Environmental Justice in Weather Forecasting

Laura Bakkensen, Derek Lemoine, Jeffrey G. Shrader*

July 2024

Abstract

Local climate significantly influences mortality rates, with extreme heat causing more fatalities annually in the United States than any other natural disaster. Weather forecasts are an important public intervention to help prevent heat-related mortality, yet their benefits may not be equally distributed across demographic groups. This paper addresses two questions: First, do temperature forecasts prevent mortality equally among different racial groups? Second, what mechanisms explain any observed disparities? Analysis of comprehensive data on mortality records and weather forecasts for a 12-year period reveals that accurate forecasts significantly reduce mortality for white individuals but not for people of color. We explore several potential mechanisms, including differences in forecast accuracy, communication, trust, and constraints on adaptive behavior. Preliminary findings suggest differences in forecast access and constraints contribute to the unequal benefits. These insights support the need for targeted policies to address these disparities. (JEL:D83,I12,Q51)

*Bakkensen: School of Government and Public Policy, University of Arizona (email: laurabakkensen@arizona.edu); Lemoine: Department of Economics, University of Arizona, NBER, and CEPR (email: dlemoine@arizona.edu); Shrader: School of International and Public Affairs, Columbia University and IZA (email: jgs2103@columbia.edu). This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2112. (CBDRB-FY22-P2112-R9939) Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. The authors gratefully acknowledge NSF grant SES-2315555 for funding. Fanyu Wang provided excellent research assistance. All errors are our own.

1 Introduction

Local climate is an important determinant of mortality (Finkelstein et al., 2021). In the United States, hot temperatures in particular lead to more fatalities in a typical year than any other natural disaster (Heutel et al., 2021, Carleton et al., 2022). Weather forecasts are a widely utilized intervention designed to help prevent mortality. Recent research has shown that accurate temperature forecasts save many lives each year and have especially helped in hot temperatures (Shrader et al., 2023). Despite the health benefits provided by forecasts, their impact may not be uniformly experienced by different demographic groups. This paper addresses two research questions: First, do temperature forecasts prevent mortality equally across demographic groups in the United States? Second, what mechanisms drive the disparities across these groups?

For the first question, we examine effects across three race groups (white; Black; or other races encompassing Asian or Pacific Islander, and American Indian or Alaska Native). The data and methodology for this result are detailed below. We find that the benefit of more accurate forecasts is greatest for white individuals: on average, more accurate forecasts substantially reduce mortality for white individuals but do not reduce mortality for any people of color. This heterogeneous effect is apparent across the temperature distribution (see Figure 1).

The heterogeneity could be due to differences in preferences, in objective quality of forecasts (which could affect reliance on forecasts), in communication of information, in subjective evaluation of forecasts, or in constraints on adaptive behavior.[1] These channels here manifest as differences in the value of reducing mortality risk, in average forecast quality, in trust in media, in exposure to forecasts, and in scope to act on forecasts. Prior work suggests that the first possibility is unlikely given that realized temperature is a relatively more important source of mortality for people of color than for white Americans (Basu, 2009). We show that the second possibility does not hold, as forecast accuracy is comparable across the demographic groups. The third and fourth possibilities would imply benefits from a policy to improve communication and outreach, and the fifth possibility would imply benefits from a policy to relax constraints on adaptation. This project uses a variety of data sources to investigate these potential explanations.

In preliminary analysis, we find evidence for all of the latter three channels. Compared to white populations in the U.S., people of color report having more trouble

[1] We plan to formalize these channels in a transparent model.

accessing the forecast information they need, and they report paying less attention to forecasts of temperature when they do access weather forecasts. We also find that white and Black populations both shift time use in response to forecasts, but in ways that differentially affect mortality. On average, people of all races shift time use out of leisure in response to forecasts of hot weather. Black populations shift into work time. White populations shift time use toward home production. These patterns hold within occupations as well. We will explore whether the wedge in weather-related mortality reflects a principal-agent problem: firms may prioritize adaptation that helps with productivity, whereas workers may prioritize adaptation that helps with hazards. This type of market failure would support policies such as the heat regulations recently promulgated by OSHA (OSHA, 2024).

2 Data

We first estimate the effect of forecasts on mortality rates for each demographic group. Forecast data come from the National Weather Service (NWS) National Digital Forecast Database (NDFD). The NDFD provides a comprehensive dataset of forecasts issued by the NWS starting in 2005. Forecasts are issued by NWS meteorologists working in forecasting offices spread across the United States. The meteorologists issue forecasts for a County Warning Area (CWA), which we use as the spatial unit for clustering standard errors. The analysis below focuses on one-day-ahead temperature forecasts.

The forecasts are combined with data on weather realizations from PRISM (Parameter-elevation Regressions on Independent Slopes) Climate Group data (PRISM Climate Group, 2004). PRISM combines weather station observations with an interpolation procedure that accounts for features such as elevation, weather inversions, rain shadows, and coastal proximity. The outputs are daily measures of weather on a 4km grid across the country. We aggregate the weather forecast and realization data to the county-day level using weighted averages with weights based on gridded population (CIESIN, 2017).

Mortality data come from the Center for Disease Control's National Center for Health Statistics Multiple Cause of Death file. It contains records of all vital events that occurred in the U.S. from 2005 to 2017. We use the restricted access version of the dataset, which includes the day and county of each vital event. From the set of all mortality events, we calculate county mortality rates for each race group per 100,000 people by dividing the total vital events within each race group each day by

by the county population in that year. Population figures are from the NIH Surveillance, Epidemiology, and End Results (SEER) Program. Decedent demographics are typically filled out by the funeral home director or a healthcare worker, typically with input from next of kin (Ver Ploeg and Perrin, 2004).

The data on mortality, forecasts of temperature, and weather realizations are merged. The resulting primary dataset is a daily, county-level balanced panel running from 2005 to 2017.

To examine effects on beliefs, information, and trust, we use data from three surveys. First, we use a nationally representative 2018 Pew survey of that asked individuals about perceptions of local news trustworthiness (Pew, 2020). Second, we use the Oklahoma Weather, Society, and Government survey, a longitudinal survey run from 2014 through 2020 that measures individual respondents' perceptions of weather and weather forecasting (The University of Oklahoma, 2020). Third, we use a nationally representative 2023 YouGov survey about weather forecasts (Orth, 2023).

To examine the scope people have to act on forecasts, we examine individual adaptation actions recorded in the American Time Use Survey (ATUS). The data structure is a repeated cross section. Individuals who have taken part in the CPS are invited to complete a time use diary for a single day's time use. The sample is gathered uniformly throughout the year and across the country. About 10,000 to 12,000 individuals participate each year, and data are available from 2005 through 2023.

3 Methods

To estimate the effect of forecast accuracy on mortality, we fit a distributed lag model that relates mortality to polynomials of forecast error. Improvements in forecast accuracy translate into reductions in mortality if mortality is convex in forecast errors (see Shrader et al. (2023) for details). We therefore estimate:

$$
\begin{aligned}
y_{dct} = & \sum_{\ell=0}^{L} \sum_{j=1}^{J} \mathbf{1}\{f_{c,t-\ell} \in B_j\} \Big(\beta_{0,\ell,j} + \beta_{1,\ell,j} e_{c,t-\ell} + \beta_{2,\ell,j} e_{c,t-\ell}^2 \\
& \qquad + \beta_{3,\ell,j} T_{c,t-\ell} + \beta_{4,\ell,j} T_{c,t-\ell}^2 + \beta_{5,\ell,j} e_{c,t-\ell} T_{c,t-\ell} \Big) \qquad (1) \\
& + \sum_{\ell=0}^{L} g_\ell(\text{prec}_{c,t-\ell}; \xi) + X_{ct}\gamma + \alpha_{cm} + \rho_t + \varepsilon_{ct}.
\end{aligned}
$$

The dependent variable, y_{ct}, is the daily mortality rate for demographic group d in county c on day t. We use L lags of forecasts and temperatures, with effects allowed to vary in each of J forecast bins. The model is fit separately for each of three race groups.

The primary right-hand side variables are realized temperature T_{ct} and forecast error $e_{ct} \triangleq f_{ct} - T_{ct}$ (based on the one-day-ahead forecast f_{ct} of temperature).[2] We include an interacted quadratic in realized temperature and forecast errors. Each coefficient $\beta_{k,\ell,j}$ gives the effect on mortality ℓ days later of marginally greater forecast errors and/or marginally greater realized temperature on a day when the forecasted temperature was in bin B_j. Controlling for temperature realizations ensures that our forecast error estimates can be interpreted as the effect of varying forecasts around a given realization. Any remaining effect of forecast errors must be based on how people use forecasts rather than on the actual weather. The coefficient $\beta_{2,\ell,j}$ corresponds to the convexity of mortality in forecast errors and thus determines whether and how a more accurate forecast saves lives ℓ periods later. We sum these coefficients over the L-lag horizon in order to obtain the cumulative number of lives saved from more accurate forecasts.

The other components of the estimating equation are controls. The function $g_\ell(\text{prec}_{c,t-\ell}; \xi)$ is specified as indicators for daily precipitation above the county's median. This control accounts for potentially correlated effects of rainfall. Date fixed effects (denoted ρ_t) remove any national, time-based confounders, including day-of-week effects, holidays, overall patterns in economic activity, changes in national policy, and large-scale weather patterns. County-by-month fixed effects (α_{cm}) remove local seasonal patterns. These fixed effects also adjust for any time-invariant differences across locations including long-run climate, average medical care availability, and fixed economic conditions. X_{ct} includes county-month fixed effects interacted with linear time trends to account for any overall trends in county- or season-specific weather or mortality. Finally, we weight the regressions by annual population in the county given that the regression is aggregated to the county level.

Standard errors are clustered at the CWA level. CWAs are collections of counties that receive weather forecasts from the same NWS forecast office. The counties are grouped based on technological considerations related to weather observation equipment and so that meteorologists can specialize in forecasting weather for particular areas of the country. The CWA is a natural level for clustering because its counties have similar weather and because its counties all receive forecasts from the same

[2] To reduce the influence of potential outliers, we Winsorize forecast errors at the 1% level.

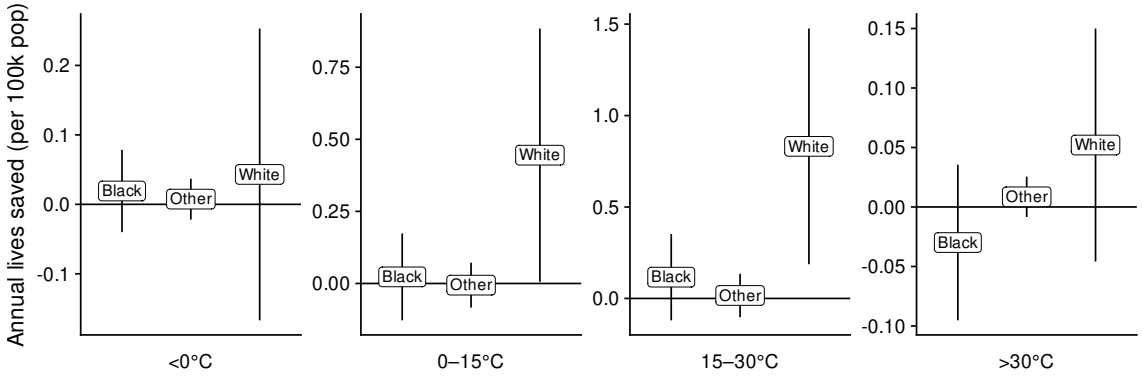
group of meteorologists.

4 Results

4.1 Effect of Forecast Accuracy on Mortality

Figure 1 reports the results of estimating Equation (1) on the primary dataset. The panel shows the results of three models—one for each of the three indicated race groups. The coefficients shown in the figure are 6-day cumulative effects of the coefficient on the square of forecast error multiplied by the days per year that the temperature in each forecasted temperature bin (indicated below each panel of the figure) is experienced in the U.S. The resulting value is the lives saved by a marginal improvement in forecast accuracy (a 1°C reduction in forecast error standard deviation).

Figure 1: Effect of Forecast Accuracy Improvement on Lives Saved by Race of the Deceased



Notes: The figure shows the annual lives saved (per 100,000 people) from a 1°C reduction in the standard deviation of forecast errors. Estimates for each demographic group come from a separate model fit using Equation (1) on the baseline data and translated into annual lives saved by multiplying the convexity by the days per year in each forecasted temperature bin. The lines are 95% confidence intervals based on standard errors clustered at the CWA level.

The figure shows that, on average, more accurate forecasts save lives for white individuals but not for people of color. This distinction holds across the temperature distribution, with especially large differences at cool and warm temperatures. The units in the figure are annual lives saved per 100,000 people. The results indicate that a marginal improvement in forecast accuracy would save just over 1.5 lives per 100,000 people, with that effect coming nearly entirely from reduction in deaths among white

individuals.

4.2 Drivers of Heterogeneity in Forecast Effectiveness

The heterogeneity in forecast effectiveness for reducing mortality documented above could be due to differences in preferences, information quality, credibility, communication, or constraints on adaptive behavior. Prior work suggests that the first possibility is unlikely given that realized temperature is a relatively more important source of mortality for people of color than for white Americans (Basu, 2009). Considering respondents to the ATUS, we calculate that white respondents experience temperatures above 15°C on 53% of days and temperatures above 30°C on 2.8% of days. Non-white respondents to the survey experience slightly more days above 15°C (56%) and a similar number of days above 30°C. On average, white respondents experience temperatures 1.1°C cooler than non-white respondents (13.9° versus 15°C). Both groups experience roughly similar weather volatility, with the standard deviation of weather across time in ATUS respondent locations of 10°C. Given that ATUS respondents are geocoded at the county level or higher, these statistics might understate the amount of extreme temperature exposure through things like the urban heat island effect, which previous research has shown non-white and low-income residents are disproportionately more likely to face (Hsu et al., 2021).

In terms of information, forecast quality is roughly evenly distributed across different demographic groups in the U.S. Calculating the root mean squared error (RMSE) of one-day-ahead temperature forecasts for ATUS respondents, we find that the RMSE of forecasts for the white respondents is 0.97°C and 0.98°C for non-white respondents. Even if forecast accuracy is similar, however, the value of forecasts will be reduced if the information is not or cannot be utilized by individuals. This might occur because of a lack of credibility for the information or because the information does not reach the individuals.

Existing evidence shows high distrust of media across most Americans, but Black and Hispanic adults are especially likely to state that the media does not understand their personal characteristics or interests (Pew, 2020). Focusing specifically on perceptions of weather forecasts, a 2018 Pew survey shows that Black and white individuals are equally likely to say that weather information is important to know (see Table 1). But Black respondents are significantly more likely to say that it is not easy to stay informed about weather. Respondents of other races are more likely than white respondents to say both that it is less important to know about weather and not easy to stay informed about weather.

Table 1: Perceptions of Weather Forecasts

	For weather information it is:	
	important to know	easy to stay informed
Black	-0.008	-0.032***
	(0.015)	(0.011)
Other	-0.077***	-0.059***
	(0.018)	(0.014)
LHS mean	0.702	0.912
Data source	Pew	Pew
N	34,407	34,407

This table shows results from estimating linear regressions on individual respondent data from (Pew, 2020). The variables are rates of responding "yes" to the indicated question relative to the rate for white respondents. All variables in the regression are shown. Estimates are weighted by sampling weights. Robust standard errors are in parentheses: * $p < .10$, ** $p < .05$, *** $p < .01$.

Orth (2023) data show that Black respondents are less likely to check weather forecasts daily (64% for white respondents compared to 51% for Black respondents, with the Pearson's Chi-square test p-value = 0.005). Black respondents are also substantially less likely to report paying attention to temperature when reading, watching, or listening to weather forecasts (64% for Black respondents versus 80% for white respondents, Pearson's Chi-squared p-value = 0.000). Black respondents were more likely to report that they pay attention to humidity and visibility.

In ongoing work, we are collecting data on the race and ethnicity of local TV meteorologists from Papper and Henderson (2022) to examine whether demographic homophily affects forecast responses within different media markets. To examine how outreach and communication can impact different groups in society, we will also examine exogenous changes in the spatial precision with which forecasts are communicated. If targeted communication is a relatively more important constraint for people of color than for white individuals, improving the granularity of forecasts could serve to close the gap in forecast value between these two groups. Even if increased granularity does not close the race gap, it could still serve to improve overall forecast value across the population.

Specifically, we will exploit two large-scale changes in forecast quality and examine the heterogeneity in resulting mortality reductions across race, ethnicity, and age. First, we will utilize changes to the spatial granularity of forecasts. On August

20, 2014, NWS forecasts for the continental United States transitioned from a 5km forecast grid to a 2.5km forecast grid (Baron, 2019). This refinement in grid size allowed forecasts to better match the local weather conditions rather than the larger regional weather conditions. Second, we will exploit an October 1, 2007 change in severe weather warnings that went from a larger county-level unit of warning analysis to storm-based polygon warning areas that significantly increased the spatial precision of warnings (NWS, 2007). We will use a version of the estimating equation (1) to examine the reductions in mortality across different racial and ethnic groups in response to these more precisely communicated forecasts.

Finally, we examine how actions in response to forecasts vary heterogeneously across race to better understand constraints on adaptive behavior. In particular, Table 2 reports results using American Time Use Survey data to examine if time use changes differ across race in response to forecasts. The results utilize a similar estimation framework to the mortality results, just with linear terms for forecast errors so that we can directly examine whether individuals increase or decrease time use in reponse to hotter or colder forecasts.

In the results shown in Table 2, we find that both Black and white populations shift time use but towards different activities. While all populations shift away from leisure from forecast error at hot temperatures, non-white populations shift towards work while white populations shift towards home production. These patterns also hold when controlling for occupation. We will explore whether a contributor to the weather-mortality relationship includes a principal-agent problem as firms may prioritize adaptation to help with productivity while workers may prioritize responses to mitigate the hazard. If present, this market failure would highlight the utility of policies such as OSHA's recent heat regulations.

5 Conclusion

This study highlights significant disparities in the health benefits of temperature forecasts across racial groups in the United States. Our findings indicate that while accurate forecasts substantially reduce mortality for white individuals, they do not offer similar benefits to people of color. This disparity appears to be driven by differences in forecast access, communication, and adaptive behavior but not differences in forecast accuracy or harms from temperature. Specifically, people of color report facing challenges in accessing and utilizing forecast information. Ongoing work is investigating the degree to which this may result from issues such as limited access to

Table 2: Effect of Forecast Accuracy on Time Use

	(1) All	(2) White	(3) Other
Panel A. Work			
< 0°C × Forecast error	-4.20	-4.39	-8.01
	(3.73)	(4.65)	(12.3)
0 to 15°C × Forecast error	0.44	1.14	-1.91
	(1.78)	(2.31)	(4.79)
15 to 30°C × Forecast error	1.35	3.12	-8.88*
	(1.78)	(2.31)	(4.79)
> 30°C × Forecast error	18.1**	7.65	69.4*
	(7.96)	(7.63)	(35.6)
Panel B. Home Production			
< 0°C × Forecast error	-1.11	1.88	-13.6
	(3.37)	(2.63)	(11.5)
0 to 15°C × Forecast error	-3.36*	-1.18	-6.87
	(1.96)	(1.88)	(4.26)
15 to 30°C × Forecast error	-0.79	-2.12	6.18
	(1.87)	(2.00)	(5.48)
> 30°C × Forecast error	16.0**	20.7**	-14.5
	(7.60)	(9.67)	(21.1)
Panel C. Leisure			
< 0°C × Forecast error	5.31	2.51	21.6*
	(4.73)	(4.72)	(12.9)
0 to 15°C × Forecast error	2.92	0.039	8.77
	(2.13)	(2.27)	(7.73)
15 to 30°C × Forecast error	-0.56	-1.00	2.70
	(2.07)	(2.08)	(7.12)
> 30°C × Forecast error	-34.2***	-28.3***	-54.9*
	(11.5)	(8.18)	(29.2)
Clusters	97	97	89
N	196,571	157,866	37,398

This table shows results from estimating three versions of Equation (1) on individual respondent data from ATUS. In all cases, the quadratic temperature and forecast error terms are excluded for parsimony. The dependent variables are indicated by the title of each panel and are time (in minutes) spent on market work, home production, or leisure. The models include 5°C bins of realized temperature, quartiles in precipitation, four bins for forecasted temperature, date fixed effects, and county-by-month fixed effects interacted with a linear time trend in years. Estimates are weighted by FIPS (county or state) population. Standard errors are clustered at the CWA level: * $p < .10$, ** $p < .05$, *** $p < .01$.

technology or differences in information dissemination. Different groups also undertake different actions in response to forecasts, with white individuals taking action that reduces exposure to temperature while people of color shift exposure from leisure time to the workplace. This difference could be due to differential employment patterns that constrain adaptation behavior for some groups more than others.

Future research should further investigate the underlying mechanisms contributing to these disparities. Specifically, studies could focus on assessing the effectiveness of various communication strategies and outreach programs designed to improve forecast access and usability for underserved communities. Additionally, examining how firms and organizations adapt their policies in response to heat forecasts, and whether these adaptations align with the needs of different demographic groups, could provide valuable insights. Understanding these dynamics will be crucial for developing more equitable policies and interventions that ensure all populations benefit equally from weather forecasts, potentially informing future regulations like those recently proposed by OSHA.

References

Baron, K. (2019). Github - historical weather map layer, using noaa ndfd data. https://github.com/nst-guide/ndfd_historical. (Accessed on 01/02/2023).

Basu, R. (2009). High ambient temperature and mortality: A review of epidemiologic studies from 2001 to 2008. *Environmental Health 8*(1), 1–13.

Carleton, T., A. Jina, M. Delgado, M. Greenstone, T. Houser, S. Hsiang, A. Hultgren, R. E. Kopp, K. E. McCusker, I. Nath, et al. (2022). Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits. *The Quarterly Journal of Economics 137*(4), 2037–2105.

CIESIN (2017). U.S. Census grids 2010. https://sedac.ciesin.columbia.edu/data/collection/gpw-v4. Accessed: 2018-08-30.

Finkelstein, A., M. Gentzkow, and H. Williams (2021). Place-Based Drivers of Mortality: Evidence from Migration. *American Economic Review 111*(8), 2697–2735.

Heutel, G., N. H. Miller, and D. Molitor (2021). Adaptation and the mortality effects of temperature across us climate regions. *Review of Economics and Statistics 103*(4), 740–753.

Hsu, A., G. Sheriff, T. Chakraborty, and D. Manya (2021). Disproportionate exposure to urban heat island intensity across major US cities. *Nature Communications 12*(1), 1–11.

NWS (2007). NWS begins era of more precise warnings. https://www.weather.gov/media/crp/STWJFall07.pdf. (Accessed on 01/02/2023).

Orth, T. (2023). How often and where Americans get information on the weather. https://today.yougov.com/topics/health/articles-reports/2023/05/08/how-and-where-americans-get-information-weather. Accessed: 2023-05-08.

OSHA (2024). Proposed rule: Heat injury and illness prevention in outdoor and indoor work settings. Accessed: 2024-07-30.

Papper, B. and K. Henderson (2022, November). TV news gets more diverse, still doesn't reflect public. https://assets-002.noviams.com/novi-file-uploads/

rtdna/Research/RTDNA-Syracuse_Diversity_in_TV_News_2022.pdf. (Accessed on 07/30/2024).

Pew (2020). Black, Hispanic, white americans feel misunderstood by media for different reasons. https://www.pewresearch.org/fact-tank/2020/06/25/black-hispanic-and-white-adults-feel-the-news-media-misunderstand-them-but-for-very-different-reasons/. (Accessed on 01/02/2023).

PRISM Climate Group (2004). http://prism.oregonstate.edu.

Shrader, J. G., L. Bakkensen, and D. Lemoine (2023). Fatal errors: The mortality value of accurate weather forecasts. *National Bureau of Economic Research w31361*, 41.

The University of Oklahoma (2020). Oklahoma meso-scale integrated sociogeographic network. http://crcm.ou.edu/epscordata/. Data were collected by the University of Oklahoma with support from the National Science Foundation under Grant No. IIA-1301789. Accessed: 2024-07-30.

Ver Ploeg, M. and E. Perrin (2004). *Eliminating Health Disparities: Measurement and Data Needs*. National Academies Press.