,999
BDHS-2011	2011	1976–2011	45,829
BDHS-2014	2014	1978–2014	43,695
BDHS-2018	2017–18	1981–2018	47,828
BDHS-2022	2022	1984–2022	64,722
Pooled total	—	1976–2022	**343,059**

Panel B. Sample-size waterfall by mortality outcome

Restriction	NN (N)	IMR (N)	U5M (N)
Pooled BDHS live births (1976–2022)	343,059	343,059	343,059
Outcome determined & window in ERA5 panel	341,011	327,608	270,115
+ 20-year recall (baseline level sample)	284,102	270,699	213,206
+ cohort window (1980–2010; decomposition)	277,074	266,323	213,206

Notes: Source: BDHS birth-recode files, nine rounds (1993–94 to 2022); one observation per live birth reported by an interviewed ever-married woman nominally aged 15–49, births in 1976 or later (earlier births are reported only in the first round, at the limit of the recall window; Section 3) with a known district. Panel A: "Birth years captured" is the range of the round's birth years in the pooled sample (recall depth of women aged 15–49, truncated at 1976); "Births contributed" is the unweighted number of the round's births, summing to the pooled total. Variable harmonization detail is in Online Appendix Section A.1. Panel B: unweighted births remaining after each cumulative restriction, by outcome: neonatal death (NN; death in the first month of life), infant death (IMR; before 12 months), and under-five death (U5M; before 60 months). "Outcome determined & window in ERA5 panel" keeps births whose full at-risk window (1, 12, or 60 months) had elapsed by the interview, with a recorded age at death for children who died, and whose 28-, 365-, or 1,825-day exposure window lies in the ERA5 panel; "20-year recall" keeps births within 20 years of the interview (the level sample of Table 2); the cohort analysis (Section 5) further restricts to birth decades 1980/1990/2000/2010, where the 1980 cohort also contains the 1976–79 births. The estimation samples are slightly smaller than these rows (e.g., NN cohort sample $N = 276{,}031$) because they drop births lacking in-utero exposure (a gestation window starting before the ERA5 panel) or a control variable. The 2020 cohort is dropped because BDHS-2022 covers only 2020–2022 with thin ≥ 40°C exposure realized so far. Online Appendix Table A.1 shows which rounds supply each birth cohort.

3.2 Mortality Outcomes

We construct three nested mortality outcomes following standard demographic convention. *Neonatal mortality* (NN) equals one if the child died in the first month of life (DHS age at death under one month), *infant mortality* (IMR) if the child died before age twelve months, and *under-five mortality* (U5M) if the child died before age five (sixty months); heat exposure is measured over the first 28, 365, and 1,825 days after birth, respectively. Because exposure is a day count over the whole window, the U5M coefficient averages the per-day effect over 1,825 days, only 28 of which fall in the neonatal period, so an effect confined to the neonatal period is diluted by roughly 28/1,825. An acute physiological mechanism should therefore show up mainly in the neonatal outcome, which is what we find. The three windows are not equally informative about the timing of exposure: for a neonatal death the 28-day window spans essentially the whole at-risk period and so contains the heat in the days before death, whereas the 365- and 1,825-day windows average over long periods in which a child may already have died. We therefore study the infant and under-five outcomes mainly as comparisons, and we check them with a design that uses heat at the time of death: a child-month panel in which death in each month of age is related to heat in that month and the month before (Section 6).

Standard child, household, and maternal covariates are harmonized across rounds: child sex, urban-rural residence, household wealth quintile computed within each round, maternal education, electricity, and birth order. Fan indicators (BDHS-2011 through BDHS-2022) and air-conditioning indicators (BDHS-2014 through BDHS-2022) used in Section 7 are recovered directly from the raw household recodes.

3.3 Heat Exposure: Inverse-Distance Weighted ERA5

Daily-maximum 2-meter air temperature is constructed from ERA5 hourly fields at 0.25° resolution, distributed by the Copernicus Climate Change Service through its Climate Data Store (Hersbach et al., 2023), for 1976–2024 at the 205 grid cells whose centroids fall inside Bangladesh. We take the daily maximum over all 24 hourly fields at each grid cell, and aggregate grid-cell daily maxima to district-day exposure by inverse-distance weighting (IDW) to the (area-weighted) centroid of each of Bangladesh's 64 districts: for each district we take the four nearest ERA5 grid cells inside Bangladesh, compute Haversine distances $\{d_i\}_{i=1}^{4}$ from the district centroid, and weight by $w_i = 1/d_i^2$, normalized so $\sum_i w_i = 1$. The median distance from a district centroid to its four contributing cells is 20.8 km; the 99th percentile is 36.0 km. Coverage is complete: every district has a daily series for every year of the panel

(Online Appendix Section A.2).

Validation against station observations. Because the analysis rests entirely on a reanalysis product, we check the ERA5 IDW panel against Bangladesh Meteorological Department (BMD) station records of daily maximum dry-bulb temperature. Matching each station to its district and comparing with the ERA5 IDW district-day maximum on the same date gives 543,418 matched station-days across 45 of the 48 BMD stations and 37 districts (1976–2024). Agreement is close: the correlation is 0.91, the mean bias is -0.20°C (station mean 30.04°C versus ERA5 29.84°C), the mean absolute error is 1.10°C, and the root-mean-square error is 1.47°C. At thresholds, the district-level IDW aggregate counts fewer exceedance days than a point station—83 % as many ≥ 40°C days (544 versus 659) and 75 % as many ≥ 35°C days—which is what one expects when a spatially smoothed district aggregate is compared with a point observation.

We do not construct exposure from the stations: they are sparse relative to the four-source IDW we require (in 1980 the fourth-nearest station is a median 94 km from a district centroid, against 28.0 km for the fourth-nearest ERA5 cell), 33 % of the 48-station daily grid over 1975–2024 is empty, and—decisively for a cohort design—the network thickens over time (24–28 reporting stations in 1976–1980 against 48 in 2023), which would attenuate the early-cohort slopes mechanically. The ERA5 grid is identically dense in every year (Online Appendix Section A.3).

Exposure measures. For each birth we count days in six 5°C bins (< 20, 20–25, 25–30 [reference], 30–35, 35–40, and ≥ 40°C) during the postnatal at-risk window of length $w \in \{28, 365, 1{,}825\}$ days from birth (the birth date through day $w - 1$). The BDHS birth histories give month and year of birth; we place each birth on the 15th of its month, so the 28-day window is approximate. Some births have an imputed month or year (DHS completeness flag); Section 6 shows that dropping them changes neither result. No daily maximum in the panel falls below 15°C within any at-risk window, so the open-ended < 20°C bin effectively covers 15–20°C. The exposure window is fixed at its planned length regardless of date of death, so heat-killed newborns are not mechanically attributed lower exposure than survivors. The BDHS does not record conception dates; we date conception nine months before birth for the gestational (in-utero) exposure, which counts days in each bin over the nine calendar months preceding the birth month, and for the conception-month variables, and use recorded months of pregnancy where available (the reproductive calendars of the 2000, 2007, 2014 and 2018 rounds and the 2022 birth history, for births in the five years

before the interview) to test whether heat shifts birth timing (Section 6). The 1994, 1997, 2004 and 2011 calendars record almost every live birth as lasting nine months.

3.4 Sample Construction and Selection

Because every birth is reported retrospectively by a mother interviewed in one of nine survey rounds, the BDHS design determines which births we observe and through whom. This subsection documents how the sample is built, how each birth cohort is observed, and what that implies for the cohort comparisons.

Who is interviewed. All nine rounds interview only ever-married women of reproductive age; every mother in our pooled data is currently or formerly married. The upper age limit of 49 applies in every round. The BDHS interviews ever-married women by design because marriage is nearly universal in Bangladesh and childbearing outside it rare: in the 2022 round fewer than 1 % of women aged 35 or older had never married (0.3 % at ages 45–49), the median age at first marriage was 16.9 years, and the DHS's own all-women adjustment treats never-married women as never having been pregnant (NIPORT and ICF, 2024, Tables 4.1 and 4.2). Restricting interviews to ever-married women therefore excludes essentially no births. A woman's marriage date determines whether she is in the interview pool, but once she is interviewed her full birth history is recorded, including births long before the survey. Each round is a fresh probability sample, so a woman's chance of being interviewed in any one round does not depend on earlier rounds, and being drawn twice is very unlikely. Pooling does, however, give more weight to women who are of interview age in more rounds—a woman aged 20 in 1994 stays eligible through 2022, while one aged 45 in 1994 ages out—which is one reason the cohorts differ in who reports them (below). Our baseline estimates are unweighted; weighting by the DHS sampling weights, which restores national representativeness within each round, leaves the cohort pattern intact (Section 6).

Sample restrictions. We apply three restrictions. First, the full exposure window must fall within the ERA5 panel, which we construct for 1976–2024; births before 1976 are dropped. Combined with the recall cap below, this cutoff binds only in the 1994 round, where it removes births from 1973–75; in every later round the recall cap is the binding restriction. Second, the full at-risk window must have been observed before the interview, whatever the child's survival status: a birth enters the NN, IMR, or U5M sample only if the child's age at interview (interview month minus birth month, in completed months) is at least 2, 13, or 61 months, respectively, and children who died must have a recorded age at

death. Survivors too young to have completed the window are therefore excluded rather than right-censored, and children who survived the window but died later are retained as window survivors. Third, the baseline analysis restricts to births within 20 years of the interview (interview year minus birth year), to limit recall error in dates of birth and death; Section 6 shows the results under a 15-year cap. Panel B of Table 1 reports the resulting sample sizes by outcome.

How each birth cohort is observed. Because births are reported retrospectively, each birth cohort reaches us through a different set of survey rounds and at a different recall length (Online Appendix Table A.1). The 1980 cohort comes almost entirely from the 1994–2004 rounds and is reported on average 13.3 years after birth; the 2010 cohort comes from the 2011–2022 rounds at an average recall of 5.3 years. Two features of this structure matter for a cohort comparison. First, round-specific differences in survey practice or reporting could differ across cohorts; adding survey-round fixed effects, which compare births reported in the same round, leaves the results essentially unchanged (Section 6). Second, because only women up to age 49 are interviewed, births in the early cohorts can be reported only by mothers who were still young enough at a later interview, so the early cohorts could over-represent births to young mothers. This does not materially change the composition of births: the mean age of the mother at birth is 23.1–23.8 years in every cohort, and the share born to teenage mothers is 29–35 %. Controlling for the mother's age at birth also leaves the cohort slopes essentially unchanged (Section 6).

Remaining selection. Two forms of selection cannot be removed with birth-history data. Births to mothers who died, or who left the household population, before the interview are never reported, and if such mothers and their children were more vulnerable, the surviving reporters are positively selected; the gap grows with the time between birth and interview, and so is largest for the earliest cohorts. Recall error in dates of birth and death also grows with that interval. Both forms of selection are plausibly tied more to the mother than to the weather a particular child was born into, although maternal deaths around a birth, or misreported dates, could in principle correlate with conditions at that birth. We address them in two ways: by capping recall at 20 years and showing the results under a 15-year cap (Section 6), and by showing that the early-cohort high-heat slope is not an artifact of long recall (Online Appendix Section D.5). Administrative data are not an alternative: civil registration of births and deaths in Bangladesh has been neither complete nor regular over our study period, which is why retrospective survey birth histories are the standard source for child mortality in Bangladesh and in most low-income countries.

4 Empirical Strategy

We estimate two specifications: a level binned dose-response, which locates the mortality effect in the temperature distribution and across outcomes, and a cohort decomposition, our test for adaptation, which lets the slopes on the margins identified in the level response vary across birth cohorts.

4.1 Level Specification

The level specification regresses each mortality outcome on the postnatal-window bin counts:

$$y_{idmy} = \alpha + \sum_{b \neq \text{ref}} \beta_b \, x^{w}_{b,idmy} + \sum_{b \neq \text{ref}} \eta_b \, x^{g}_{b,idmy} + \mathbf{X}'_{idmy}\boldsymbol{\gamma} + \mu_{dm} + \delta_{\text{bord}_i} + \tau_y + \sum_{d \neq d_0} \lambda_d \, \mathbf{1}[\text{dist}_i = d] \, t_y + \varepsilon_{idmy}, \tag{1}$$

where i indexes the child, d the district, m the calendar month of birth, and y the year of birth. μ_{dm} are district $\times$ calendar-month fixed effects ($64 \times 12 = 768$ cells) and τ_y birth-year fixed effects; δ_{bord_i} are birth-order-group dummies, a control. In addition, $\lambda_d \, \mathbf{1}[\text{dist}_i = d] \, t_y$ are district-specific linear trends. The bin index b runs over the six 5°C bins (< 20, 20–25, 25–30, 30–35, 35–40, and ≥ 40°C), with 25–30°C as the omitted reference bin; $x^{w}_{b,idmy}$ is the count of days in bin b during the postnatal at-risk window of length $w \in \{28, 365, 1{,}825\}$ days from birth, with w matched to outcome (NN, IMR, U5M). Furthermore, $x^{g}_{b,idmy}$ is the count of days in bin b during the gestational period, measured as the nine calendar months preceding the birth month (conception is dated nine months before birth). The control vector $\mathbf{X}_{idmy}$ collects child characteristics (sex), household characteristics (urban/rural residence, within-round wealth quintile, electricity), and maternal characteristics (educational attainment, entered as categories). Coefficients β_b are per +10 days throughout.

Standard errors are clustered at the district level (64 districts). We also check robustness to Conley spatial HAC standard errors (Conley, 1999), two-way clustering at the district and birth-year level (Cameron et al., 2011), and the restricted wild cluster bootstrap (Cameron et al., 2008; Webb, 2023).

The β_b coefficients, reported in Section 5, show which bins carry the mortality effect; we use them to choose the two exposure margins of the cohort decomposition below.

4.2 Cohort Decomposition

The level response puts the mortality effect in the extreme tail and finds no robust effect between 30 and 40°C (Section 5). We therefore split hot exposure at 40°C and let the slope on each part vary by birth decade. The 30–40°C margin pools the two bins where we find no robust level effect ($x^w_{\text{mid}} = x^w_{30\text{–}35} + x^w_{35\text{–}40}$); we track it because a small pooled level effect is what one would see if a positive early-cohort response had since been adapted away, so this band is where adaptation would show. The extreme-tail margin is the bin where we do find one (≥ 40°C):

$$\begin{aligned} y_{idmy} = \alpha + \beta_{\text{mid}} x^w_{\text{mid}} + \sum_c \pi_c x^w_{\text{mid}} \mathbf{1}[D_i = c] + \beta_{\geq 40} x^w_{\geq 40} + \sum_c \kappa_c x^w_{\geq 40} \mathbf{1}[D_i = c] \\ + \sum_{b \in \{<20,\, 20\text{–}25\}} \beta_b \, x^w_{b,idmy} + \sum_{b \neq \text{ref}} \eta_b \, x^g_{b,idmy} + \mathbf{X}'_{idmy}\boldsymbol{\gamma} + \mu_{dm} + \delta_{\text{bord}_i} \\ + \tau_y + \sum_{d \neq d_0} \lambda_d \, \mathbf{1}[\text{dist}_i = d] \, t_y + \varepsilon_{idmy}, \end{aligned} \tag{2}$$

where D_i is child i's birth decade, $c \in \{1990, 2000, 2010\}$, and 1980 is the omitted base cohort; the < 20 and 20–25°C postnatal bins enter as controls. We write $\beta_c^{30\text{–}40} \equiv \beta_{\text{mid}} + \pi_c$ and $\beta_c^{\geq 40} \equiv \beta_{\geq 40} + \kappa_c$ for the decade-specific slopes on the 30–40°C margin and on the tail (with $\pi_{1980} = \kappa_{1980} = 0$). Adaptation at high heat means that the 1980 slope exceeds the later ones, so against each later cohort c we test $H_1 : \beta_{1980}^{30\text{–}40} > \beta_c^{30\text{–}40}$, that is $\pi_c < 0$, one-sided with $t(63)$ critical values (64 district clusters). At the tail the question is whether the effect has declined by the most recent cohort, so we test $H_1 : \beta_{2010}^{\geq 40} < \beta_c^{\geq 40}$ against each earlier cohort, also one-sided. The fixed effects are as in equation (1); only the two hot margins, not the full six-bin spline, are interacted with cohort, and Section 6 shows that also interacting the colder bins changes neither margin's cohort pattern.

4.3 Identification

The design controls for place and seasonality jointly, so that a hot day is interpreted as heat that is unexpected for that time of year *and* that place. The district × calendar-month fixed effects μ_{dm} give every district its own month-of-year profile. They absorb all time-invariant differences across districts (climate, baseline mortality risk, health infrastructure, sociodemographic composition) and, within each district, the regular seasonal cycle of both temperature and newborn survival: the pre-monsoon heat of April and May in the northwest, the monsoon disease season, and seasonal patterns in births, agricultural work and food availability. A

birth in Rajshahi in May is therefore compared only with births in Rajshahi in May of other years, never with births in December or with births elsewhere in May. Because conception month is birth month minus nine, the same cells also absorb the conception-month cycle. Identification rests on the assumption that, conditional on these fixed effects and the controls, year-to-year deviations in the temperature distribution within a district-month cell are orthogonal to other determinants of newborn survival. Online Appendix Table F.1 reports the baseline results under weaker, additive district and season effects. The birth-year fixed effects τ_y absorb national-level trends in child survival, including rising maternal education and immunization coverage, declining fertility, expansion of primary care, and the secular decline in infant mortality. Moreover, we include 63 district-specific linear trends, so that the cohort interactions are identified net of differential district-level secular change (rising health infrastructure, urbanization). We also test robustness to adding district-specific quadratic trends. Birth-order fixed effects absorb the well-documented gradient in birth-order risk. Finally, the in-utero heat bins ensure that the postnatal-window coefficients do not capture temporally correlated gestational heat. We address the remaining threats empirically in Section 6.

What the cohort slopes identify. The cohort comparison does not compare mortality levels across decades, which changed for many reasons; it compares how strongly mortality responds to year-to-year heat variation within district-month cells, net of everything common to a birth year and of gradual differential change across districts. Bangladeshi society did change greatly over these decades, but that change is the object of the test rather than a confound of it: adaptation, in the sense that damage functions assume, is precisely a weakening of the heat–mortality link as incomes, infrastructure and behavior evolve, and the policy-relevant quantity is how the slope moves over time, not the slope of any one cohort. What must be ruled out is that the cohorts differ in how they were observed rather than in how they respond to heat. Each BDHS round is an independent, nationally representative probability sample (Section 3.4), and the cohort slopes are essentially unchanged when survey-round fixed effects, sampling weights or controls for the mother's age at birth are added (Section 6), so neither sampling design nor the retrospective collection of births drives the comparison. A robust decline in the heat response under these controls is therefore a decline in the sensitivity of newborn survival to heat—that is, adaptation—whatever its channel; Section 7 asks which channel. For the extreme tail we make a narrower claim: the effect is present in the pooled (all-cohort) regression and remains significant for the most recent cohort, and no test detects a decline from earlier cohorts; we do not interpret differences among the earlier cohorts' tail slopes, which are imprecise.

Measurement error and birth timing. Classical measurement error, larger for the early cohorts, would attenuate the 1980 high-heat slope toward zero and so work against the decline we find (Section 6). Because exposure and the fixed effects are anchored on the date of birth, heat-induced early delivery could make postnatal heat partly proxy for preterm birth; Section 6 tests this directly and re-anchors exposure at expected delivery (Online Appendix Section E.1).

5 Results

5.1 Where the Mortality Effect Lies in the Temperature Distribution

Before testing for adaptation, we locate the mortality effect in the temperature distribution and across outcomes, which determines the margins we track across cohorts. Table 2 reports the binned dose-response for each outcome, under both the full and the 20-year recall sample.

We find two things in the level response. First, the heat effect sits mainly in the extreme tail. The ≥ 40°C coefficient is positive and significant in both samples (+0.0187, $p = 0.013$, full sample; +0.0196, $p = 0.031$, 20-year recall), so ten additional days above 40°C are associated with about 19 to 20 additional deaths per 1,000 live births, or about 2 per 1,000 per day, a tenth of the current neonatal mortality rate of 20 per 1,000 in Bangladesh. We find no robust evidence of neonatal mortality effects of exposure below 40°C: the only significant hot-bin (30–40°C) coefficient is the full-sample 30–35°C bin (+0.0023, $p = 0.021$), which is smaller and not significant in the 20-year recall sample (+0.0020, $p = 0.148$).[1] Second, the effect is concentrated in the neonatal window. The ≥ 40°C coefficients for infant and under-five mortality are much smaller: only the full-sample infant coefficient is significant, and only at the 10 % level (+0.0053, $p = 0.0999$, under a third of the neonatal coefficient), and none is significant in the 20-year recall sample. This is the expected signature of an acute effect, diluted in the longer windows (Section 3). Splitting each of the two upper hot bins into isolated days and heat-wave days (≥ 3 consecutive days in the band) does not settle whether the tail risk comes from sustained heat: for neonatal mortality, estimated on the 1980–2010 birth cohorts without a recall cap ($N = 331{,}709$) rather than on the level sample of Table 2, the ≥ 40°C heat-wave-day component is significant (+0.0187, $p = 0.028$) and

[1] The < 20°C bin is negative, i.e., protective, though significant only in the 20-year recall sample and only at the 10 % level: ten additional days below 20°C are associated with about 18 fewer neonatal deaths per 1,000 in the 20-year recall sample ($p = 0.055$) and 8 fewer in the full sample ($p = 0.392$). Bangladesh has no true winter: no daily maximum in our sample falls below 15°C, so this bin captures mild days, which are colder than usual for the country but physiologically favorable. Given our focus on adaptation to extreme heat, we do not pursue it.

Table 2: **Level binned-temperature response.**

	NN		IMR		U5M	
	No cap	20-yr	No cap	20-yr	No cap	20-yr
< 20°C	−0.0083	−0.0175*	−0.0021	−0.0064	−0.0000	−0.0014
	(0.0096)	(0.0089)	(0.0050)	(0.0061)	(0.0026)	(0.0029)
20–25°C	+0.0003	−0.0006	−0.0011	−0.0004	+0.0003	+0.0009
	(0.0017)	(0.0018)	(0.0010)	(0.0010)	(0.0006)	(0.0007)
25–30°C (ref.)	—	—	—	—	—	—
30–35°C	+0.0023**	+0.0020	−0.0000	+0.0003	+0.0004*	+0.0004
	(0.0010)	(0.0013)	(0.0005)	(0.0006)	(0.0002)	(0.0003)
35–40°C	−0.0003	−0.0000	−0.0006	−0.0013*	−0.0001	−0.0003
	(0.0020)	(0.0022)	(0.0007)	(0.0008)	(0.0005)	(0.0005)
≥ 40°C	+0.0187**	+0.0196**	+0.0053*	+0.0032	+0.0028	+0.0013
	(0.0073)	(0.0089)	(0.0032)	(0.0036)	(0.0021)	(0.0022)
N	338,737	283,059	325,341	269,663	267,883	212,205
Districts	64	64	64	64	64	64

Notes: Linear probability models of regression (1), one observation per birth, of neonatal death (NN; first month of life), infant death (IMR; before 12 months) and under-five death (U5M; before 60 months). Sample: births in the nine BDHS rounds whose full at-risk window had elapsed by the interview, no birth-cohort restriction (births 1976 through 2022, 2021 and 2017 for NN, IMR and U5M); "20-yr" keeps births within 20 years of the interview, "No cap" imposes no recall restriction. N (No cap, 20-yr): NN 338,737, 283,059; IMR 325,341, 269,663; U5M 267,883, 212,205. Entries: days in each 5°C bin of district daily-maximum temperature over the outcome's 28-, 365- or 1,825-day window from the birth date, per +10 days, relative to 25–30°C; multiplying a coefficient by 1,000 gives deaths per 1,000 live births per +10 days. Fixed effects: district × birth calendar month and birth year. Controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile, maternal education (none, primary, secondary, higher), child sex, urban residence, household electricity, 63 district-specific linear birth-year trends and in-utero heat bins (days in each temperature bin over the nine calendar months before the birth month). Standard errors clustered by district (64) in parentheses; $t(63)$. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

its isolated-day component larger but imprecisely estimated (+0.0243, $p = 0.302$; Online Appendix Table B.1).

We therefore study neonatal mortality and track two margins across birth cohorts: the high-heat band, 30 to 40°C, where the responses are small and not robust across samples, and the extreme tail, above 40°C, where they are large and robust across samples. The tail margin follows from where the level effect is; the 30–40°C margin is where adaptation would show (Section 4).

5.2 Tests for Adaptation across Cohorts

If rising income and technology access mitigate heat-related mortality, the dose-response function should flatten across birth cohorts. Our baseline specification is the joint decomposition in equation (2): we interact both the high-heat (30–40°C) and the extreme-tail ($\geq$ 40°C) exposure with birth decade in a single regression, so each margin is estimated net of the other and the two compete for the same cohort variation. Table 3 reports the decade-specific slopes on both margins together with the cross-cohort tests.

The two margins behave differently (Figure 1). At high heat the positive response is confined to the earliest cohort: $\widehat{\beta}_{1980}^{30\text{–}40} = +0.0045$ ($p = 0.015$, significant at the 5 % level), about 4.5 additional neonatal deaths per 1,000 live births per +10 days of exposure between 30 and 40°C, against a 73 per 1,000 base rate in that cohort. Every later cohort is indistinguishable from zero, and the one-sided test for a decline from the 1980 cohort rejects at the 5 % level against each of them ($p = 0.012$, 0.027 and 0.013 for the 1990, 2000 and 2010 cohorts). At the extreme tail no test detects such a decline, and we take that up below. Since both margins are estimated in the same regression, neither pattern is an artifact of attributing the cohort movement to one margin or the other; nor is the decline a feature of how each cohort was sampled, since it is essentially unchanged with survey-round fixed effects and DHS sampling weights (Section 6).

Because 1980 is the base cohort in the decomposition, each cohort interaction equals the contrast $\widehat{\beta}_c - \widehat{\beta}_{1980}$ (Online Appendix Table C.1). The contrasts are uniformly negative and of similar size (1990: −0.0036; 2000: −0.0026; 2010: −0.0036), with one-sided 95 % upper bounds of −0.0010, −0.0004 and −0.0010, so the 1980 slope of +0.0045 falls to essentially zero by the 1990 cohort and stays there; the three later cohorts' slopes are jointly indistinguishable from zero ($F = 0.91$, $p = 0.443$).

At the extreme tail we find a positive effect that is statistically significant only for the most recent (2010) cohort. The 2010 slope is +0.0432 ($p = 0.009$): ten additional days of $\geq$ 40°C exposure during the 28-day neonatal window are associated with about 43 additional deaths per 1,000 live births, against a 2010-cohort base rate of 30 per 1,000. The 2000 point estimate is of similar size (+0.0441) but less precise and not statistically significant ($p = 0.173$). The 1980 and 1990 slopes are small positive estimates that are not statistically significant (+0.0078 and +0.0076). Because adaptation would mean a decline, Table 3 tests $H_1 : \beta_{2010}^{\geq 40} < \beta_c^{\geq 40}$ against each earlier cohort. No one-sided test comes close to rejecting ($p = 0.927$, 0.983 and 0.489 against 1980, 1990 and 2000). No test therefore detects a decline in the tail effect, although the earlier tail slopes, estimated on sparse extreme-heat

Table 3: **Joint cohort decomposition, neonatal mortality.**

	Panel A. High heat (30–40°C)		Panel B. Extreme tail (≥ 40°C)	
Birth decade	$\widehat{\beta}$	p	$\widehat{\beta}$	p
1980	+0.0045**	0.015	+0.0078	0.672
	(0.0018)		(0.0182)	
1990	+0.0009	0.563	+0.0076	0.511
	(0.0016)		(0.0116)	
2000	+0.0018	0.157	+0.0441	0.173
	(0.0013)		(0.0320)	
2010	+0.0009	0.537	+0.0432***	0.009
	(0.0014)		(0.0161)	
High heat: one-sided tests for a decline, p				
$\beta_{1980} > \beta_{1990}$	0.012			
$\beta_{1980} > \beta_{2000}$	0.027			
$\beta_{1980} > \beta_{2010}$	0.013			
Extreme tail: one-sided tests for a decline, p				
$\beta_{2010} < \beta_{1980}$			0.927	
$\beta_{2010} < \beta_{1990}$			0.983	
$\beta_{2010} < \beta_{2000}$			0.489	
N	276,031			
Districts	64			

Notes: Linear probability model of neonatal death (death in the first month of life), equation (2), one observation per birth. Sample: the 20-year recall sample of the 1980–2010 birth cohorts (births 1976–2019 within 20 years of the interview in the nine BDHS rounds, 1993–94 to 2022, whose first month had elapsed by the interview; 1976–79 births are in the 1980 cohort). Exposure: days in each 5°C bin of district daily-maximum temperature over the 28 days from the birth date, per +10 days; the 30–40°C and ≥ 40°C counts are both interacted with 1990, 2000, and 2010 birth-decade indicators (1980 base) in one regression, so each margin is estimated net of the other, and the < 20 and 20–25°C counts are controls (25–30°C reference). Entries are total cohort slopes (main effect plus decade interaction); multiplying by 1,000 gives deaths per 1,000 live births. Fixed effects: district × birth calendar-month and birth-year. Controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile (within round) and maternal education (none, primary, secondary, higher), each as categorical dummies; child sex; urban residence; household electricity; 63 district-specific linear birth-year trends; and in-utero heat bins (days per bin over the nine calendar months before the birth month). β_c denotes $\beta_c^{30\text{–}40}$ in Panel A and $\beta_c^{\geq 40}$ in Panel B. Panel A, high heat: one-sided p-values for a decline from the 1980 cohort, $H_1 : \beta_{1980} > \beta_c$; Panel B, extreme tail: one-sided p-values for a decline to the 2010 cohort, $H_1 : \beta_{2010} < \beta_c$; both from $t(63)$. Standard errors clustered by district (64) in parentheses; p-values beside coefficients are two-sided tests of $H_0 : \beta = 0$ from $t(63)$. p-values that would round onto 0.010, 0.050 or 0.100 are shown to four decimals. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

exposure, are too noisy to rule one out, or for us to interpret differences among them. Under the wild cluster bootstrap and Conley standard errors the 2010 effect is significant only at the 10 % level (bootstrap $p = 0.094$; Section 6), and it is identified from the most extreme district-years in the northwest (Section 5.5). The dose-response flattening that we observe

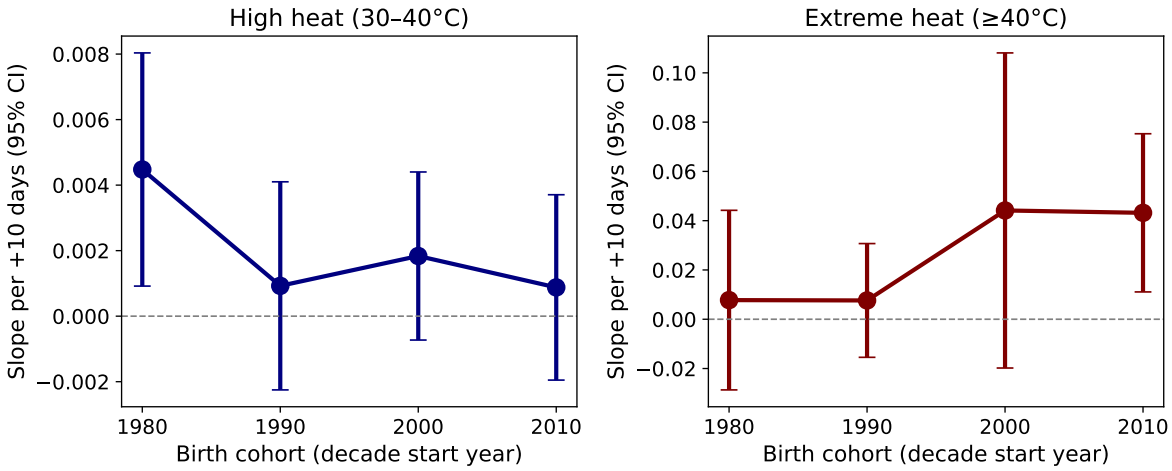


Figure 1: **Neonatal heat-mortality slope by birth decade.** *Left:* high heat (30–40°C); *Right:* extreme heat (≥ 40°C). Total cohort slopes (main effect plus decade interaction) per +10 days in the 28-day neonatal window, from the single joint decomposition of Table 3 (linear probability model of neonatal death; 20-year recall sample of the 1980–2010 birth cohorts, N = 276,031), so each margin is net of the other. Fixed effects: district × birth calendar month and birth year. Controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile, maternal education (none, primary, secondary, higher), child sex, urban residence, household electricity, 63 district-specific linear birth-year trends and in-utero heat bins (days in each temperature bin over the nine calendar months before the birth month). Standard errors clustered by district (64); $t(63)$. Bars: 95 % confidence intervals.

between 30 and 40°C is therefore not detectable in the extreme hot tail of the temperature distribution.

Online Appendix Section B repeats the extreme-tail cohort interaction on heat-wave-day exposure (≥ 40°C days inside runs of at least three consecutive such days). The cohort pattern on heat-wave days resembles that for all ≥ 40°C days: the 2010 slope is +0.0687 ($p < 0.001$) and the 2000 slope +0.0726 ($p = 0.077$), against +0.0038 and −0.0111 for the 1980 and 1990 cohorts. No one-sided test for a decline from an earlier cohort to the 2010 cohort rejects (smallest $p = 0.463$; Online Appendix Table B.2), so here too the tail effect is present in the most recent cohort. Because a level decomposition cannot reliably separate heat-wave days from isolated extreme days (Online Appendix Table B.1), we do not attribute the tail effect to heat waves specifically.

5.3 Humidity

A dry-bulb ≥ 40°C effect could reflect humid heat stress. Barreca (2012) shows that humidity, entered as specific-humidity bins alongside temperature, independently affects US

mortality, although the effect of the hottest days survives humidity controls; our control is relative humidity, which he avoids because it is mechanically related to temperature, so the test is in the spirit of his analysis rather than a replication. In Bangladesh the extreme-heat days are the drier days: mean relative humidity is about 80 % from the 25–30 through the 30–35°C bin, then falls to 67 % at 35–40°C and 49 % at $\geq$ 40°C, and 95 % of $\geq$ 40°C district-days fall in the pre-monsoon months of April and May rather than in the humid monsoon. Adding the mean relative humidity over the 28-day postnatal window as a control to the joint decomposition moves neither margin much. On the 269,407 births with humidity data, the 2010 tail slope goes from +0.0431 (on the same humidity subsample, without the control) to +0.0458, significant at the 1 % level; the 1980 high-heat slope goes from +0.0040 (same subsample, without the control) to +0.0045, significant at the 5 % level, and the one-sided tests for a decline from it still reject against the 1990 and 2010 cohorts at the 5 % level and against the 2000 cohort at the 10 % level ($p = 0.062$). The humidity coefficient is itself small and insignificant (+0.00017, $p = 0.242$). Hence neither the high-heat adaptation nor the extreme-tail effect is a humid-heat-stress artifact. Online Appendix Section C.2 reports the humidity profile by temperature bin and season and the full estimates (Online Appendix Figure C.1 and Online Appendix Tables C.2 and C.3).

5.4 Threshold Sensitivity

Both headline effects rest on a temperature cutoff, so we test each for sensitivity to that choice. We first vary the upper tail threshold for the 2010-cohort extreme-heat effect. Figure 2 plots the estimated 2010-cohort slope $\widehat{\beta}^{\geq T}_{2010}$ on $\geq T$°C postnatal exposure across T. The slope is indistinguishable from zero at $T = 38$°C (+0.0088), is smaller but already significant at the 5 % level at $T = 39$°C (+0.0194), is much larger at $T = 40$°C (+0.0432), and larger still at higher T. The break therefore lies at 38–40°C. This dry-bulb threshold does not coincide with core body temperature ($\sim$37°C), and we do not offer a physiological interpretation of where it falls.

We run the analogous exercise for the 1980-cohort 30–40°C effect, whose base case is that band. Shifting the band up in 1°C steps—28–38, 29–39, 30–40, 31–41—leaves the 1980 slope positive in every band (Figure 3). It is +0.0025 and +0.0030 in the two lower bands, positive but imprecise, and +0.0045 and +0.0043 in the 30–40 and 31–41°C bands, significant at the 5 % level. The adaptation result is therefore not an artifact of the exact cutpoints; the 1980 slope is larger, and significant, only once the band's lower edge reaches 30°C.

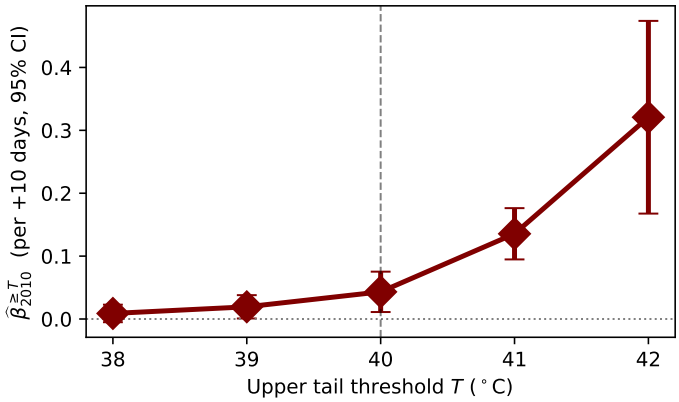


Figure 2: 2010**-cohort neonatal heat-mortality slope by tail threshold T.** Each point is $\widehat{\beta}^{\geq T}_{2010}$ (per +10 days in the 28-day neonatal window) with its 95 % confidence interval, from a separate joint decomposition (linear probability model of neonatal death; 20-year recall sample of the 1980–2010 birth cohorts) in which the high-heat margin is redefined as $[30, T)$°C and the tail as $\geq T$°C, both interacted with birth decade; $T = 40$ reproduces Table 3 (N = 276,031). Fixed effects: district × birth calendar month and birth year. Controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile, maternal education (none, primary, secondary, higher), child sex, urban residence, household electricity, 63 district-specific linear birth-year trends and in-utero heat bins (days in each temperature bin over the nine calendar months before the birth month). Standard errors clustered by district (64); $t(63)$. Beyond $T = 40$ the slope is +0.136 ($T = 41$) and +0.321 ($T = 42$).

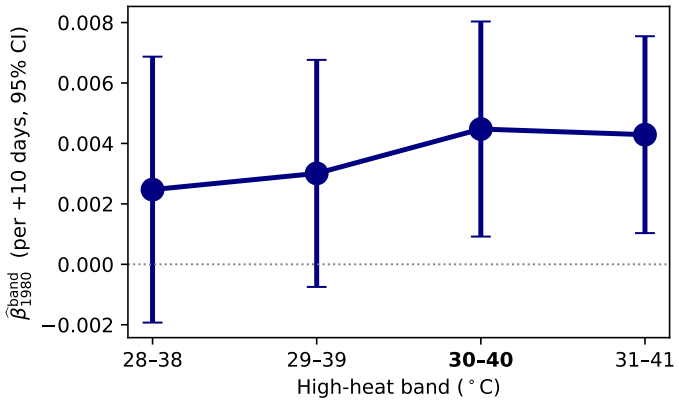


Figure 3: 1980**-cohort neonatal slope by temperature band.** Each point is $\widehat{\beta}_{1980}$ on days in the indicated 10°C band (per +10 days in the 28-day neonatal window) with its 95 % confidence interval, from a separate joint decomposition (linear probability model of neonatal death; 20-year recall sample of the 1980–2010 birth cohorts) in which days in the band and days above it are both interacted with birth decade (days from 25°C to the band's lower edge are the reference); the 30–40°C base case (bold) reproduces Table 3 (N = 276,031). Fixed effects: district × birth calendar month and birth year. Controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile, maternal education (none, primary, secondary, higher), child sex, urban residence, household electricity, 63 district-specific linear birth-year trends and in-utero heat bins (days in each temperature bin over the nine calendar months before the birth month). Standard errors clustered by district (64); $t(63)$.

Table 4: **Where the 2010-cohort extreme-heat effect comes from.**

Sample	$\widehat{\beta}^{\geq 40}_{2010}$	p	N
Baseline (all districts)	$+0.0432^{***}$	0.009	276,031
	(0.0161)		
Excluding a region			
− Rajshahi division	+0.0348	0.316	244,961
	(0.0344)		
− Rangpur division	+0.0327	0.112	244,632
	(0.0203)		
− Khulna division	$+0.0575^{***}$	< 0.001	243,169
	(0.0146)		
− Chapainawabganj district	+0.0272	0.111	273,168
	(0.0168)		
Excluding the most extreme exposure cells (district-years)			
− max cell (Chapainawabganj 2012, 30 days)	+0.0135	0.458	275,979
	(0.0181)		
− top 1 % exposure cells ($\geq$ 18 days, 8 cells)	+0.0165	0.358	275,616
	(0.0178)		
− top 5 % exposure cells ($\geq$ 13 days, 25 cells)	$+0.0318^{*}$	0.095	274,499
	(0.0188)		

Notes: Each row re-estimates the joint cohort decomposition of Table 3 (linear probability model of neonatal death; 20-year recall sample, 1980–2010 cohorts) on the indicated sample and reports the 2010-cohort ≥ 40°C slope (total cohort slope per +10 days), its p-value, and N. Fixed effects: district × birth calendar month and birth year. Controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile, maternal education (none, primary, secondary, higher), child sex, urban residence, household electricity, 63 district-specific linear birth-year trends (those of dropped districts drop out) and in-utero heat bins (days in each temperature bin over the nine calendar months before the birth month). Standard errors clustered by district (64, fewer when a division is dropped) in parentheses; p-values beside coefficients are two-sided tests of $H_0 : \beta = 0$ from $t(G-1)$. Exposure cells are district-years (birth year), ranked by the number of ≥ 40°C district-days, among cells with any such days and with births in the estimation sample; deleting a cell drops its births. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5.5 Where the Extreme-Heat Effect Comes From

Table 4 shows where in the sample the 2010-cohort tail slope is identified. Geographically, it comes from the northwest: dropping Rajshahi division lowers the slope to +0.0348 and leaves it imprecise ($p = 0.316$), and dropping Rangpur division lowers it to +0.0327 ($p = 0.112$), whereas dropping Khulna division raises it to +0.0575 ($p < 0.001$) and dropping any of the other five divisions leaves it between +0.0420 and +0.0452. Dropping any one of the 64 districts leaves the slope between +0.0272 and +0.0492 (median +0.0432), significant at the 5 % level in 63 of the 64 re-estimations; the exception is Chapainawabganj (+0.0272, $p = 0.111$), the district that holds the most extreme district-year. Within the exposure

distribution, the slope comes from the most extreme district-years: excluding the single maximum cell (Chapainawabganj in 2012, 30 days at or above 40°C) lowers it to +0.0135 ($p = 0.458$) and excluding the top 1 % of cells (8 district-years with at least 18 such days) to +0.0165 ($p = 0.358$), while excluding the top 5 % (25 cells with at least 13 days) leaves +0.0318 ($p = 0.095$). These cell deletions are not robustness tests in the usual sense: dropping the most extreme cells removes the very exposure whose effect is being estimated, so a smaller estimate there shows where the effect is identified rather than that it is fragile. Online Appendix Table C.4 reports every division and summarizes the district leave-one-out distribution (range, median and significance counts). We carry the leave-one-district-out band into the projection of Section 8, where it moves each projected figure by about −37 to +14 % (Online Appendix Table H.1).

6 Additional Robustness

Table 5 collects the robustness checks of the joint cohort decomposition. Each row re-estimates the baseline of Table 3 with one change and reports what the two findings rest on: the 1980-cohort high-heat slope and its tests against each later cohort, and the 2010-cohort extreme-tail slope. Panel A varies the sample and the comparison group, including survey-round fixed effects, sampling weights, heat interactions with household characteristics (so that cohorts are compared at a fixed composition), a 10-year recall cap, and the 21 districts with $\geq$ 40°C days in every decade (so that every tail slope is identified off the same districts). Panel B re-times exposure—anchoring the window at expected delivery (conception +9 months), using conception-month fixed effects, controlling for heat just before birth, keeping nine-month gestations, and letting the in-utero heat bins vary by decade—and adds rainfall or humidity controls. Panel C varies the fixed effects, the district trends and the set of temperature bins interacted with birth decade; Panel D varies the inference.

Both results survive. Outside the 10-year recall cap (+0.0008), the 1980 high-heat slope lies between +0.0029 (additive fixed effects) and +0.0051 and is significant at the 5 % level in every row. Outside that row, the one-sided tests for a decline from the 1980 cohort reject at the 5 % level against the 1990 and 2010 cohorts in every row; against the 2000 cohort, the weakest contrast in the baseline ($p = 0.027$), they reject at the 10 % level in every row and at the 5 % level outside the two conception-anchored rows, the rainfall and humidity controls and two-way clustering (p between 0.062 and 0.089). The 2010 tail slope lies between +0.0414 and +0.0621 in every row, is significant at the 5 % level or better in all rows except three in which it is significant at the 10 % level (the 10-year recall cap, +0.0463,

and the Conley and wild-bootstrap inference rows of Panel D), and stays close to or above its baseline size in every row, so no check suggests that the tail effect has faded for the most recent cohort. Interacting every temperature bin with birth decade changes neither margin, and the added cold-bin × cohort interactions are jointly insignificant ($F = 1.41$, $p = 0.224$), so the two-margin interaction basis is not restrictive.

Who is observed. Every birth is reported retrospectively by a mother who survived to the interview and was ever married, through survey rounds that differ by cohort (Section 3.4), so the cohorts could differ in how they were sampled, in the mothers who report them, or in household composition rather than in how newborns respond to heat. We add survey-round fixed effects, so that cohorts are compared only within a round; control for the mother's age at birth; weight by the DHS sampling weights; and let every heat term vary with household wealth, maternal education, urban residence and household electricity, so that a change in who is exposed cannot masquerade as a change in the response. The 1980 high-heat slope is +0.0044 to +0.0048 across these rows, against +0.0045 in the baseline, and every one-sided test for a decline from it rejects at the 5 % level; the 2010 tail slope is +0.0433 to +0.0621, significant at the 5 % level or better in each case (Table 5, Panel A). Neither the survey design nor the composition of the sample drives the cohort comparison.

Within-infancy mortality harvesting. If ≥ 40°C exposure in the first 28 days pulls forward deaths that would have occurred later in infancy, it should lower mortality at days 29–365 among neonatal survivors. For recent cohorts it does, weakly: the slope is −0.0121 ($p = 0.088$) for the 2000 cohort and −0.0154 ($p = 0.064$) for the 2010 cohort, although the 2010 and 1990 slopes are not statistically distinguishable ($p = 0.188$, two-sided; Online Appendix Table D.1), so the offset is not clearly specific to recent cohorts. Partial displacement therefore cannot be ruled out. The point estimate of the 2010-cohort offset is about a third (36 %) of the 2010-cohort neonatal tail effect (0.0154 against 0.0432), so at the point estimates the net first-year effect of extreme heat remains positive. Displacement would change the timing of the extra deaths, not the persistence of the neonatal tail effect.

Fetal loss and selection into the live-birth sample. The DHS records only live births, so heat-induced fetal loss could change the composition of the live-birth pool. In the BDHS-2022 pregnancy recode, which records stillbirths, ten additional third-trimester days at or above 40°C raise the stillbirth probability by about 1.7 percentage points among 66,424 viable pregnancies (stillbirth rate 2.6 %), significant at the 10 % level ($p = 0.087$; Online Appendix

Table 5: **Robustness of the cohort decomposition.**

	High heat (30–40°C)					Extreme tail		
			One-sided p, $\beta_{1980} > \beta_c$			($\geq$ 40°C)		
	$\widehat{\beta}_{1980}$	(SE)	c=1990	2000	2010	$\widehat{\beta}_{2010}$	(SE)	N
Baseline (Table 3)	+0.0045**	(0.0018)	0.012	0.027	0.013	+0.0432***	(0.0161)	276,031
Panel A. Sample, selection and recall								
Survey-round FE	+0.0045**	(0.0018)	0.012	0.027	0.013	+0.0435***	(0.0162)	276,031
Mother's age at birth	+0.0044**	(0.0018)	0.014	0.032	0.015	+0.0433***	(0.0160)	276,031
DHS sampling weights	+0.0046**	(0.0019)	0.022	0.045	0.048	+0.0535**	(0.0203)	276,031
Heat × household traits	+0.0048**	(0.0018)	0.012	0.026	0.0097	+0.0621***	(0.0142)	276,031
Complete birth dates only	+0.0046***	(0.0017)	0.014	0.029	0.014	+0.0419**	(0.0167)	273,657
Recall $\leq$ 10 years	+0.0008	(0.0024)	0.263	0.551	0.455	+0.0463*	(0.0236)	147,471
$\geq$ 40°C days in every decade	—		—	—	—	+0.0437**	(0.0167)	73,223
Panel B. Exposure timing and weather controls								
Expected-delivery window	+0.0043**	(0.0019)	0.019	0.063	0.040	+0.0524***	(0.0162)	276,031
Conception-month FE	+0.0044**	(0.0019)	0.018	0.070	0.043	+0.0449***	(0.0162)	276,031
Pre-birth heat control	+0.0044**	(0.0018)	0.0098	0.019	0.012	+0.0488***	(0.0167)	276,031
9-month gestations only	+0.0048***	(0.0018)	0.012	0.019	0.008	+0.0464***	(0.0158)	274,172
In-utero heat × decade	+0.0051**	(0.0021)	0.021	0.039	0.011	+0.0414**	(0.0161)	276,031
Precipitation control	+0.0047**	(0.0018)	0.027	0.089	0.021	+0.0440**	(0.0166)	266,341
Relative-humidity control	+0.0045**	(0.0019)	0.023	0.062	0.024	+0.0458***	(0.0163)	269,407
Panel C. Fixed effects, trends and interaction basis								
District + season FE	+0.0029**	(0.0014)	0.015	0.024	0.014	+0.0433***	(0.0145)	276,031
District quadratic trends	+0.0045**	(0.0018)	0.008	0.022	0.012	+0.0474***	(0.0172)	276,031
All bins × birth decade	+0.0048**	(0.0019)	0.024	0.045	0.006	+0.0421**	(0.0159)	276,031
Panel D. Inference (baseline estimates)								
Two-way clustering	+0.0045**	(0.0021)	0.036	0.067	0.042	+0.0432***	(0.0122)	276,031
Conley spatial HAC	+0.0045***	(0.0017)	0.005	0.023	0.005	+0.0432*	(0.0249)	276,031
Wild cluster bootstrap	+0.0045**		0.015	0.033	0.013	+0.0432*		276,031

Notes: Each row re-estimates the baseline of Table 3 (linear probability model of neonatal death; 20-year recall sample of the 1980–2010 birth cohorts; 28-day-window 30–40°C and $\geq$ 40°C counts both interacted with birth decade) with the one change named in the row. Baseline fixed effects: district × birth calendar month and birth year. Baseline controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile, maternal education (none, primary, secondary, higher), child sex, urban residence, household electricity, 63 district-specific linear birth-year trends and in-utero heat bins (days in each temperature bin over the nine calendar months before the birth month). Standard errors clustered by district (64; 21 in the common-support row) in parentheses; $t(G-1)$. Columns, left to right: the 1980-cohort high-heat slope and its SE; one-sided p-values for a decline from it, $H_1 : \beta_{1980} > \beta_c$, against the 1990, 2000 and 2010 slopes, $t(G-1)$; the 2010-cohort extreme-tail slope and its SE; N. Entries are total cohort slopes per +10 days; "—" marks a statistic the row does not estimate. p-values that would round onto 0.010, 0.050 or 0.100 are shown to four decimals. Panel A rows add survey-round fixed effects; add six mother's-age-at-birth categories; weight by the DHS woman-level sampling weight; interact both heat margins with maternal education, wealth quintile, urban residence and electricity (each centered at its mean; "household traits"); keep births with reported birth month and year; keep births within 10 years of the interview; or keep the 21 districts with $\geq$ 40°C days in every decade (tail slope only). Panel B rows use district × conception-month and conception-year fixed effects with the 28-day window anchored at expected delivery (conception +9 months, or recorded gestation; "expected-delivery window") or, in the next row, at birth; add the temperature bins of the 30 days before birth ("pre-birth heat control"); drop births whose recorded gestation is not nine months; interact the in-utero 30–40 and $\geq$ 40°C margins with birth decade ("in-utero heat × decade"); or add Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) wet days ($\geq$ 1 mm) and total rainfall (births before 1981 drop) or mean daily relative humidity, each over the 28-day window. Panel C rows use additive district, birth-season and conception-season fixed effects instead of district × month ("district + season FE"); add district quadratic trends; or also interact the $<$ 20 and 20–25°C bins with birth decade ("all bins"). Panel D keeps the baseline estimates with two-way clustering (district, birth year; variance made positive semi-definite by the eigenvalue correction of Cameron et al., 2011), Conley spatial HAC standard errors (35 km, 3-year Bartlett lag; normal p-values for these two rows' stars and tests), or restricted wild cluster bootstrap p-values (Cameron et al., 2008; Webb weights, Webb, 2023; B = 9,999; one-sided for the tests of a decline), which that row's stars reflect. N falls where a row drops births. Full results: Panel A, Online Appendix Section D (Online Appendix Tables D.6, D.9 and D.11); Panel B, Online Appendix Section E (Online Appendix Tables E.2, E.3 and E.6), except the humidity row, Online Appendix Section C.2 (Online Appendix Table C.3); Panels C and D, Online Appendix Section F (Online Appendix Tables F.1, F.2, F.3 and F.4). Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table D.2). Such fetal loss removes the frailest fetuses before birth, so the live births we observe are positively selected and our neonatal estimates, including $\widehat{\beta}^{\geq 40}_{2010} = +0.043$, are if anything lower bounds.

Conception timing, fertility and the sex ratio. Heat could also change who is conceived, born and sampled: if couples in hotter districts shifted conception away from the hot season, if extreme heat lowered the number of births or of births that survive into the sample, or if it removed female fetuses selectively, the cohort slopes would be estimated on differently selected births. We test each. A March–May conception indicator does not respond differentially across decades to the district's ≥ 40°C climatology (joint $F = 0.23$, $p = 0.87$). We find no fall in births either. The joint test rejects ($F = 4.90$, $p = 0.004$), but only because sampled births in a district-year are higher in and after years with more ≥ 40°C days: the coefficients on current, once- and twice-lagged exposure are all positive and significant at the 1 % level. That is the opposite sign to a heat-induced loss of births; we do not read the association causally, since sampled births are an imperfect proxy for fertility. Relative to the 1980 cohort, a district one standard deviation hotter shifts the cohort male share at birth by at most about half a percentage point on a 51.2 % baseline; the joint test rejects ($F = 5.46$, $p = 0.002$), but the pattern is non-monotonic, and its only significant coefficient (1990: -0.0055, $p = 0.023$) has the sign opposite to heat-induced loss of female fetuses, so sex-differential fetal loss is unlikely to confound the results (Online Appendix Tables D.3, D.4 and D.5).

Recall. The 1980 cohort is observed mainly through long maternal recall. Recall error confounds the high-heat slope only if it is systematic with respect to heat exposure, which is unlikely: the identifying heat anomalies are measured from ERA5, not reported by the respondent, and misclassification of the outcome that is unrelated to heat exposure (for example, early deaths omitted at a constant rate) scales the linear-probability slope by a factor below one, toward zero (Hausman et al., 1998). Nor is there a sign of gross omission of early deaths: the neonatal mortality rate we recover for the 1980 cohort (births 1976–89; 73 per 1,000) lies above the UN IGME estimate for 1990 (about 66 per 1,000; World Health Organization, 2024b), as it should for births before 1990. Shorter recall caps remove the births that carry the high-heat result. Under a 15-year cap (218,679 births) the 1980 slope falls to +0.0038 ($p = 0.054$), but the one-sided tests for a decline still reject against the 1990 and 2010 cohorts at the 5 % level and against the 2000 cohort at the 10 % level ($p = 0.006$, 0.074 and 0.012); the 10-year cap drops every 1976–82 birth, leaving a 1980 cohort of about 16,400 births from 1983–89, and the 1980 slope is then small and insignificant (Table 5).

Restricting to the 1994–2004 rounds, which supply almost all 1980-cohort births at their shortest recall and contain no 2010s births, the 1980 slope is +0.0063 ($p = 0.020$). We read the weaker results under shorter caps as a loss of power rather than evidence of recall bias (Online Appendix Section D.5).

Measurement error. Early cohorts are reported at longer recall and, for the earliest years, from a reanalysis that assimilates fewer observations (Section 3), so their exposure and dates could be measured with more error. Classical error of this kind biases slopes toward zero, so noisier 1980-cohort data would work against finding the decline in the high-heat slope rather than produce it: the 1980 slope is the largest despite being the most exposed to such error. The tail claim rests on the pooled regression and on the most recent cohort, whose data are the best measured. From the 1980s on, ERA5 tracks the station record similarly well in every decade (correlation 0.90–0.91 in the 1980s through 2000s and 0.93 in the 2010s; mean absolute error 1.0–1.1°C), and its count of ≥ 40°C days relative to all matched stations is closest to one in the 2010s (Online Appendix Table D.10); dropping births with partly imputed birth dates changes neither result (Table 5).

Exposure timing. Both the exposure window and the month in the fixed effects are anchored on the date of birth. If heat in late pregnancy triggered early delivery, babies would be born earlier, into weather correlated with the heat that triggered delivery and with the higher mortality risk of preterm birth, so postnatal heat would partly proxy for heat-induced preterm birth and conditioning on birth month would condition on an outcome of heat. Anchoring exposure at expected delivery (conception +9 months) avoids this; we do so for all births (the window moves only where recorded gestation differs from nine months) and test the mechanism directly on the 42,852 births with recorded gestation in the regression (Section 3). Heat in the expected third trimester, dated from conception, does not raise the probability of preterm birth (+0.0005 per +10 days ≥ 40°C, SE 0.0088; base rate 4.9 %; Online Appendix Table E.1), and controlling for heat in the 30 days before birth, dropping births whose recorded gestation is not nine months or using conception-month fixed effects does not weaken the tail estimate (Panel B). The tail slope in the Panel B row "Exposure anchored at expected delivery" is the largest in that panel (+0.0524), but the window moves for only 1,859 births, and the increase comes mainly from a few preterm babies who died just before a hot spell being assigned heat that arrived after their death, so we do not read it as stronger evidence (Online Appendix Section E.1).

Gestational heat and heat at the time of death. The cohort decline could reflect a changing response to heat *in utero* rather than after birth, since the baseline holds the in-utero response fixed across cohorts; and an effect we call neonatal could be mis-timed if heat acts at other ages. On the first, letting the in-utero bins vary by decade barely moves the postnatal slopes: with both in-utero margins interacted with birth decade, the 1980 high-heat slope is +0.0051 and the one-sided tests for a decline from it to the 1990, 2000 and 2010 cohorts still reject ($p = 0.021$, 0.039 and 0.011), and the 2010 tail slope is +0.0414 ($p = 0.013$; Online Appendix Table E.3), so gestational heat does not carry the early-cohort result. With trimester-specific gestational bins, the postnatal $\geq$ 40°C level coefficient is +0.0200 ($p = 0.030$), no trimester's $\geq$ 40°C coefficient is significant at the 5 % level (the third-trimester one is negative, −0.0092, $p = 0.082$), and the three are jointly insignificant ($F = 1.73$, $p = 0.170$; Online Appendix Table E.4). On the second, in a child-month panel relating death in each month of age to heat in that month and the month before, extreme heat has no detectable effect on post-neonatal infant deaths (−0.14 deaths per 1,000 child-months per +10 days, SE 0.58, against 2.3 deaths per 1,000 child-months) or on second-year deaths (Online Appendix Table E.5; Online Appendix Section E.3), so the extreme-heat effect is neonatal also when exposure is measured at the time of death.

Specification. The cohort contrast could reflect differential secular change across districts that the baseline's 63 linear district trends do not absorb, or depend on the two temperature cutoffs that define the margins. On the first, adding district-specific quadratic trends to the baseline leaves the 1980 high-heat slope at +0.0045, with the one-sided tests for a decline from it rejecting at the 5 % level, and the 2010 tail slope at +0.0474 (Table 5, Panel C). Letting every temperature bin, not only the two margins, vary by birth decade changes neither result (+0.0048 and +0.0421), and the added cold-bin × cohort interactions are jointly insignificant ($F = 1.41$, $p = 0.224$), so the two-margin interaction basis is not restrictive. Nor do the results depend on the district × calendar-month structure itself: with additive district and season effects in its place, both results and their one-sided tests hold (1980 slope +0.0029, $p = 0.048$; 2010 tail slope +0.0433; Online Appendix Section F.1). The cutoffs are examined in Section 5.4: the tail result holds at thresholds of 40 to 42°C and the 1980 high-heat slope stays positive when the 10°C band is shifted.

Other weather and the districts at risk. Hot days could be deadly because they are humid or because they coincide with rainfall, and the districts with extreme heat change over time, so the cohort tail slopes could be identified from different places. Section 5.3 shows that $\geq$ 40°C days are dry and that controlling for postnatal relative humidity leaves

the 2010 tail slope at +0.0458 and the 1980 high-heat slope at +0.0045; adding district-day rainfall over the window gives +0.0440 and +0.0047 (Panel B). Restricting to the 21 districts with ≥ 40°C days in every decade, so that every tail slope is identified off the same districts, leaves the 2010 tail slope at +0.0437 ($p < 0.05$; Panel A).

Inference. With 64 districts, weather that is correlated across neighboring districts and within a year, and a tail effect identified from few district-years, district-clustered standard errors could overstate precision. We therefore re-run inference three ways. Two-way clustering by district and birth year (Cameron et al., 2011), which allows arbitrary within-year correlation across districts (with the variance made positive semi-definite by their eigenvalue correction), does not inflate the 2010 tail standard error (0.0122, against 0.0161 with district clustering) and raises the 1980 high-heat one only from 0.0018 to 0.0021. Conley spatial HAC standard errors (Conley, 1999) (35 km) inflate the tail standard error to 0.0249 ($p = 0.083$), and the restricted wild cluster bootstrap (Cameron et al., 2008; Webb, 2023), which guards against over-rejection with 64 clusters, gives $p = 0.094$: under both the 2010 tail slope is significant at the 10 % level rather than the 1 % level. Under none of the three does a one-sided test for a decline from an earlier cohort to the 2010 cohort reject (bootstrap $p = 0.880$, 0.940 and 0.491 against 1980, 1990 and 2000; smallest p across methods 0.486). The high-heat results are essentially unchanged: the bootstrap p is 0.018 for the 1980 slope, and the one-sided bootstrap tests for a decline from it give $p = 0.015$, 0.033 and 0.013 against the 1990, 2000 and 2010 cohorts (0.036, 0.067 and 0.042 under two-way clustering; Online Appendix Section F.4).

Functional form and cohort resolution. Neonatal death is a rare outcome, so a linear probability model could misstate effects that are proportional to the base rate, and decade bins could hide movement within a decade. Poisson pseudo-maximum likelihood (PPML), which drops the 266 births in fixed-effect cells with no neonatal death, gives the same cohort patterns: at the pooled-sample base rate the 2010 tail effect is +47 deaths per 1,000 live births per +10 days ($p < 0.001$; linear +43), and the high-heat slope is again largest and significant only for the 1980 cohort (+3.2 deaths per 1,000 per +10 days, $p = 0.039$; linear +4.5), although the one-sided tests for a decline from it are weaker ($p = 0.014$, 0.107 and 0.046 against the 1990, 2000 and 2010 cohorts; Online Appendix Table F.5). At seven 5-year cohort bins the 2010-cohort tail effect (+0.0433) and the early-cohort high-heat slope, localized to 1980–84, both survive (Online Appendix Table F.6).

7 A Candidate Mechanism for High-Heat Adaptation: Rural Electrification

This section probes the mechanism behind the adaptation to the 30 to 40°C range; the mechanism test is not informative at the extreme tail (Section 7.1). The most plausible observable driver is the sharp rise in access to electricity over our sample period, which we test directly using the staggered rollout of the REB program. Electrification is a broad treatment, however. Active convective cooling from household fans is the most direct candidate—fan ownership rose from 52 % to 96 % of households between BDHS-2011 and BDHS-2022, while air conditioning, which cools regardless of the heat–humidity environment, remains negligible at 1.4 % (Online Appendix Section G.5)—but electrification also travels with refrigeration for food and medicine, energy that supports hospital services and health-seeking, treated water, and broader housing and infrastructure upgrades. We cannot distinguish these channels, so the test is deliberately agnostic about which does the work: it asks whether the *timing* of a district's electrification, rather than any single technology's adoption, moves the neonatal heat-mortality slope.

Connection dates. REB delivers electricity through Palli Bidyut Samities (PBSs; singular Palli Bidyut Samity, PBS), district-level rural electric cooperatives that it sets up and energizes one at a time. We date a district's program connection by the first energization (start of commercial operation) of the district's own PBS, taken from that PBS's official "at a glance" fact sheet (Palli Bidyut Samities, 2026) or, where the fact sheet gives no energization date, from the commercial-operation dates in the Management Information System (MIS) report of the Bangladesh Rural Electrification Board (BREB), as REB is now called (Bangladesh Rural Electrification Board, 2026). Every date is an energization date, never a registration or founding date, documented with a citation and a verbatim quote and independently re-checked (for seven districts the MIS report is the only source). Where a source documents that the district's area was served earlier by a predecessor PBS, we use that PBS's date. Narayanganj was served by Narsingdi PBS-1 from 1986; Rajshahi's Charghat upazila (sub-district) lay in the original service area of the PBS now called Natore PBS-2, energized in 1981; and REB's thana list from before Gazipur PBS existed (Rural Electrification Board, 2004) assigns Gazipur's upazilas to Dhaka PBS-1, energized in 1980, and Mymensingh PBS-2, energized in 1996. Where two predecessors served parts of a district, we use the earlier. Gazipur PBS-1's own date, 1 July 2005 (the start of a fiscal year), is the administrative creation of a new PBS in an area these PBSs already served. On-

line Appendix Table G.3 reports the 2005 date as a robustness check. Districts without a PBS of their own take the date of the earliest documented PBS serving part of the district: Barguna, Pirojpur PBS (1986), which by 2004 served its Bamna and Patharghata upazilas (Rural Electrification Board, 2004), Patuakhali PBS serving the rest from 1992; Narail, Jashore PBS-2; Chuadanga, Meherpur PBS; Panchagarh, Thakurgaon PBS. The three Chittagong Hill Tracts districts lie outside REB's district service area and drop out, leaving 61 districts, with connection years (calendar year) from 1980 to 2001.

Design, identifying assumption, and threat. The staggered REB rollout provides plausibly exogenous timing variation in district-level electrification access conditional on the fixed effects. We define $\text{post}_{dt} = \mathbf{1}[\text{birth year} \geq \text{district } d \text{ program-connection year}]$ and interact it with the high-heat (30–40°C) and tail ($\geq$ 40°C) heat-day counts in the standard postnatal window. The regression otherwise follows the baseline specification: district $\times$ birth calendar-month and birth-year fixed effects, the $<$ 20 and 20–25°C postnatal bins, categorical controls for birth order, wealth quintile, and maternal education, indicators for male and urban, 60 district-specific linear trends (one for each of the 61 panel districts but the base district) and the in-utero heat bins, with standard errors clustered by district. The one departure is that it omits the household electricity control, which is an outcome of the program and hence a bad control here. With 99 % of 1980–2010 cohort births in the 61 districts with a connection date, and essentially no never-connected comparison group, the design is identified off the staggered timing of connection.

The estimand is a reduced-form, intent-to-treat object: the effect of the timing of a district's program connection on the neonatal heat-mortality slope, not the effect of household electrification or fan ownership, which the BDHS records only as a survey-time snapshot, with no year of acquisition, and which we therefore cannot use in a household first stage (Online Appendix Section G.5). Identification requires that, conditional on the fixed effects and controls, the timing of a district's REB connection be orthogonal to within-district trends in the unobserved determinants of the neonatal heat-mortality slope. The leading threat is that districts were connected in a sequence plausibly correlated with development (market access, road and health infrastructure, baseline income), so program timing could proxy a broader "development arrived" bundle rather than electrification itself. We address it with three tests: a rural/urban placebo exploiting the program's rural design (the large cities were electrified before it), a pre-trend test on the heat slope, and a development-indicator placebo that asks whether program timing moves other markers of rural development.

Estimators. With staggered connection and effects that may differ across connection cohorts (connection-year groups, not birth decades) and with time since connection, the conventional two-way fixed-effects (TWFE) estimate of the post × heat interaction partly compares later-connected with already-connected districts, the source of its potentially negative weights (Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2020). Table 6 therefore reports TWFE (column 1) next to a heterogeneity-robust cohort-pooled interaction-weighted (IW) estimator in the spirit of Sun and Abraham (2021) (column 2), which avoids such comparisons. It interacts a single post-connection indicator and its heat interactions with four connection-cohort bins (connected by 1985, 1986–1990, 1991–1995, and 1996 or later) and averages the four cohort-specific slope changes with weights equal to each bin's share of post-connection births. Because the post-connection births of every connection cohort enter their own interaction terms, the baseline heat slope is estimated from not-yet-connected births only, so treated births never serve as comparisons. The assumption is that, without the program, the heat slope would not have drifted over time; the birth-decade × heat check below relaxes it. It uses all birth years, 1976–2019. It is not Sun and Abraham's exact estimator: it pools the post-connection years within a connection-cohort bin before aggregating and uses no last-treated control cohort. As a robustness check, Online Appendix Table G.1 reports the imputation estimator of Borusyak et al. (2024) in its exact two-step form: the untreated model, with heat slopes varying by birth year and by connection cohort, is fitted on not-yet-connected births only, the slope changes are estimated from connected births' deviations from their imputed untreated outcomes, and standard errors come from a leave-one-district-out jackknife over both steps. Because every district is connected by 2001, it uses births through 2000.

Both estimates point to a flattening of the high-heat slope at connection: −0.0039 under TWFE ($p = 0.004$) and −0.0058 under the cohort-pooled IW estimator ($p < 0.001$), which drops the comparisons with already-connected districts and is larger in magnitude. The imputation estimate is similar in size to the IW estimate (−0.0055) but imprecise (jackknife SE 0.0040, $p = 0.179$), and its joint test of the four high-heat leads two to five years before connection, run on not-yet-connected births, does not reject ($p = 0.183$; Online Appendix Table G.1). These estimates, about 4 to 6 deaths per 1,000 live births per +10 days at 30–40°C, are larger than the cross-cohort decline in the high-heat slope (0.0026–0.0036 per +10 days; Section 5), and the IW estimate is close to, and slightly above, the 1980-cohort slope itself (+0.0045). That slope is a lower bound on the pre-electrification slope rather than a ceiling on the flattening: connections began in 1980 and the 1980 cohort spans births in 1976–89, so part of that cohort is already post-connection and +0.0045 understates the slope that electrification acted on. We nonetheless read the estimates as evidence on the

Table 6: **REB program × heat: TWFE and cohort-pooled IW estimates.**

	(1) TWFE	(2) Cohort-pooled IW
Post × high heat (30–40°C)	−0.0039***	−0.0058***
	(0.0013)	(0.0016)
Post × extreme heat (≥ 40°C)	+0.0073	+0.0107
	(0.0250)	(0.0259)
Birth years	1976–2019	1976–2019
N	273,520	273,520
Districts	61	61

Notes: Change in the neonatal heat slopes at a district's REB connection. Linear probability models of neonatal death, one observation per birth; 20-year recall sample of the 1980–2010 birth cohorts (births 1976–2019) in the 61 districts with an REB connection date (connection years 1980–2001; the three Chittagong Hill Tracts districts have none). High heat: 30–40°C, extreme heat: ≥ 40°C day counts over the 28-day neonatal window, per +10 days (< 20 and 20–25°C counts as controls). Post = 1 if the birth year is at or after the district's connection year. (1) TWFE: post and post × each heat margin; already-connected births also serve as comparisons. (2) Cohort-pooled interaction-weighted (IW) estimator, in the spirit of Sun and Abraham (2021): post and post × each heat margin interacted with four connection-cohort bins (connection year ≤ 1985, 1986–90, 1991–95, ≥ 1996), averaged with each bin's share of post-connection births; baseline heat slopes from not-yet-connected births only. Fixed effects: district × birth calendar month and birth year. Controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile, maternal education (none, primary, secondary, higher), child sex, urban residence, 60 district-specific linear birth-year trends and in-utero heat bins (days in each temperature bin over the nine calendar months before the birth month); household electricity is excluded as an outcome of the program. Pre-trend test: Figure 4; imputation-estimator check: Online Appendix Table G.1. Standard errors clustered by district (61) in parentheses; $t(60)$. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

direction and timing of the flattening rather than on its exact size.

Neither estimator says anything reliable about the ≥ 40°C margin. The TWFE and cohort-pooled IW tail estimates are +0.0073 and +0.0107, neither significant, with standard errors (0.0250 and 0.0259) more than an order of magnitude larger than at high heat, because few births in the estimable cells have any such days; the imputation tail estimate (+0.0700, jackknife SE 0.150) is less informative still. We return to this in Section 7.1.

The high-heat result does not hinge on how the connection dates are defined. Online Appendix Table G.3 re-estimates the TWFE and cohort-pooled IW regressions and the event study under the baseline dates and five alternatives: Gazipur dated by its own PBS's 2005 creation; each district's own PBS only; the earliest energization of any PBS serving part of the district today, from the fact sheets; the same from BREB's MIS dates; and an earlier connection file compiled for this project. Across these six definitions the cohort-pooled IW high-heat estimate ranges from −0.0050 to −0.0058 and is significant at the 1 % level under every one, and neither pre-program event-study coefficient is significant at the 10 % level under any of them. Nor does the result reflect a general decline in the heat slopes across birth cohorts that happens to coincide with connection, the drift that the cohort-pooled IW estimator assumes away: adding birth-decade × heat interactions, which let both heat slopes change across birth cohorts for reasons unrelated to the program, leaves its high-heat

estimate at −0.0048 (SE 0.0018, $p = 0.011$; Online Appendix Table G.4).

Rural/urban placebo. REB is, by design, a *rural* electrification program, and the large cities had grid power well before it. If the flattening at connection reflects electrification rather than an unrelated secular trend, it should be absent among births in these cities. We therefore divide births by the mother's place of residence at the interview into rural areas, towns (all other urban areas), and the large cities (the city corporations and statistical metropolitan areas of the nine city-corporation districts in the sample). We estimate one pooled regression in which each group has its own baseline heat slopes and its own post-connection changes, so the differences between groups are contrasts within a single regression (Table 7). Large-city births show no post-connection change in the high-heat slope (+0.0023, $p = 0.62$, under the cohort-pooled IW estimator; +0.0034, $p = 0.42$, under TWFE), while rural births (−0.0055, $p = 0.003$; TWFE −0.0036, $p = 0.021$) and births in towns (−0.0081, $p = 0.012$; TWFE −0.0065, $p = 0.039$) flatten. That the towns flatten is not at odds with the placebo: many towns and municipalities lie inside PBS service areas (the PBS fact sheets report the municipalities, or *pourashavas*, within each PBS area) and gained electricity later than the large cities, which is also why pooling towns with cities gives urban flattening about as large as rural (Online Appendix Table F.1, Panel D, baseline columns). Taking rural and town births together as the PBS areas, the high-heat slope flattens by 0.0061 ($p < 0.001$; TWFE 0.0043, $p = 0.002$). The differences between the cities and the rest have the predicted sign but are imprecise: city minus rural is +0.0078 (SE 0.0047, $p = 0.11$) under the cohort-pooled IW estimator and +0.0070 (SE 0.0045, $p = 0.12$) under TWFE, and city minus PBS areas is +0.0084 (SE 0.0046, $p = 0.075$) and +0.0077 (SE 0.0044, $p = 0.083$), significant only at the 10 % level. The city estimates rest on nine district clusters, and residence is measured at the interview, up to 20 years after the birth, so the placebo is consistent with the electrification interpretation but does not on its own establish it. The cities have few days at or above 40°C, so we do not interpret the significant city coefficients at the tail.

Pre-trend test. Figure 4 reports the high-heat 5-year-bin event-study coefficients. We define τ = birth year − district program-connection year and assign each child to one of five 5-year bins: two pre-program bins, the reference bin covering the five years just before connection, and two post-program bins. This conventional TWFE check accompanies the TWFE and cohort-pooled IW estimates; the imputation estimator's own placebo is in Online Appendix Table G.1. Over the 15 years before connection, neither pre-program coefficient ($\tau \in [-15, -6]$, relative to $[-5, -1]$) is significant (−0.0036, $p = 0.304$, and −0.0020, $p = 0.486$). Both are negative, so if anything the high-heat slope was lower, not higher, in the

Table 7: **Rural/urban placebo: REB program × heat by mother's residence.**

	Rural	Towns	PBS areas (rural + towns)	Large cities	Cities − rural	Cities − PBS areas
Panel A. TWFE: post-connection change in the heat slope						
High heat (30–40°C)	−0.0036**	−0.0065**	−0.0043***	+0.0034	+0.0070	+0.0077*
	(0.0015)	(0.0031)	(0.0013)	(0.0042)	(0.0045)	(0.0044)
Extreme heat (≥ 40°C)	−0.0012	+0.0280	+0.0056	+0.0456	+0.0468*	+0.0400*
	(0.0288)	(0.0218)	(0.0248)	(0.0280)	(0.0261)	(0.0218)
Panel B. Cohort-pooled IW: post-connection change in the heat slope						
High heat (30–40°C)	−0.0055***	−0.0081**	−0.0061***	+0.0023	+0.0078	+0.0084*
	(0.0018)	(0.0031)	(0.0016)	(0.0047)	(0.0047)	(0.0046)
Extreme heat (≥ 40°C)	+0.0027	+0.0263	+0.0082	+0.1219***	+0.1192***	+0.1137***
	(0.0325)	(0.0588)	(0.0254)	(0.0269)	(0.0418)	(0.0364)
Births	195,384	54,683	250,067	23,453		
Districts	61	61	61	9		

Notes: Change in the neonatal heat slopes at a district's REB connection, by the mother's place of residence at the interview. Linear probability models of neonatal death, one observation per birth; 20-year recall sample of the 1980–2010 birth cohorts in the 61 districts with an REB connection date. Heat margins: 30–40 and ≥ 40°C day counts over the 28-day neonatal window, per +10 days (< 20 and 20–25°C counts as controls). Post = 1 if the birth year is at or after the district's connection year. Residence, from the mother's DHS place of residence at the interview (the round-by-round variable coding is in Online Appendix Section A.1): *rural* is DHS rural residence; *large cities* are urban births in the city corporations and statistical metropolitan areas of the nine city-corporation districts in the sample (Dhaka, Chattogram, Khulna, Rajshahi, Sylhet, Barishal, Rangpur, Narayanganj and Gazipur); *towns* are every other urban birth (towns and municipalities, many of them inside PBS service areas); *PBS areas* pools the rural and town births. For each estimator, one regression on all births gives each residence group its own level, its own baseline high- and extreme-heat slopes, and its own post-connection changes in the level and in both heat slopes; the entries are these post-connection slope changes. TWFE: post × group and post × heat × group. Cohort-pooled IW estimator, in the spirit of Sun and Abraham (2021): these terms interacted with the four connection-cohort bins, each group's bin coefficients averaged with that group's shares of post-connection births. The PBS-areas entry is the average of the rural and town changes weighted by their numbers of post-connection births. "Cities − rural" and "Cities − PBS areas" are the large-city change minus the rural change and minus the PBS-areas change, linear combinations of coefficients from the same regression. Fixed effects: district × birth calendar month and birth year. Controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile, maternal education (none, primary, secondary, higher), child sex, 60 district-specific linear birth-year trends and in-utero heat bins (days in each temperature bin over the nine calendar months before the birth month); urban residence enters through the residence-group indicators, and household electricity is excluded as an outcome of the program. Births and districts: births in each residence group and the districts in which they occur; the regressions use the $N = 273{,}520$ births with complete covariates in 61 districts. The large-city estimates rest on 9 district clusters, and few large-city births have any ≥ 40°C days, so the extreme-heat rows are not informative about the cities. Standard errors clustered by district (61) in parentheses; $t(60)$. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

earlier pre-program years; a pre-existing decline that the post-program coefficients merely continued would instead show up as positive pre-program coefficients. The post-program estimates show the flattening appearing in the first 5-year window after connection (−0.0047, $p = 0.020$) and persisting into the second (−0.0075, $p < 0.001$). All coefficients from this regression, including the extreme-tail margin the figure omits, are tabulated in Online Appendix Table G.2.

Development-indicator placebo. If REB rollout order proxied a broad "development arrived" bundle rather than electrification itself, connection should coincide with gains in *other* markers of rural development, not electrification alone. We test this on the two de-

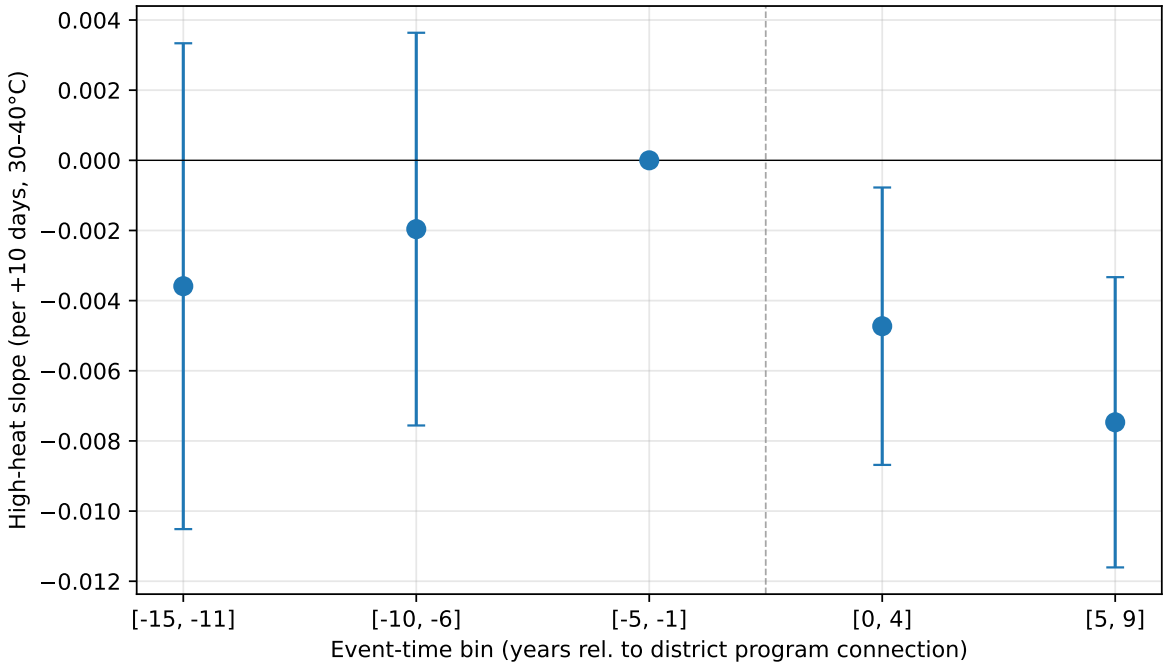


Figure 4: **Pre-trend test for the REB program effects.** Coefficients on the high-heat (30–40°C) slope × event-time-bin interactions relative to the $\tau \in [-5, -1]$ reference bin (just before program connection), per +10 days in the 28-day neonatal window from the birth date, with 95 % confidence intervals from district-clustered standard errors (61 districts) and $t(60)$ critical values. One linear probability model of neonatal death on the 20-year recall sample of the 1980–2010 birth cohorts in the 61 REB districts, births with τ = birth year − connection year $\in [-15, 9]$ (N = 118,454); both heat margins enter as levels and interacted with the four non-reference event-time bins. Fixed effects: district × birth calendar month and birth year. Controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile, maternal education (none, primary, secondary, higher), child sex, urban residence, 60 district-specific linear birth-year trends and in-utero heat bins (days in each temperature bin over the nine calendar months before the birth month); household electricity is excluded as an outcome of the program. The dashed line marks the pre/post threshold; the two pre-program coefficients are the pre-trend test.

velopment indicators the BDHS records at the household level: the (within-round) wealth quintile and maternal education (Table 8). Restricting to the births of rural households, we collapse each indicator to a district-round mean, standardize it, and regress it on the REB `post` indicator, set to one if the district was connected by the median calendar year of interview in that district-round cell, with district and survey-round fixed effects. Neither indicator shows a significant change at connection (wealth quintile −0.320 standard deviations, $p = 0.154$; maternal education −0.083, $p = 0.548$); the wealth estimate is not small, but it is negative and imprecise, pointing, if anything, away from a development gain. The test has limited power: 42 of the 61 districts were connected by the first survey round (1993–

94), so only 33 of the 545 district-round cells precede connection and the `post` coefficient is identified from the 19 districts connected during the survey period. The confidence intervals rule out development gains larger than about 0.12 standard deviations in wealth and 0.19 in maternal education, but not smaller ones. Within these limits, we find no evidence that REB-connected rural districts gained differentially in development.

Table 8: **Placebo: other markers of rural development.**

	Wealth quintile (within-round)	Maternal education
Post (REB connected)	−0.320	−0.083
	(0.222)	(0.138)
p-value	0.154	0.548
Rural mean (pre → post)	2.67 → 2.63	0.50 → 0.84
N (district-rounds)	545	545

Notes: Source: BDHS births of rural households (nine rounds, 1993–94 to 2022) in the 61 districts with an REB connection date (the three Chittagong Hill Tracts districts have none). Each column is a development indicator (the dependent variable), the household wealth quintile (within round, 1–5) or maternal education (0 none to 3 higher), averaged over the rural birth records of each district-round cell and z-standardized across the 545 district-round cells; one observation per district-round. `post` = 1 if the district's REB connection year is at or before the median calendar year of interview in the district-round cell. Estimates are from indicator ∼ `post` + district fixed effects + round fixed effects, with no other controls, so the "Post (REB connected)" coefficient $\widehat{\beta}$ is the within-district jump in the standardized development indicator at connection, in standard deviations. "Rural mean (pre → post)" is the mean of the unstandardized indicator over cells before and after connection. Standard errors clustered by district (61 districts) in parentheses; p-values are two-sided tests of $H_0 : \beta = 0$ from $t(G-1)$, $G = 61$. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

7.1 What the Mechanism Test Can and Cannot Say About the Extreme Tail

The mechanism test is informative only at high heat. Why the cohort evidence shows adaptation at high heat but not at the extreme is a separate question, for which we have no direct test; we set out what we can and cannot say.

The obvious candidate is a capacity limit of the cooling technology that electrification actually delivers. Fans are near-universal, while air conditioning reaches only 4.0 % of urban and 0.4 % of rural households (NIPORT and ICF, 2024). If fans relieve heat strain up to some threshold and only refrigerated cooling helps above it, an electrification program that diffuses fans would flatten the 30–40°C margin and leave the extreme untouched. The penetration gap is a fact in our data, and it is large enough that any tail protection operating through air conditioning would be invisible in this population.

In a randomized crossover trial, Morris et al. (2019) find that fan use *worsens* core temperature and cardiovascular strain in very hot, dry conditions (47°C, 10 % relative humidity) but *improves* both in hot, humid conditions (40°C, 50 % relative humidity). They suggest that fans could be encouraged where temperatures rarely exceed 40°C and relative humidity is moderately high (above 30–40 %). Bangladesh's $\geq$ 40°C days average 49 % relative humidity (Section 5.3), close to the trial's hot, humid condition, in which fans helped; the trial, run on 12 healthy young men at rest, does not test hotter days at that humidity, or infants. Taken at face value, it does not support a fan-capacity explanation at Bangladesh's $\geq$ 40°C humidity levels.

Barreca et al. (2016) show that residential air conditioning explains essentially the entire post-1960 decline in US hot-day mortality, in a test at the extreme: days with a daily mean above 90°F, whose daily maximum averaged 106°F (about 41°C), close to our $\geq$ 40°C bin. They also find that residential electrification, identified mainly from rural variation over 1929–1959 and so acting through the technology then available, was not statistically related to reductions in heat-related mortality, and in developing-country DHS samples household electricity appears protective against cold but not hot days, though with wide confidence intervals (Geruso and Spears, 2018). If the channel is cooling, the relevant comparison is not electricity access as such but the cooling it enables, and the cooling device electricity has brought to Bangladeshi households is the fan, not air conditioning. Read together, the evidence is consistent with electricity-enabled fan cooling flattening the 30–40°C slope, but it leaves open whether air conditioning, rather than fans, would be needed above 40°C, and our data cannot establish the channel.

The most important caveat is that the tail estimate is too imprecise to carry the contrast between the two margins. The cohort-pooled IW tail coefficient is +0.0107 with a standard error of 0.0259, so its 95 % confidence interval runs from −0.041 to +0.062, and the TWFE and imputation estimates are no more informative (Table 6 and Online Appendix Table G.1). The point estimate is positive, opposite to flattening, but the interval contains zero and much else besides, including a flattening of −0.041, which would offset almost the entire 2010-cohort tail slope (+0.0432). The data therefore cannot distinguish no program effect at the tail from one nearly as large, in proportional terms, as the effect we find at high heat, and the tail null is not evidence that electrification failed at the extreme.

Other explanations remain open and we cannot separate them. Surviving an extreme-heat episode may depend on acute clinical care rather than on household cooling, and the REB program was not designed to expand acute clinical capacity, although electrification may have improved facility services. Distinguishing these would require exposure-time appliance

ownership, which the BDHS does not record, or a setting with independent variation in air-conditioning access. We therefore report the asymmetry as an empirical regularity and leave its explanation open (Section 9).

8 Neonatal Mortality Costs: Historical and Projected

We translate the estimates into neonatal mortality costs in three steps: a *historical* accounting of the deaths averted by the REB rollout; the deaths attributable to extreme-tail (at or above 40°C) heat in the 2010s, using the 2010-cohort slope, against a zero-effect counterfactual; and a *mid-century* projection that combines the 2010-cohort tail coefficient with the ensemble-median ≥ 40°C day counts for 2041–2060 that we compute from the bias-corrected CMIP6 data of Mishra et al. (2020) for 13 general circulation models (GCMs), under different assumptions about AC adoption. Table 9 reports the results; the formulas, caveats, and VSL, AC-effectiveness, climate-model and tail-slope sensitivity are in Online Appendix Section H.

8.1 Historical Accounting

We compute the neonatal mortality *averted* by the REB rollout from the change in the high-heat (30 to 40°C) slope at connection estimated by the cohort-pooled IW estimator, -0.0058 (column 2 of Table 6). We apply this estimate because it is heterogeneity-robust and uses all birth years, 1976–2019, matching the accounting horizon; other estimates rescale the totals proportionally (below). The slope change is applied to realized decade-mean exposure in days at or above 30°C, of which $\geq$ 40°C days are under 0.5 %, and scaled by the unweighted share of the 64 districts connected to the program in each year of the decade (the three Chittagong Hill Tracts districts, which have no connection date, count as unconnected). This implies approximately 718,000 neonatal deaths averted cumulatively over the 1990s–2010s (95 % CI 323,000 to 1,112,000; ∼18,300/yr in the 1990s, ∼26,000/yr in the 2000s, ∼27,600/yr in the 2010s), or Int$143.5 billion in mortality costs at a central VSL of Int$200,000 for Bangladesh. All dollar values are in 2015 international (purchasing-power-parity) dollars: the VSL applies the preferred benefit-transfer rule of Robinson et al. (2019), a US VSL scaled by the ratio of Bangladesh's to US gross national income (GNI) per capita, both in 2015 international dollars, with an income elasticity of 1.5 (derivation in Online Appendix Section H.1). The slope change is close to, and slightly above, the 1980-cohort high-heat slope (+0.0045; Table 3). That slope is not the pre-electrification slope, however: connections began in 1980 and the 1980 cohort spans births in 1976–89, so part of that cohort

is already post-connection and +0.0045 understates the slope that electrification acted on. The comparison therefore does not bound the flattening. In the 2010s the point figure (∼27,600/yr) is nearly a third of the ∼90,000 annual neonatal deaths implied by the 2010-cohort base rate, which is why we read these magnitudes as evidence on the direction and timing of the flattening rather than its exact size. Holding births at 3.0 million is conservative for the 1990s and 2000s, when they averaged 3.65 and 3.59 million (United Nations, 2022). The figure is proportional to the slope change, so TWFE (−0.0039) would scale it by about two-thirds and the imputation estimate (−0.0055; Online Appendix Table G.1) by about 0.95. Since the REB rollout plausibly bundles other correlated rural improvements (roads, income, health infrastructure), we interpret this as the mortality averted by the *program* as a whole, of which cooling access is one possible channel.

The 2010-cohort extreme-tail effect ($\widehat{\beta}^{\geq 40}_{2010} = +0.0432$) applies to a realized exposure distribution that is sparse nationally but concentrated in the northwest, with a secondary cluster in the southwest: 71 of the 640 district-years of the 2010s (11 %) and 24 of 64 districts had any ≥ 40°C day; the median district-year has none, the 99th percentile ∼12 days, and the maximum 30 (Chapainawabganj, 2012), against 18 for the second-highest cell. Applied to this distribution against a zero-effect counterfactual, the estimate implies roughly 5,340 neonatal deaths attributable to extreme heat over the decade, an aggregate mortality cost of ∼Int$1.07 billion at the central VSL. If part of the effect brings forward deaths from later infancy (Section 6), these attributable deaths, and the projected deaths below, could be overstated by roughly a third. The magnitude is modest because ≥ 40°C exposure remains rare today; the policy-relevant quantity is what the same per-day effect implies for the much larger projected exposure at mid-century.

8.2 Mid-Century Projections

For 2041–2060, we combine the 2010 tail slope with the ensemble-median additional days at ≥ 40°C for Bangladesh under three Shared Socioeconomic Pathways (SSPs), computed from the 13-GCM bias-corrected CMIP6 data of Mishra et al. (2020): +1.5 days under SSP1-2.6, +1.7 under SSP2-4.5, and +2.4 under SSP5-8.5, relative to 1995–2014. Because the bias-corrected data have more such days than ERA5 over 1995–2014 (1.11 against 0.81 per year) and our slope is estimated per ERA5 day, we carry the change over in proportion to the bias-corrected baseline, which gives 1.94, 2.03 and 2.57 ERA5 days per year at ≥ 40°C under the three pathways, against 0.81 over 1995–2014. Air conditioning attenuates the slope linearly with household penetration at a per-household slope-reduction factor $\rho = 0.75$, a calibration motivated by the ∼75 % decline in the US heat-mortality response over the twentieth century,

Table 9: **Mortality cost: historical accounting and mid-century projections.**

	Annual deaths	Decade total	Int$ bn (2015 PPP) (VSL $200K)
Panel A. Neonatal deaths averted by the REB rollout (cohort-pooled IW slope change)			
1990s	18,258	182,577	36.5
2000s	25,953	259,529	51.9
2010s	27,557	275,572	55.1
Cumulative 1990s–2010s **(point)**	—	**717,678**	**143.5**
95 % CI		[323,124, 1,112,232]	[64.6, 222.4]
Panel B. Deaths attributable to $\geq 40\,°C$ *heat,* 2010 *cohort (zero counterfactual)*			
2010s	534	5,342	1.07
Panel C. Mid-century projection, 2041–2060 *average (annual deaths, Int$ bn, 2015 PPP)*	Current AC (5 %)	Moderate AC (30 %)	Ambitious AC (75 %)
SSP1-2.6 (optimistic)	1,445 ($0.29)	1,164 ($0.23)	657 ($0.13)
SSP2-4.5 (median)	1,517 ($0.30)	1,221 ($0.24)	689 ($0.14)
SSP5-8.5 (pessimistic)	1,920 ($0.38)	1,546 ($0.31)	873 ($0.17)

Notes: Calculated neonatal deaths (death in the first month of life), not regression estimates. "Annual deaths" is the decade-average number of deaths per year, "Decade total" its ten-year sum, and "Int$ bn (2015 PPP)" the decade total valued at the VSL in billions of 2015 international (purchasing-power-parity) dollars (Panel C: annual deaths and, in parentheses, their annual value in the same units). The slopes applied come from linear probability models of neonatal death on the 20-year recall sample of the 1980–2010 birth cohorts (fixed effects: district × birth calendar month and birth year; controls: birth-order group (1, 2, 3, 4–5, 6+), household wealth quintile, maternal education (none, primary, secondary, higher), child sex, urban residence, district-specific linear birth-year trends and in-utero heat bins (days in each temperature bin over the nine calendar months before the birth month); standard errors clustered by district): Panel A uses the REB regression of Table 6 (61 districts, 60 trends, no household electricity), Panels B and C the joint cohort decomposition of Table 3 (63 trends, plus household electricity). Panel A applies the cohort-pooled interaction-weighted (IW) high-heat slope change at program connection (−0.0058 per +10 days, i.e., −0.00058 per day) to the realized ERA5 decade-mean ≥ 30°C exposure in each decade, multiplied by 28/365 (neonatal window share) × 3,000,000 annual births and by the unweighted share of the 64 districts connected to the program, averaged over the years of the decade; the 95 % CI propagates its standard error. Panel B applies the 2010-cohort extreme-tail slope (+0.0432 per +10 days, i.e., +0.0043 per day) to realized 2010s ≥ 40°C exposure (0.54 days per district-year), multiplied by 28/365 × 3,000,000 annual births (zero-effect counterfactual). Panel C applies the same per-day slope to projected 2041–2060 exposure in ERA5 units: the ERA5 1995–2014 mean (0.81 days per district-year) scaled by $1 + \Delta/B$, where Δ is the ensemble-median change in ≥ 40°C days from 1995–2014 to 2041–2060 and B (1.11 days) the ensemble-median 1995–2014 baseline, from the bias-corrected projections of Mishra et al. (2020); multiplied by 28/365 × 2,340,000 annual births (UN medium variant, 2041–2060 mean; United Nations, 2022), with the slope reduced by $\rho = 0.75$ times the AC penetration share (sensitivity in Online Appendix Section H.2). Dollar values use the central VSL of Int$200,000 in 2015 international dollars: option (a) of Robinson et al. (2019), a US VSL of $9.4 million scaled by $(4{,}440/57{,}900)^{1.5}$, the Bangladesh-to-US ratio of 2015 GNI per capita in 2015 international dollars (World Bank, 2026) raised to an income elasticity of 1.5.

essentially all of which Barreca et al. (2016) attribute to residential AC and by the odds ratio of 0.3 for heat-related death with a working air conditioner in Semenza et al. (1996) (Online Appendix Section H.2).

AC penetration matters more than the emissions pathway. The "current-AC" scenario sets household penetration at 5 %, allowing for some diffusion since the 1.4 % measured in BDHS-2022 (NIPORT and ICF, 2024), which puts that column about 3 % below what measured coverage would imply. The current-AC column is not an upper bound: the climate-model spread alone carries the projection well above it. Holding the SSP pathway fixed and

moving AC penetration from current (5 %) to ambitious (75 %) roughly halves projected neonatal deaths in every SSP scenario: 1,445 → 657 under SSP1-2.6, 1,517 → 689 under SSP2-4.5, and 1,920 → 873 under SSP5-8.5. Holding AC fixed and varying the SSP pathway moves the projection by less, about 31 % of the median-pathway level at any AC scenario; averaged across scenarios (the AC-driven range over the three pathways, the SSP-driven range over the three AC levels), the AC-driven range is roughly two and a half times the SSP-driven range, and about 1.9 times the SSP-driven range measured in the current-AC column alone. Moderate penetration (30 %) gives 1,164 to 1,546 deaths per year, about a fifth below the current-AC column, and the full grid runs from ∼657/year (SSP1-2.6 with 75 % AC) to ∼1,920/year (SSP5-8.5 with 5 % AC), or Int$0.13–0.38 billion per year at the central VSL. Within the median SSP2-4.5 pathway, ambitious AC adoption would avert roughly 828 neonatal deaths per year (∼Int$0.17 billion) relative to continued 5 % penetration.

How much the projection depends on its parameters. The 657 to 1,920 range is a scenario grid, not a confidence interval. In Online Appendix Table H.1, VSL rescales every scenario by the same factor; the leave-one-district-out band on the tail slope, $[+0.0272, +0.0492]$, moves each figure by about -37 to $+14$ %; and, under SSP2-4.5, ρ moves the ambitious-AC figure from 394 ($\rho = 1$) to 985 ($\rho = 0.5$) deaths per year but barely moves the current-AC figure. The widest source of uncertainty is the climate projection itself: across the 13-GCM ensemble the SSP2-4.5 increase in ≥ 40°C days ranges from 0.03 (10th percentile) to 3.17 (90th percentile) on the scale of the bias-corrected data of Mishra et al. (2020), carrying the current-AC projection from 621 to 2,326 deaths per year. The level of the mid-century burden should therefore be read as an order of magnitude. The ordering is what survives: moving from current to ambitious AC cuts projected deaths by about 55 % at the baseline ρ (and 74 % at $\rho = 1$), more than the gap between the optimistic and pessimistic emissions pathways; even at $\rho = 0.5$ the cut, about 36 % (1,536 → 985), still exceeds that gap (about 31 % of the median-pathway level at any ρ), although only narrowly.

9 Conclusion

The neonatal heat-mortality dose-response in our sample has two parts. At 30 to 40°C it falls from +0.0045 in the 1980 cohort to a small, statistically undetectable level from the 1990 cohort onward, and one-sided tests detect the decline against every later cohort. At ≥ 40°C it is positive and significant in the pooled regression and large and statistically significant

for the most recent (2010) cohort (identified mainly from the most extreme district-years in the northwest), and no test detects a decline from earlier cohorts, whose tail estimates are imprecise. Applied to realized 2010s extreme-heat exposure, the tail effect implies about 5,340 neonatal deaths attributable to extreme heat over the decade, or roughly a third fewer, at the point estimate of the offset, if part of the effect brings forward deaths from later in infancy.

We find evidence on what drove the first part but not the second. In the staggered REB rollout, connection flattens the high-heat neonatal slope by 0.0039 per +10 days under conventional two-way fixed effects and by 0.0058 under a heterogeneity-robust cohort-pooled IW estimator, both significant at the 1 % level; the imputation estimator gives a similar but imprecise 0.0055. These estimates exceed the observed cross-cohort decline in that slope (0.0026–0.0036 per +10 days), and the cohort-pooled IW estimate slightly exceeds even the 1980-cohort slope (+0.0045), which, because connections began in 1980 and that cohort spans births in 1976–89, understates the pre-electrification slope; we read them as evidence on the direction and timing of the flattening rather than its exact size. The rollout averted about 718,000 neonatal deaths over the 1990s–2010s (95 % CI 323,000 to 1.11 million). Large-city births, which had grid power before the program, show no flattening, as the electrification interpretation predicts, though the city contrasts are imprecise, and no pre-connection test rejects. At the tail the REB estimates are too imprecise (cohort-pooled IW +0.0107, SE 0.0259) to tell whether electrification reduced extreme-heat mortality; because every program district was connected by 2001, the large 2010-cohort tail slope shows only that it did not eliminate it. It is tempting to attribute the asymmetry to the gap between near-universal fan ownership and negligible air conditioning, but we have no direct evidence for it. Why adaptation is not detectable above about 40°C is the main question we leave open.

This two-part pattern matters for how the cost of warming is projected. The Carleton et al. (2022) framework, for instance, projects climate-attributable mortality, including under-five mortality, with a response that flattens smoothly with income and the local climate. Such smooth attenuation can overstate adaptation at the extreme tail, where we detect none, and understate it at high heat, where it has occurred. The two errors need not cancel: the margin left unadapted is the one at which each additional day is most lethal.

The urgency comes from the projections. Days above 40°C are rare in our panel and have not trended upward over four decades, but the ensemble median we compute from the 13-GCM bias-corrected CMIP6 data of Mishra et al. (2020) implies that, on the ERA5 scale of our estimates, they rise from about 0.8 per district-year over 1995–2014 to between 1.9 and 2.6 by 2041–2060, depending on the emissions pathway. Applied to the tail slope of the most

recent cohort we observe, this implies 657 to 1,920 excess neonatal deaths per year across emissions pathways and air-conditioning scenarios, 1,445 to 1,920 at current (5 %) coverage, of which moderate (30 %) coverage would avert 281 to 374 (Section 8). Averaged across scenarios, the projections are about two and a half times as sensitive to air-conditioning coverage as to the emissions pathway, but that comparison rests on a calibrated attenuation of the slope rather than on an estimate from our data, which we cannot verify (Section 8 and Online Appendix Section H.2). We do not claim to have identified the technology that would protect at the extreme. What our estimates indicate is that the margin at which we detect no adaptation is the margin warming is set to expand: after four decades of rising income and electrification, the tail effect is still large and statistically significant for the most recent cohort.

Data availability

The Bangladesh Demographic and Health Survey (BDHS) birth histories are distributed by the DHS Program and the ERA5 reanalysis by the Copernicus Climate Data Store; both are free to any registered user, and the paper uses them as distributed. CHIRPS precipitation, the bias-corrected CMIP6 projections of Mishra et al. (2020) and the UN World Population Prospects 2022 birth series (United Nations, 2022) are publicly available. The rural-electrification connection dates were compiled from the official fact sheets of the Palli Bidyut Samities (PBS), the Bangladesh Rural Electrification Board's MIS reports and the Board's 2004 PBS-wise thana list (Rural Electrification Board, 2004), with a citation for every district. The full replication code, which rebuilds every table and figure from the raw files, and the compiled connection dates will be deposited in the Harvard Dataverse upon acceptance.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the authors used Claude (Anthropic) to assist with coding and debugging of the analysis code and with drafting, including grammar and sentence-structure fixes and trimming of the text. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

Maximilian Auffhammer. Quantifying economic damages from climate change. *Journal of Economic Perspectives*, 32(4):33–52, 2018. doi: 10.1257/jep.32.4.33.

Rakesh Banerjee and Riddhi Maharaj. Heat, infant mortality, and adaptation: Evidence from India. *Journal of Development Economics*, 143:102378, 2020. doi: 10.1016/j.jdeveco.2019.102378.

Bangladesh Rural Electrification Board. Management information system (MIS) report, July 2026. Monthly report, Bangladesh Rural Electrification Board (BREB), Dhaka, Bangladesh, 2026. URL https://reb.gov.bd/pages/static-pages/6922e00f933eb65569e25473. MIS Statement, column "Date of Commercial Operation" for all 80 Palli Bidyut Samities. Accessed October 2026.

Alan Barreca, Karen Clay, Olivier Deschênes, Michael Greenstone, and Joseph 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. doi: 10.1086/684582.

Alan I. Barreca. Climate change, humidity, and mortality in the United States. *Journal of Environmental Economics and Management*, 63(1):19–34, 2012. doi: 10.1016/j.jeem.2011.07.004.

Bill Bell, Hans Hersbach, Adrian Simmons, Paul Berrisford, Per Dahlgren, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Raluca Radu, Dinand Schepers, Cornel Soci, Sebastien Villaume, Jean-Raymond Bidlot, Leopold Haimberger, Jack Woollen, Carlo Buontempo, and Jean-Noël Thépaut. The ERA5 global reanalysis: Preliminary extension to 1950. *Quarterly Journal of the Royal Meteorological Society*, 147(741):4186–4227, 2021. doi: 10.1002/qj.4174.

Kirill Borusyak, Xavier Jaravel, and Jann Spiess. Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6):3253–3285, 2024. doi: 10.1093/restud/rdae007.

Robin Burgess, Olivier Deschênes, Dave Donaldson, and Michael Greenstone. Weather, climate change and death in India. Working paper, London School of Economics, 2017. URL https://www.lse.ac.uk/economics/Assets/Documents/personal-pages/robin-burgess/weather-climate-change-and-death.pdf. Version of April 2017. Accessed October 2026.

A. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427, 2008. doi: 10.1162/rest.90.3.414.

A. Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249, 2011. doi: 10.1198/jbes.2010.07136.

Tamma Carleton, Amir Jina, Michael Delgado, Michael Greenstone, Trevor Houser, Solomon Hsiang, Andrew Hultgren, Robert E. Kopp, Kelly E. McCusker, Ishan Nath, James Rising, Ashwin Rode, Hee Kwon Seo, Arvid Viaene, Jiacan Yuan, and Alice Tianbo Zhang. Valuing the global mortality consequences of climate change accounting for adap-

tation costs and benefits. *Quarterly Journal of Economics*, 137(4):2037–2105, 2022. doi: 10.1093/qje/qjac020.

François Cohen and Antoine Dechezleprêtre. Mortality, temperature, and public health provision: Evidence from Mexico. *American Economic Journal: Economic Policy*, 14(2): 161–192, 2022. doi: 10.1257/pol.20180594.

Timothy G. Conley. GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45, 1999. doi: 10.1016/S0304-4076(98)00084-0.

Clément de Chaisemartin and Xavier D'Haultfœuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996, 2020. doi: 10.1257/aer.20181169.

Olivier Deschênes and Michael Greenstone. Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. *American Economic Journal: Applied Economics*, 3(4):152–185, 2011. doi: 10.1257/app.3.4.152.

Michael Geruso and Dean Spears. Heat, humidity, and infant mortality in the developing world. NBER Working Paper 24870, National Bureau of Economic Research, Cambridge, MA, 2018.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021. doi: 10.1016/j.jeconom.2021.03.014.

Jerry A. Hausman, Jason Abrevaya, and Fiona M. Scott-Morton. Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics*, 87(2):239–269, 1998. doi: 10.1016/S0304-4076(98)00015-3.

Hans Hersbach, Bill Bell, Paul Berrisford, Gionata Biavati, András Horányi, Joaquín Muñoz Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Iryna Rozum, Dinand Schepers, Adrian Simmons, Cornel Soci, Dick Dee, and Jean-Noël Thépaut. ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2023. URL https://doi.org/10.24381/cds.adbb2d47. Accessed October 2026.

Garth Heutel, Nolan H. Miller, and David Molitor. Adaptation and the mortality effects of temperature across U.S. climate regions. *Review of Economics and Statistics*, 103(4): 740–753, 2021. doi: 10.1162/rest_a_00936.

Vimal Mishra, Udit Bhatia, and Amar Deep Tiwari. Bias-corrected climate projections for South Asia from Coupled Model Intercomparison Project-6. *Scientific Data*, 7:338, 2020. doi: 10.1038/s41597-020-00681-1.

Nathan B. Morris, Timothy English, Lily Hospers, Anthony Capon, and Ollie Jay. The effects of electric fan use under differing resting heat index conditions: A clinical trial. *Annals of Internal Medicine*, 171(9):675–677, 2019. doi: 10.7326/M19-0512.

NIPORT and ICF. Bangladesh Demographic and Health Survey 2022: Final report. Dhs final report, National Institute of Population Research and Training (NIPORT) and ICF, Dhaka, Bangladesh, and Rockville, Maryland, USA, January 2024. URL https://dhsprogram.com/publications/publication-fr386-dhs-final-reports.cfm. DHS Program publication FR386. Accessed October 2026.

Palli Bidyut Samities. "At a Glance" fact sheets of the 80 Palli Bidyut Samities. Official websites of each Palli Bidyut Samity (rural electricity cooperative), linked from

the Bangladesh Rural Electrification Board's list of PBS websites, 2026. URL https://reb.gov.bd/pages/static-pages/6922dfe0933eb65569e24707. Accessed October 2026.

Lisa A. Robinson, James K. Hammitt, and Lucy O'Keeffe. Valuing mortality risk reductions in global benefit-cost analysis. *Journal of Benefit-Cost Analysis*, 10(S1):15–50, 2019. doi: 10.1017/bca.2018.26.

Rural Electrification Board. PBS wise thana. Rural Electrification Board website, http://www.reb.gov.bd/pbs_thana.htm, 2004. List of the thanas (upazilas) included in each Palli Bidyut Samity; site copyright 2004. Archived 21 February 2006, https://web.archive.org/web/20060221220212/http://www.reb.gov.bd/pbs_thana.htm. Accessed October 2026.

Jan C. Semenza, Carol H. Rubin, Kenneth H. Falter, Joel D. Selanikio, W. Dana Flanders, Holly L. Howe, and John L. Wilhelm. Heat-related deaths during the July 1995 heat wave in Chicago. *New England Journal of Medicine*, 335(2):84–90, 1996. doi: 10.1056/NEJM199607113350203.

Cornel Soci, Hans Hersbach, Adrian Simmons, Paul Poli, Bill Bell, Paul Berrisford, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Raluca Radu, Dinand Schepers, Sebastien Villaume, Leopold Haimberger, Jack Woollen, Carlo Buontempo, and Jean-Noël Thépaut. The ERA5 global reanalysis from 1940 to 2022. *Quarterly Journal of the Royal Meteorological Society*, 150(764):4014–4048, 2024. doi: 10.1002/qj.4803.

Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021. doi: 10.1016/j.jeconom.2020.09.006.

United Nations. World Population Prospects 2022, online edition. Online database, Department of Economic and Social Affairs, Population Division, New York, 2022. URL https://population.un.org/wpp/. Accessed October 2026.

Eline van de Kamp and Hein A. M. Daanen. Narrative review on infants' thermoregulatory response to heat. *International Journal of Environmental Research and Public Health*, 22 (8):1265, 2025. doi: 10.3390/ijerph22081265.

Matthew D. Webb. Reworking wild bootstrap-based inference for clustered errors. *Canadian Journal of Economics*, 56(3):839–858, 2023. doi: 10.1111/caje.12661.

World Bank. GNI per capita, PPP (current international $), indicator NY.GNP.PCAP.PP.CD. World Development Indicators, 2026. URL https://data.worldbank.org/indicator/NY.GNP.PCAP.PP.CD. Accessed October 2026.

World Health Organization. Newborn mortality. Fact sheet, World Health Organization, Geneva, March 2024a. URL https://www.who.int/news-room/fact-sheets/detail/newborn-mortality. Fact sheet dated 14 March 2024. Accessed October 2026.

World Health Organization. Neonatal mortality rate (per 1,000 live births). WHO Global Health Observatory, UN IGME estimates, 2024b. URL https://data.who.int/indicators/i/E3CAF2B/A4C49D3. Accessed October 2026.