

# Striving for Improvement

## The Perceived Value of Improving Hurricane Forecast Accuracy

Renato Molina, David Letson, Brian McNoldy,  
Pallab Mozumder, and Matthew Varkony

**ABSTRACT:** Hurricanes are the costliest type of natural disaster in the United States. Every year, these natural phenomena destroy billions of dollars in physical capital, displace thousands, and greatly disrupt local economies. While this damage will never be eliminated, the number of fatalities and the cost of preparing and evacuating can be reduced through improved forecasts. This paper seeks to establish the public's willingness to pay for further improvement of hurricane forecasts by integrating atmospheric modeling and a double-bounded dichotomous choice method in a large-scale contingent valuation experiment. Using an interactive survey, we focus on areas affected by hurricanes in 2018 to elicit residents' willingness to pay for improvements along storm track, wind speed, and precipitation forecasts. Our results indicate improvements in wind speed forecast are valued the most, followed by storm track and precipitation, and that maintaining the current annual rate of error reduction for another decade is worth between \$90.25 and \$121.86 per person in vulnerable areas. Our study focuses on areas recently hit by hurricanes in the United States, but the implications of our results can be extended to areas vulnerable to tropical cyclones globally. In a world where the intensity of hurricanes is expected to increase and research funds are limited, these results can inform relevant agencies regarding the effectiveness of different private and public adaptive actions, as well as the value of publicly funded hurricane research programs.

**KEYWORDS:** Communications/decision making; Economic value; Hurricanes/typhoons; Policy; Social Science; Societal impact

<https://doi.org/10.1175/BAMS-D-20-0179.1>

Corresponding author: Renato Molina, [renato.molina@rsmas.miami.edu](mailto:renato.molina@rsmas.miami.edu)

Supplemental material: <https://doi.org/10.1175/BAMS-D-20-0179.2>

In final form 16 February 2021

©2021 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

**AFFILIATIONS:** **Molina**—Department of Environmental Science and Policy, Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida, and Department of Economics, Miami Herbert Business School, University of Miami, Coral Gables, Florida; **Letson and Varkony**—Department of Environmental Science and Policy, Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida; **McNoldy**—Department of Atmospheric Science, Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida; **Mozumder**—Department of Earth and Environment, and Department of Economics, Florida International University, Miami, Florida

Landfalling hurricanes in the continental United States are both common and devastating. Between 1900 and 2017, 197 hurricanes have resulted in about \$18 trillion in economic losses (Weinkle et al. 2018). Some cost more than \$100 billion per landfall, such as Hurricanes Andrew and Katrina. In an effort to reduce these impacts, the Hurricane Forecast Improvement Project (HFIP) was established in 2008 with the explicit goal of reducing track and wind speed forecast error over the following 10 years (Gall et al. 2013). By any measure, the project was a success. Storm track and wind speed forecast errors have been cut by more than 40% each, and the project is still in place striving to produce even better forecasts. But, are these further improvements still valuable? Would the public benefit from an even more accurate hurricane forecasts? We shed light on these questions by using an experimental setting that relies on both atmospheric science and contingent valuation methods. Using 2018 as benchmark, our results show the general public significantly values further improvement in forecast accuracy, and that they value a more accurate wind speed prediction the most. As scarce public funds sustain many of these research efforts, these estimates provide a baseline for establishing the value of hurricane research, and the relative importance of different forecast products.

The need for better forecasts follows from recognizing the massive economic impact of hurricanes (Weinkle et al. 2018), as well as their overall threat to sustainable development (Adams and Judd 2016; Gaddis et al. 2007). In fact, previous evidence suggests that officially distributed hurricane forecasts are actively used to inform effective individual adaptive behavior such as timely evacuation decisions and the purchase of protective capital (Emanuel 2005; Mozumder et al. 2015). In addition, more accurate forecasts help emergency responses when it comes to the deployment of manpower and the preparation of critical infrastructure (Marks et al. 1998; Quiring et al. 2014). Indeed, the advantage of having more accurate forecasts is widely recognized by both the public and the private sector, hence the current supply of research efforts and the variety of products offered by the relevant agencies aimed at increasing resiliency from hurricanes (Ewing et al. 2007). The problem, however, is that there are still many open questions regarding the private and social value of these forecast products (Gladwin et al. 2007; Letson et al. 2007).

The gap in our understanding regarding the value of hurricane forecast accuracy can be traced to the way in which the public accesses this information. Because forecasts are not traded in a market setting, people may not have a well-defined preference for these products, and if they do, their preferences may be influenced by their subjective risk perceptions (Trumbo et al. 2016). Moreover, people's risk-averting behavior and the value they put on an accurate forecast are oftentimes intertwined (Letson et al. 2007). To overcome these difficulties, contingent valuation, which relies on survey instead of observational data, can be used to infer the public's preferences regarding these services (Lazo and Waldman 2011). In particular, contingent valuation methods have been used to investigate the value of weather information (Lazo et al. 2009) and improved hurricane forecasts (Lazo and Waldman 2011) in the United States, as well as the value of tropical cyclone early warning systems in Australia (Anaman et al. 1998), Bangladesh (Ahsan et al. 2020), and Vietnam (Nguyen et al. 2013).

While these previous studies have provided extremely valuable insights, we contribute to this literature by further integrating atmospheric science in a contingent valuation effort. Our goal is to credibly create alternative scenarios that relate to the decision-making process of the average user when faced with the threat of a hurricane. These scenarios are then imposed on current forecast products and relatable measures of impact so as to extract an estimate for the value an individual gives to different attributes of the forecast. We then implement a large-scale survey that is both geographically and demographically heterogeneous. In the following section, we lay out the core of this approach and quantify the monetary value of improving forecast accuracy along storm track, wind speed, and precipitation forecasts.

## Methods

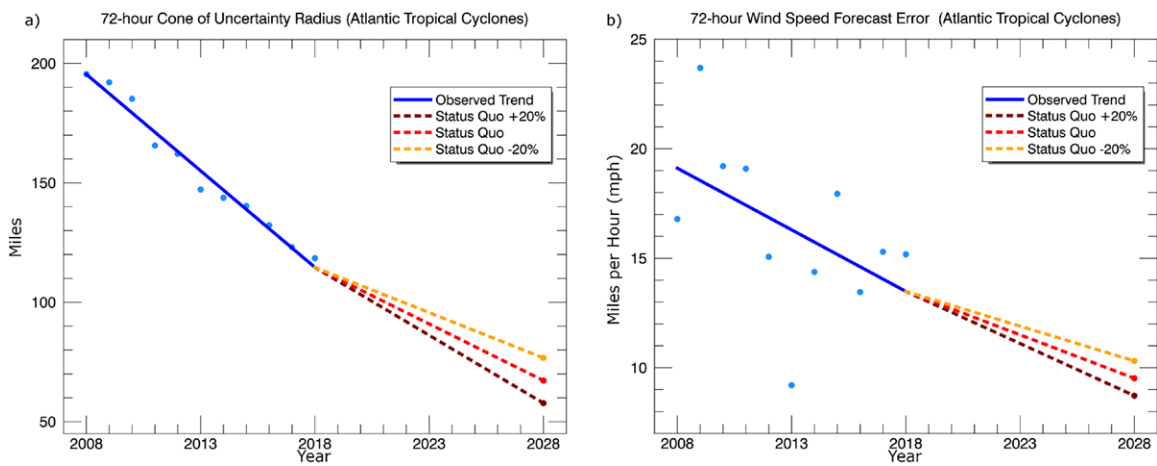
**Hurricane forecast improvement scenarios.** Our goal is to evaluate how the public values further improvement of the forecast product. To do so, we construct hypothetical scenarios using the historical forecast errors from the National Hurricane Center (NHC). NHC calculates and provides its annual average error statistics of track and wind speed. These errors come from comparing the values in all of the real-time forecasts against the corresponding observed “best-track values.” Taking the error reduction trend from 2008 to 2018, we then project our hypothetical improvement scenarios for the next decade (2018–28). We focus on the critical 72-h lead time before landfall (Regnier 2008), and work with the forecast for Florence from 1200 coordinated universal time (UTC) 11 September (landfall was 1115 UTC 14 September) and the forecast for Michael from 1800 UTC 7 October (landfall was 1730 UTC 10 October) as an illustration for how the forecast might look if it were more accurate.

We construct three potential scenarios for the 2018–28 period. The first scenario is the status quo, which assumes the rate of forecast improvement, or error reduction, observed in 2008–18 will continue for another 10 years. The second and third scenarios are either a 20% acceleration or a 20% deceleration with respect to the status quo. We evaluate these projections for the errors of three attributes of a typical forecast: storm track, wind speed, and precipitation.

For track errors, we use the “cone of uncertainty” (just “cone” hereafter) because of the public’s familiarity with it since its introduction in 2002. The NHC updates the size of the cone each year based on track errors over the past five hurricane seasons. Because of the sliding 5-yr averages, variations in the size of the cone are quite smooth from year to year. The observed and the three hypothetical trends of track forecast error are shown in Fig. 1a, with the status quo rate of improvement as the red dashed line, the most aggressive track forecast improvement as the maroon dashed line (20% acceleration of the status quo), and the reduced rate of forecast improvement as the orange dashed line (20% deceleration of the status quo).

Wind speed errors are handled slightly differently. Rather than using a sequence of sliding 5-yr averages, we calculate a linear trend through individual annual error values. The improvement trend is shown in Fig. 1b by the solid blue line. From the individual annual values denoted by light blue dots, it can be seen that a substantial amount of interannual variability arises. Similar to the treatment of the track forecast error improvements, we project the three scenarios into the coming decade (2018–28).

Hypothetical future errors are calculated using a percentage rather than a static number because errors can only approach zero, not reach zero or become negative. So the slope of the blue line and the slope of the red line in Fig. 1 are not equal by design. In addition to reducing the future errors, we also adjust the forecast values closer to the observed values by the same percentage. In other words, we expect that forecasts in future decades will be more accurate and with less uncertainty surrounding them. It has previously been pointed out that forecasts have been generally improving over the past several decades, but there will be a time when forecasts can no longer be improved due to the inherent limit of predictability of



**Fig. 1. Historical and hypothetical hurricane forecast errors.** The figure shows the projections assumed for the construction of the hypothetical scenarios in the survey. (a) The trend in the size of the cone of uncertainty for 72-h track forecasts; (b) the trend in errors for wind speed. The same average percentage of improvement from 2008 to 2018 is extrapolated to 2028 using the same rate of improvement (Status Quo, red dashed line), a 20% increase in that rate (Status Quo + 20%, maroon dashed line), and a 20% decrease in that rate (Status Quo – 20%, orange dashed line).

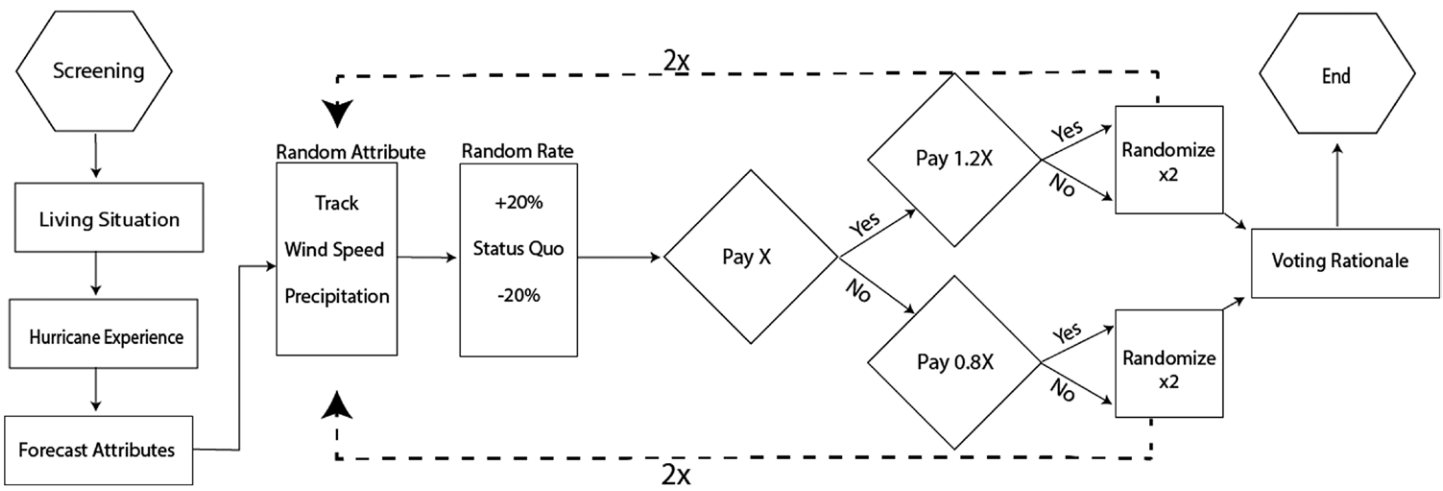
chaotic systems such as the atmosphere (Landsea and Cangialosi 2018). For the sake of this study, we assume the predictability limit will not be reached in the coming decade and that forecasts will continue to become more accurate during that time span.

While track and wind speed are fairly simple metrics to calculate and verify, precipitation metrics are more complex. Precipitation depends on the hurricane location and speed, as well as the storm's intensity and size, and the topography of the affected area. To tackle these difficulties, we rely on the parametric hurricane rainfall model (PHRaM) (Lonfat et al. 2007). PHRaM accounts for storm size, intensity, wind shear-based storm asymmetry, and topographic effects, so we utilize the hypothetical values of intensity and location that were defined above, and use wind shear values from the operational Statistical Hurricane Intensity Prediction Scheme (SHIPS) model output (DeMaria and Kaplan 1994).

To address current and future uncertainty in the precipitation forecast, we use the Monte Carlo ensemble that NHC creates every 6 h for each active storm to produce its suite of wind speed probability forecasts (DeMaria et al. 2009). PHRaM is run on all of the 1,000 realizations, which in turn allows us to ask questions related to the probability of over- or underforecasting rainfall compared to a deterministic forecast (Marks et al. 2020). We extend this analysis for all potential error reduction scenarios (status quo  $\pm$  20%).

**Survey.** To elicit the value of improved hurricane forecasts, we implement the insights above into a web-based survey questionnaire. We target individuals recently affected by our chosen hurricanes, Michael and Florence, so participants can compare the forecast products familiar to them, with those derived from an improved forecast's capability. In particular, we identified coastal counties that were under at least a tropical storm warning in the NHC advisories for Florence and Michael. In addition, to identify inland affected counties, we used the Federal Emergency Management Agency (FEMA) designations of counties eligible for assistance. Those designations include areas in Florida, Georgia, and North and South Carolina. Respondents answer the sequence of questions described in Fig. 2, which seek to extract relevant information on their backgrounds and attitudes toward forecast improvements.

Participants are initially screened by the zip code where they live, as well as their ability to provide thoughtful and honest answers in the survey. Following the introduction, respondents are briefed on the nature of the survey and its potential policy implications. The survey



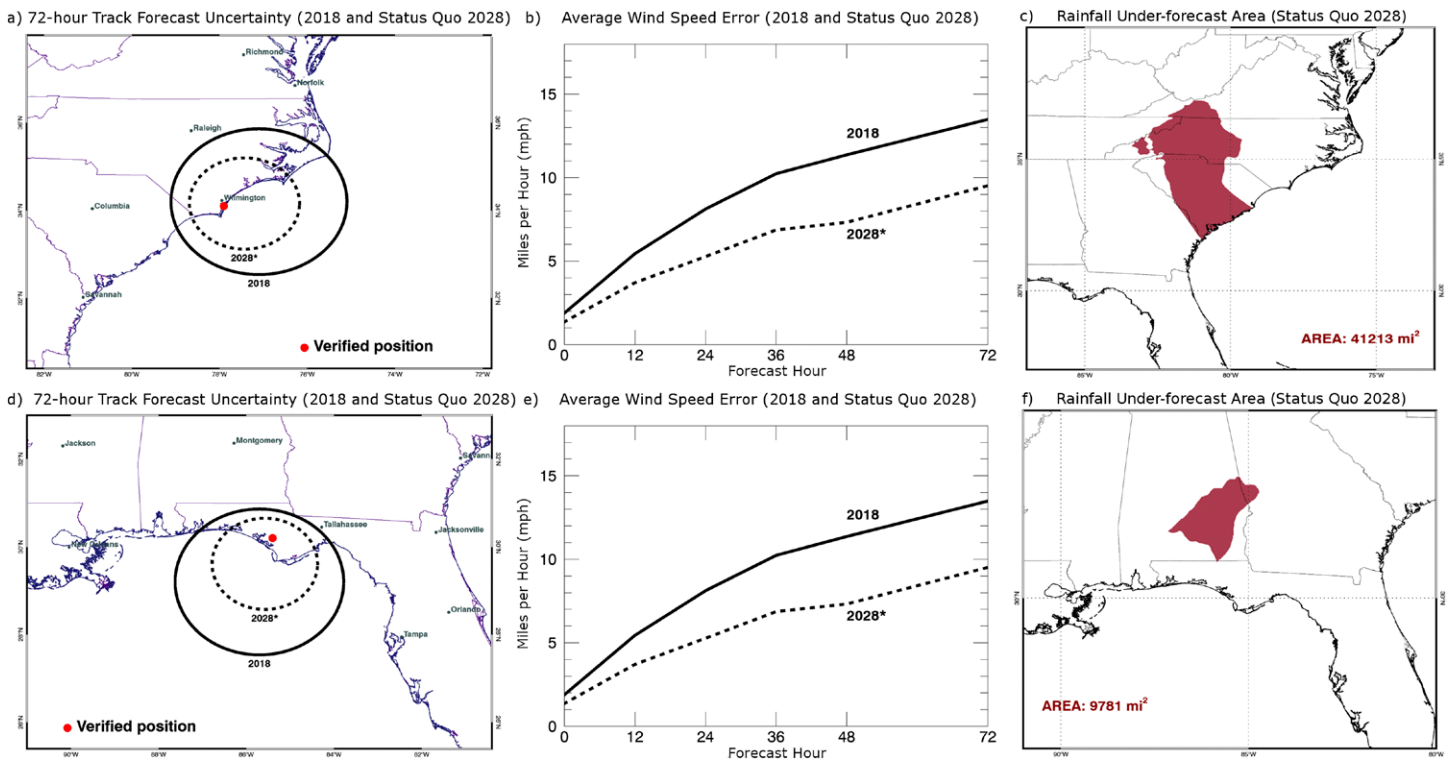
**Fig. 2. Survey flowchart.** The figure describes the structure of the survey. The rectangular sections represent questions pertaining to participant background. The middle of the diagram displays our dichotomous choice design. A random attribute is matched with a random improvement rate. Each attribute and each rate are only used once. The participants must vote in favor or against an annual tax between \$1 and \$50. The tax is then adjusted based on the previous answer, and the respondents are presented with another yes/no vote. This process is repeated two times. Both hexagons represent the beginning and end of the survey.

questions solicit participant information on residential living status and the extent of insurance coverage for their homes. Respondents are also asked to describe their familiarity with hurricane risk and the governmental programs created to protect against hurricane-related damages in the United States.

The next set of questions asks respondents to recount their experience with their respective storm (i.e., Florence or Michael). These questions are preceded with a statement describing the acceptable limits of hurricane experience, informing participants that their experience is not only limited to physical impacts. Depending on their experience, a set of follow-up questions is presented inquiring about evacuation decisions and damages to property. The section concludes with general questions about evacuation plans and the number of individuals living in the residence.

After documenting the respondents' experience with the past hurricane, the survey describes the role of federal agencies in providing hurricane information. We briefly explain the process of fund allocation for hurricane research. In addition, we mention our motivation in collecting individual attitudes toward tax increases to support funding for hurricane forecasting research.

We then conduct our experiment by providing respondents with a set of three random scenarios needed for the contingent valuation. Examples of these forecast components are shown in Fig. 3 for the status quo scenario. Moving along columns from left to right, these figures represent changes in decadal trends of 72-h track forecast uncertainty, wind speed forecast error, and underforecasted precipitation. As shown in Fig. 2, the survey randomly generates a scenario combining one forecast attribute with one rate of forecast improvement. The respondents are provided with a brief description of the randomly selected forecast attribute. A visual and a probabilistic measure of improvement are also included in the description to demonstrate the change in accuracy of the given forecast as a result of the randomly assigned rate of improvement. As stated earlier, the change in forecast abilities is related to the baseline improvements from 2008 to 2018. Respondents are then asked to answer yes or no to a randomly generated annual tax increase meant to pay for these forecast improvements. To decrease ambiguity, we specify that the tax increase takes place at the household level and lasts for 10 years.



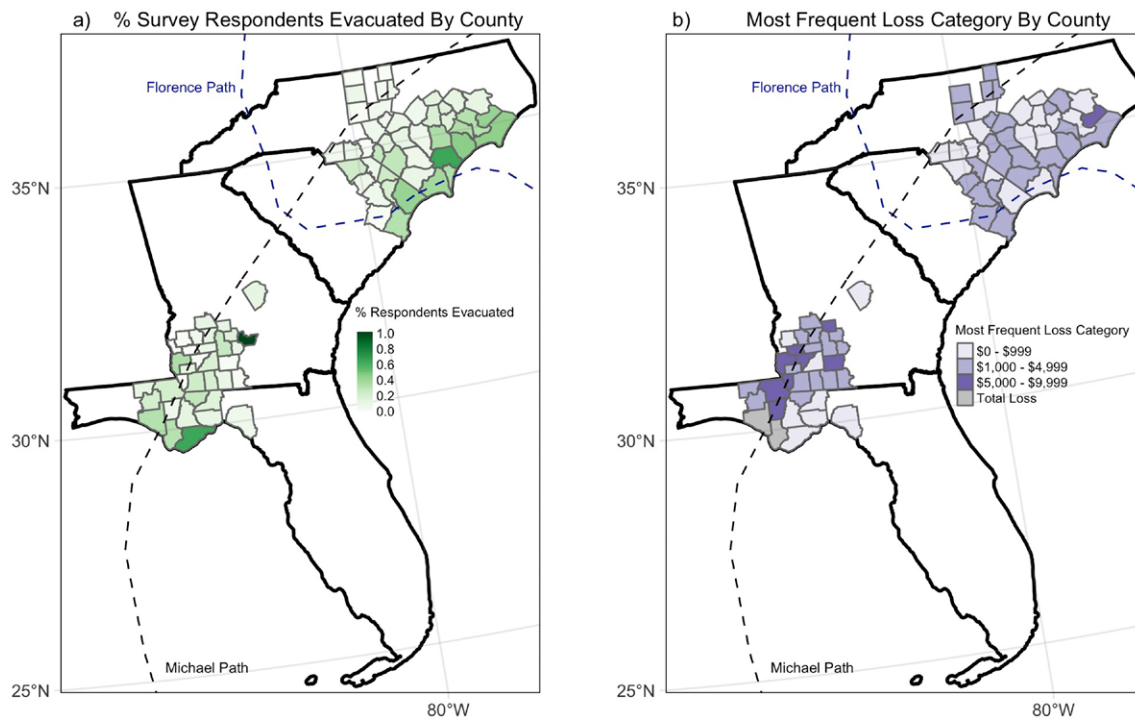
**Fig. 3. Hypothetical forecast components.** The figure shows the maps and charts shown to survey participants in (top) the Florence survey and (bottom) the Michael survey. Each of the three columns contains figures representing forecast improvements corresponding to the status quo improvement. (a),(d) The track uncertainty defined by the size of the cone for 72-h forecasts. (b),(e) The average wind speed error, and are the same because average wind speed forecast error is the same regardless of location. (c),(f) The rainfall underforecast area.

A follow-up yes or no question with an increased or decreased tax, relative to the original tax, is then presented to the respondent. If the respondent answered yes to the initial tax increase, then the follow-up tax increase is 20% greater; if the respondent's initial answer is no, then the follow-up tax is reduced by 20% instead. This random process combining a forecast attribute and rate of improvement is repeated three times. Each survey participant observes all three of the forecast attributes combined with a unique rate of improvement.

Finally, the survey concludes with a sequence of questions asking respondents to explain their own decision making process, including their level of belief that public officials will use the survey information to guide policy implementation. The assumptions underlying the survey as well as the statistical analysis to elicit the willingness to pay forecast improvement are detailed in appendix A.

## Results

The survey is deployed using the Qualtrics platform, and the final sample encompasses a total of 4,650 respondents: 3,150 from the region affected by Hurricane Florence and 1,500 from the region affected by Hurricane Michael as defined by FEMA. From here onward we refer to these two samples simply as Florence and Michael. The mean respondent is 43 years old, earns \$65,176 per year, and lives about 87 km from the coast. The total sample has 70% self-identified females and 53% home owners, with varying degrees of beliefs regarding short- and long-term risk of experiencing a hurricane (see the online supplemental material for more details). Geographically, the span of the survey is shown in Fig. 4, and shows the distribution of the rate of evacuation and the range of the self-reported losses in the sample. Figure 4a shows that evacuations were more prevalent in coastal areas. However, inland counties in the path of the storms experienced partial evacuation as well. In terms of losses, Fig. 4b shows



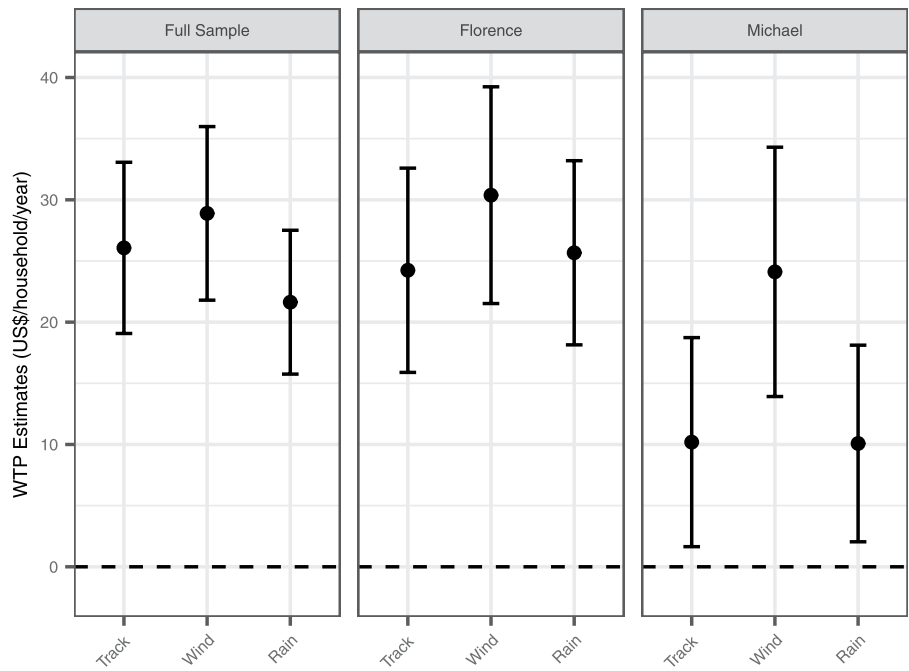
**Fig. 4. Loss and evacuation maps.** The figure shows the geographical distribution of responses to evacuation and total loss questions in the survey: (a) the percent of respondents who evacuated within a county (percent represents the survey sample) and (b) the frequency of total losses (categorized) for survey respondents by county. Dashed lines denote Florence's and Michael's paths.

that self-reported capital losses are widely distributed across the sample, but total losses are more prevalent for coastal counties hit by Michael in the panhandle. In fact, comparing the unconditional mean reported losses reveals that they were larger for the Michael sample ( $p < 0.001$  for a two-sided  $t$  test).

According to our experiment, respondents are willing to pay for further improvement within  $\pm 20\%$  of the rate experienced between 2008 and 2018. This willingness to pay expands to all attributes tested (i.e., storm track, wind speed, and precipitation). Out of the three forecast attributes, respondents across both surveys value further improvements in wind speed forecast the most. These estimates are shown in Fig. 5 and report an average willingness to pay (WTP) of \$26.07, \$28.89, and \$21.63 per household per year for continued improvement in storm track, wind speed, and precipitation forecast accuracy, respectively (see supplemental material for estimation tables). The overall ranking of WTPs is maintained across the two samples, but the Florence sample exhibits a higher average WTP for all attributes when considered in isolation. This pattern is not explained by differences in income or damage experienced. Likely, these differences in preferences follow from the fact that damages caused by Florence were mostly associated with flooding, while the damages from Hurricane Michael were mostly associated with strong winds and storm surge (Stewart and Berg 2019), but also because locations might be correlated with other underlying preferences (see supplemental material for estimations using additional sets of responses, which reduce the discrepancy between samples, but also decrease the precision of the estimates). Because of the large size of our sample, however, we are able to statistically control for these differences and appropriately account for them in the full sample estimates.

Projecting these estimates out of sample relies on the assumption that survey responses are representative of out-of-sample preferences. Other statistically significant covariates (see supplemental material for estimation tables) are obtained from the U.S. American Community Survey (ACS). Accordingly, extrapolating to the regions affected by both Hurricanes Florence and Michael indicates an annual total WTP of \$60 million, \$67 million, and \$50 million

for improvements in storm track, wind speed, and precipitation forecast, respectively. These results are shown in Table 1. Further, we can extend these WTPs to all areas that are exposed to the effects of hurricanes in the continental United States. This extrapolation, however, depends heavily on a key assumption about what makes a certain area exposed or not. Our criterion is having experienced hurricane-related wind speeds above a certain threshold from 2006 onward. In particular, we focus on areas that have experienced sustained winds above 20, 30, 40, and 50 mph due to Hurricanes Florence, Harvey, Ike, Irma, or Michael. The geographical span of these areas is shown in Fig. 6.



**Fig. 5.** Average willingness to pay for hurricane forecast improvement. The figure displays the average willingness to pay (WTP) for further improvements in the precision of storm track, wind speed, and precipitation forecasts. The figure is split into three panels according to the sample used to estimate the respective WTPs: the full sample, respondents from counties affected by Hurricane Florence, and counties affected by Hurricane Michael, respectively. Bars denote the 95% confidence interval.

Highly exposed areas border the Gulf of Mexico and the southern states along the Atlantic coast. As the threshold wind speed is reduced, inland areas are further classified as exposed. Using household data from the 2013–17 ACS, we then project the total WTP, WTP per capita (based on average household occupancy as reported by the ACS), and the present value (PV)

**Table 1.** Extrapolation of willingness to pay for hurricane forecast improvement. All monetary units are in 2018 U.S. dollars (USD). Extrapolation based on the number of households occupied as reported in the 2013–17 American Community Survey (ACS). Exposed counties are those that have experienced at least 20-, 30-, 40-, and 50-mph wind due to a hurricane between 2006 and 2018, respectively. Per capita values only consider individuals over 18 years old as per the 2013–17 ACS, and future values are discounted at 2% yr<sup>-1</sup> for 10 years.

	Sampled	>50 mph	>40 mph	>30 mph	>20 mph
Total WTP (USD × 10 <sup>6</sup> yr <sup>-1</sup> )					
Storm track	\$60 ± 16	\$376 ± 98	\$566 ± 149	\$946 ± 248	\$1,443 ± 378
Wind speed	\$67 ± 17	\$428 ± 101	\$645 ± 152	\$1,078 ± 254	\$1,637 ± 386
Precipitation	\$50 ± 14	\$327 ± 89	\$493 ± 135	\$824 ± 225	\$1,253 ± 342
Per capita WTP (USD per year per person)					
Storm track	\$12.56 ± 3.37	\$11.31 ± 2.07	\$11.47 ± 3.01	\$11.57 ± 3.04	\$11.72 ± 3.07
Wind speed	\$13.92 ± 3.42	\$12.89 ± 3.04	\$13.07 ± 3.09	\$13.18 ± 3.10	\$13.30 ± 3.13
Precipitation	\$ 10.42 ± 2.83	\$9.85 ± 2.69	\$9.99 ± 2.73	\$10.07 ± 2.75	\$10.17 ± 2.78
PV (USD per person)					
Storm track	\$115.08 ± 30.89	\$103.62 ± 27.22	\$105.09 ± 27.61	\$106.01 ± 27.84	\$107.38 ± 28.08
Wind speed	\$127.54 ± 31.32	\$118.10 ± 27.81	\$119.75 ± 28.30	\$120.76 ± 28.44	\$121.86 ± 28.70
Precipitation	\$95.47 ± 25.96	\$90.25 ± 24.68	\$91.53 ± 25.03	\$92.26 ± 25.23	\$93.18 ± 25.46



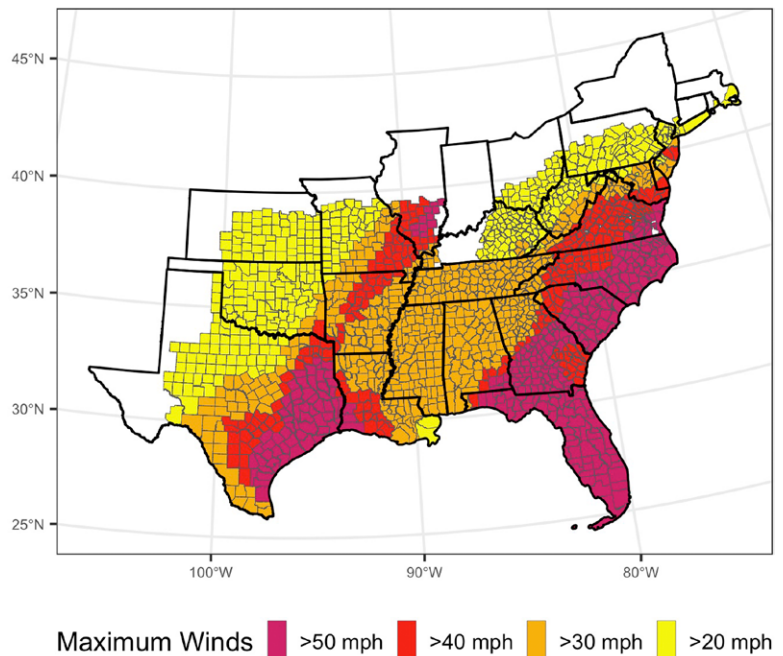
per capita for increased forecast accuracy in these affected areas. These results are also shown in Table 1. In the most exposed areas (historical wind speed > 50 mph) the total annual WTP is \$376 million, \$428 million, and \$327 million for storm track, wind speed, and precipitation forecast, respectively. Extending the WTP estimates to less exposed areas increases these estimates to \$1.4 billion, \$1.6 billion, and \$1.3 billion, respectively. On a per capita basis, the WTP for improvements along the forecast attributes range between \$9.85 and \$13.30 per person per year. Finally, projecting these WTPs over a 10-year time horizon, and discounting at 2% (Drupp et al. 2018), indicates a present value of improvements between \$90.25 and \$121.86 per person.

To better interpret the practical significance of these results, it is useful to contrast them with the current tax load devoted to fund meteorological products. Using 2018 as a baseline, the federal weather enterprise report estimates the total meteorological funding to be \$5.3 billion, which is about 0.081% of the \$6.6 trillion total federal obligations in fiscal year (FY) 2018 (OFCM 2020). Given that the median income household (\$63,179) paid about \$4,320 in taxes for FY 2018, about \$3.49 of that amount can be attributed to meteorological services and research. Out of this total, however, only a fraction is specifically allocated toward hurricane forecasting. Therefore, and when compared to this fiscal allocation, our results highlight that (i) hurricane forecast is perceived to be valuable, (ii) the perceived value greatly exceeds the operational cost of running the current programs, and (iii) further improvements are also valuable and present the regulator with a potentially sizable upside when it comes to funding research efforts.

## Discussion

We analyze the perceived valuation of hurricane forecast improvements across space and forecast attributes. We document the plausible policy argument for why the public may assign value to such improvements, and then test for the existence of that value. While we find that improvements in forecast accuracy are strictly positively valued, improvements in wind speed forecasts are consistently the most valued.

This result is perhaps related to the way in which respondents process different sources of forecast information. Wind speed is directly related to the well-known Saffir–Simpson category of a storm, which is both single-dimensional and crudely related to the damages associated with a given hurricane (Murnane and Elsner 2012). Arguably, this measure allows an individual to rapidly assess the potential danger of a storm and engage in adaptive behavior accordingly. If true, the decision-maker’s relative ease in thinking about wind speed is an availability effect, and as such, it is important even if wind speed is not objectively the most threatening hurricane attribute to human life. Moreover, Saffir–Simpson categories are often used in the media as well, so individuals are likely more familiar with that index. Track and precipitation forecasts, on the other hand, are not as effective in describing the potential



**Fig. 6. Past hurricane exposure.** This map displays the maximum wind speed experienced, due to hurricanes, in the United States between 2006 and 2018. Wind speed is in miles per hour (mph), and the unit of observation is a county.

damage of a storm, and it has previously been documented that many people may not even understand how to interpret these forecast products in the first place (Broad et al. 2007).

Besides documenting the public's perception of further hurricane forecast improvement, this study also demonstrates the potential for further interdisciplinary collaborations in hurricane research. In addition to being the first large-scale and multilocation contingent valuation of forecast improvements, it demonstrates the value of these improvements by integrating key insights from both atmospheric science and the stated preferences literature. We believe these interdisciplinary efforts are important to the body of research regarding hurricane forecasts (Lazo and Waldman 2011; Nguyen et al. 2013; Ahsan et al. 2020; Martinez 2020), and a valuable contribution to the public debate regarding the overall value of publicly funded science.

Our results, however, are not free of caveats. First, we use observed forecast errors of track and wind speed during the prior decade (2008–18) to inform what forecast errors could be a decade in the future. Specifically, we assume that further improvement continues along the historical path. The implication is that forecasts made in the future will not only be more accurate than those made in 2018, but will also have less uncertainty surrounding them. While the choice of a specific decade as the baseline is arbitrary, we assert that all three scenarios are plausible for both track and intensity and that the results would not differ noticeably if some other length of time was used to construct the error reduction values.

Second, forecast attributes are treated separately and ranked differently, but all three of them are intertwined in reality. For example, if hurricane research looking at internal processes, air–sea interactions, or sensitivity to wind shear leads to an improvement to track forecasts, it should also lead to better wind speed and rainfall forecasts. These linkages suggest that joint improvements are likely to follow in research efforts, and thus can compound the value of improvements across different attributes; these compounded values will not be additively separable. Accordingly, adding the estimates across forecast attributes can be considered as the upper bound for joint improvements. In other words, efforts resulting in improvements across multiple attributes are likely to be even more valuable than these individual estimates, but less than their respective sums.

Third, there could be concerns regarding the geographical span of our sample respondents and their relatively recent experience with Hurricanes Florence and Michael. Our choice of locations was on purpose. By targeting areas in which respondents are already familiar with the National Hurricane Center forecast products, we ensure that they are able to give a more informed reply when presented with the hypothetical improvement scenarios from the atmospheric models. To avoid potential bias, we explicitly control for respondents' having experienced and suffered damage from the storms. Economic theory (Letson et al. 2007) suggests that if better information allows them to take better decisions in terms of adaptive behavior in the face of a hurricane, then there would be an associated value assigned to having an improved forecast. Our results provide quantitative evidence that this relationship exists. Other potential issues and robustness checks are covered in appendix B.

Finally, the magnitudes of our WTP estimates exceed those of Lazo and Waldman (2011), who calculate WTP of \$15 (adjusted for inflation) per household per year for hurricane forecast improvement. This difference is large enough to raise concerns regarding which estimates should be considered when evaluating public policy related to hurricane research. We assert that this discrepancy follows from methodological differences and timing. In particular, Lazo and Waldman (2011) implement a two-step choice experiment with a relatively small number of respondents from a small area of South Florida at the beginning of the implementation of the Hurricane Forecast Improvement Project (HFIP). Our study directly builds on their seminal effort, but it seeks to improve some of the shortcomings in their study (i.e., sample size, geographical span, and understanding of improvement implications). The timing of the study is also important; by 2020, 8 of the past 10 hurricane seasons have produced an

above-average number of named storms, and the public is likely aware of this trend. It is possible that our estimates are also capturing a meaningful increase in awareness that translates into the perceived higher value of hurricane forecast improvement.

Bearing in mind the points above, the results in the analysis are consistent across multiple specifications and robustness checks, with noticeable implications for policy making; namely, the public values further hurricane forecast improvement, even after the remarkable progress observed since the start of the HFIP in 2008. This result is encouraging, and highlights the relevance of the ongoing efforts to make the hurricane forecast even more accurate. Nonetheless, our results also raise questions regarding the adequacy of the mandated standards that focus on track and wind speed (Gall et al. 2013). While justifiable in political discourse, it is unclear if such a strict goal would pass a cost–benefit analysis for optimal allocation of public resources. This analysis sheds a light on this problem and provides estimates that can be considered for such evaluation.

**Acknowledgments.** We thank Andrea Schumacher and Frank Marks for support with atmospheric modeling. We also thank Gina Eosco and participants of the Weather Program Office’s Weather Economic Research Workshop for feedback on earlier stages of the project. We are grateful to James Hammit, Daniel Herrera-Araujo, and Christopher Parmeter for guidance and support with the statistical analysis. Finally, we thank the three anonymous reviewers whose comments and suggestions helped improve this manuscript. This project was funded by the National Oceanic and Atmospheric Administration (NOAA) through Grant NA15OAR4320064. The statements, findings, and conclusions in this work are solely those of the authors and do not necessarily reflect the views of NOAA.

**Data availability statement.** Survey, data, and scripts to replicate the analysis are available at [https://github.com/renatomolinah/cv\\_hurricane\\_forecast](https://github.com/renatomolinah/cv_hurricane_forecast).

## Appendix A: Theory and empirics

This section illustrates the core of our analysis, and follows previously established literature (Carson and Hanneman 2005; Letson et al. 2007; Lazo and Waldman 2011). Let  $u(\mathbf{z}|f, h)$  be the utility function of an individual that enjoys a consumption bundle  $\mathbf{z}$  in the face of a hurricane with a given forecast accuracy  $f$  and probability of occurrence  $h$ . A rational utility maximizing individual will choose bundle  $\mathbf{z}$ , so as to maximize her utility subject to her budget  $y$ . Let  $v(\mathbf{p}, y|f, h)$  denote the indirect utility of the individual under price vector  $\mathbf{p}$ ;  $u(\mathbf{z}|f, h)$  is increasing and quasi-concave in  $\mathbf{z}$ , which implies  $v(\mathbf{p}, y|f, h)$  is decreasing in  $\mathbf{p}$  and increasing in  $y$ .

Let  $f^0$  and  $f^1$  be two different forecast accuracy levels, such that  $f^0 < f^1$ . The dollar value of the change in  $f$ ,  $w$ , is then given by

$$v(\mathbf{p}, y|f^0, h) = v(\mathbf{p}, y - w|f^1, h) \quad (\text{A1})$$

Therefore, willingness to pay (WTP) for improving from  $f^0$  and  $f^1$  can be written as  $w(f^0, f^1, \mathbf{p}, y)$ . Let  $m(\mathbf{p}, u, f, h)$  be the expenditure function for the direct utility function  $u(\mathbf{z}|f, h)$ . It follows that the expenditure function is increasing in  $u$ , nondecreasing, concave and homogenous of degree 1 in  $\mathbf{p}$ . Implicitly, this formulation assumes improvements are desired, so the expenditure function is also decreasing in  $f$ . The implication is that  $m(\mathbf{p}, u, f, h) > 0$  for any  $f$ , and that  $w < y$ .

Depending on the structural assumptions on  $u(\mathbf{z}|f, h)$ ,  $w(f^0, f^1, \mathbf{p}, y)$  could be derived in several different ways (Carson and Hanneman 2005). We will assume that the individual WTP for a better forecast,  $w(f^0, f^1, \mathbf{p}, y)$ , can be represented by

$$w_i(f^0, f^1, \mathbf{p}, y) = \mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i, \quad (\text{A2})$$

with  $\mathbf{x}_i$  as a vector of observables at the individual level. Parameter  $\varepsilon$  is a zero-mean idiosyncratic random component, and it is additive to the difference in indirect utility.

To establish this WTP, we implement a double-bounded dichotomous choice design (Hanemann et al. 1991). In our design, each respondent is presented with a randomly selected hurricane forecast attribute and a randomly selected potential rate of improvement, relative to the progress experienced between 2008 and 2018. For each of these dimensions and their respective rates of improvement, each individual  $i$  is presented with a first bid,  $b_i^1 \sim \mathcal{U}[1,50]$ . We use a uniform distribution following recommendations for a continuous distribution of bids from Boyle et al. (1988) and Lewbel et al. (2011). The support is chosen to be between \$1 and \$50 following the suggestion from Kanninen (1993) for a wide range of bids, and drawing from previous estimates available from Lazo et al. (2009), Lazo and Waldman (2011), and Anaman et al. (1998), and our own calculations from the pretest.

Depending on the respondent's answer to the bid, she would be presented with a follow-up bid,  $b_i^2$ . If the answer in the first round is positive (i.e., she accepts the additional tax burden on her household), the bid is then increased by 20%. If the answer in the first round is negative, the bid is decreased by 20% instead. The 20% adjustment is arbitrary, but because the bids are randomly generated over a wide range, there is enough variation in the bid structure that the precision benefits of the design can still be captured.

It follows that for each forecast attribute, a respondent would then fall into one of four possible scenarios. Let  $Y_i^j \in \{0,1\}$  denote the individual response for bid  $j = \{1, 2\}$ , and  $\mathbf{Y}_i = [Y_i^1, Y_i^2]$  the tuple representing her response to both questions for a given forecast attribute. Further, suppose that for individual  $i$ , the willingness to pay for a certain rate of improvement is given by

$$\text{WTP}_i(\mathbf{x}_i) = \mathbf{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \quad (\text{A3})$$

with  $\mathbf{x}_i$  as the vector including the order in which the forecast attribute is shown to the respondent and the rate of improvement, along with all other individual observables. Furthermore, let  $\varepsilon_i \sim N(0, \sigma^2)$ . The four possible scenarios, as a function of the individual survey responses, are then given by

$$\Pr(\mathbf{Y}_i = [0,0] | \mathbf{x}_i) = 1 - \Phi\left(\frac{b_i^2 - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\right), \quad (\text{A4})$$

$$\Pr(\mathbf{Y}_i = [0,1] | \mathbf{x}_i) = \Phi\left(\frac{\mathbf{x}_i^\top \boldsymbol{\beta} - b_i^2}{\sigma}\right) - \Phi\left(\frac{\mathbf{x}_i^\top \boldsymbol{\beta} - b_i^1}{\sigma}\right), \quad (\text{A5})$$

$$\Pr(\mathbf{Y}_i = [1,0] | \mathbf{x}_i) = \Phi\left(\frac{\mathbf{x}_i^\top \boldsymbol{\beta} - b_i^1}{\sigma}\right) - \Phi\left(\frac{\mathbf{x}_i^\top \boldsymbol{\beta} - b_i^2}{\sigma}\right), \text{ and} \quad (\text{A6})$$

$$\Pr(\mathbf{Y}_i = [1,1] | \mathbf{x}_i) = \Phi\left(\frac{\mathbf{x}_i^\top \boldsymbol{\beta} - b_i^1}{\sigma}\right). \quad (\text{A7})$$

Finally, the log-likelihood function for parameters  $\boldsymbol{\beta}$  and  $\sigma$  is characterized as

$$\begin{aligned} \mathcal{L}(\boldsymbol{\beta}, \sigma | \mathbf{x}_i) = & +\mathbb{1}_{\mathbf{Y}_i=[0,0]} \ln \left[ 1 - \Phi\left(\frac{b_i^2 - \mathbf{x}_i^\top \boldsymbol{\beta}}{\sigma}\right) \right] + \mathbb{1}_{\mathbf{Y}_i=[0,1]} \ln \left[ \Phi\left(\frac{\mathbf{x}_i^\top \boldsymbol{\beta} - b_i^2}{\sigma}\right) - \Phi\left(\frac{\mathbf{x}_i^\top \boldsymbol{\beta} - b_i^1}{\sigma}\right) \right] \\ & + \mathbb{1}_{\mathbf{Y}_i=[1,0]} \ln \left[ \Phi\left(\frac{\mathbf{x}_i^\top \boldsymbol{\beta} - b_i^1}{\sigma}\right) - \Phi\left(\frac{\mathbf{x}_i^\top \boldsymbol{\beta} - b_i^2}{\sigma}\right) \right] + \mathbb{1}_{\mathbf{Y}_i=[1,1]} \ln \left[ \Phi\left(\frac{\mathbf{x}_i^\top \boldsymbol{\beta} - b_i^1}{\sigma}\right) \right]. \end{aligned} \quad (\text{A8})$$

The estimates for the parameters of interest,  $\hat{\boldsymbol{\beta}}$  and  $\hat{\sigma}$ , maximize Eq. (A8). Further, recall from Eq. (A3) that  $E[\text{WTP}_i | \mathbf{x}_i] = \mathbf{x}_i^\top \boldsymbol{\beta}$ . The estimate for the average willingness to pay is then given by

$$\widehat{\text{WTP}} = \bar{\mathbf{x}}^T \hat{\boldsymbol{\beta}}, \quad (\text{A9})$$

with  $\bar{\mathbf{x}}$  as the vector of mean values for the order in which the forecast attribute is presented, the rate of improvement, and all other observables for a given respondent. Our strategy is then to perform this analysis for each individual hurricane forecast attribute (i.e., track, wind speed and precipitation). Specifically, our regression model follows Eq. (A3), and can be written as

$$\begin{aligned} \text{WTP}_i(\mathbf{x}_i) = & \beta_0 + \beta_1 \text{ORDER}_i + \beta_2 \text{RATE}_i + \beta_3 \text{INCOME}_i + \beta_4 \text{FLORENCE}_i + \beta_5 \text{FEM}_i \\ & + \beta_6 \text{EXP}_i + \beta_7 \text{EVAC}_i + \beta_8 \text{VOICE}_i + \beta_9 \text{ACTION}_i + \beta_{10} \text{LRISK}_i \\ & + \beta_{11} \text{AGE}_i + \beta_{12} \text{OWNER}_i + \beta_{13} \text{TENURE}_i + \beta_{14} \text{SRISK}_i + \beta_{15} \text{HURR}_i \\ & + \beta_{16} \text{FEMA}_i + \beta_{17} \text{NFIP}_i + \beta_{18} \text{DAM}_i + \beta_{19} \text{SIZE}_i + \beta_{20} \text{CDIST}_i + \varepsilon_i. \end{aligned} \quad (\text{A10})$$

Parameter  $\beta_0$  is the constant of the model. ORDER and RATE are the order in which the attribute is shown to the response (from 1 to 3), and its respective rate of improvement (−20%, no change, or +20% relative to the improvement rate from 2008 to 2018). INCOME is the mean income in thousands of dollars of the respondent’s zip code. FLORENCE is an indicator variable that takes a value of 1 if the respondent is from the Florence sample (this variable is dropped when working with individual samples), and FEM is an indicator variable for when the respondent identifies herself as a female. EXP is an indicator variable representing when respondents declare having experienced either Florence or Michael, while EVAC is an indicator variable that takes a value of 1 if they report having evacuated. VOICE and ACTION are categorical variables taking values from 1 to 5 depending on the respondent’s beliefs that the survey will be considered by the authorities and if the results will lead to actual policy change, respectively. LRISK is the perceived chance that the respondent will experience a hurricane in the next 10 years. These variables are what we define in the analysis as Control Set 1.

Moreover, AGE is the age of the respondent in years. OWNER and TENURE are indicator variables in case the respondent currently owns her residence, and for how long she has lived there, respectively. SRISK is the perceived chance that the respondent will experience a hurricane in the next 5 years. HURR, FEMA, and NFIP are categorical variables taking values from 1 to 5 depending on the respondent’s familiarity with hurricane, Federal Emergency Management Agency (FEMA) and National Flood Insurance Program (NFIP) insurance options, respectively. DAM is binary variable that indicates if the respondent experienced damages due to the storm. SIZE is the household size in number of individuals, and CDIST is the distance from the coast of the respondents’ zip code centroid. These variables are what we define in the analysis as Control Set 2.

## Appendix B: Technical discussion

This section covers details of the analysis that were not explicitly discussed in the main text, including robustness checks included in the supplemental material. Our elicitation device is a web-based survey targeting populations affected by Hurricanes Florence and Michael. Table ES1 shows the breakdown of responses collected, as well as the filters implemented to ensure quality responses. Our average qualified response rate is 49%.

The summary statistics for the qualified answers are shown in Table ES2. The table shows no statistically significant differences between the unconditional mean for the referendum answers across samples. The Florence sample has a higher income, 5% more participants that self-identify as female, and about 10% fewer participants that experienced the storm. Respondents for Florence are also less confident that their responses will be considered or lead to actual policy changes. Respondents in Florence also have a higher average long-term risk perception. All of these differences are significant at the 0.01% (two-sided *t* test). Respondents evacuated at a similar rate (about 18% for both samples). In addition, respondents in Florence

are on average 5 years older, show a higher rate of ownership, report higher average short-term (5-yr) risk perception, and a higher average hurricane insurance awareness. Florence also has a lower fraction of respondents reporting having experienced damages due to the storm, and report a lower household size in average. All of these differences are statistically significant at least at the 95% confidence level (two-sided *t* test). There are no detectable differences between the samples in terms of awareness of FEMA and NFIP.

Using these data, our study elicits the willingness to pay (WTP) for further forecast improvement through a double-bounded dichotomous choice experiment. The breakdown of responses that allows us to estimate this model is shown in Tables ES3 and ES4 in the supplemental material. In particular, they present the number of responses that fit any of the four possible outcomes (i.e., no/no, no/yes, yes/no, and yes/yes). The tables show that respondents were less likely to say yes when the bids were higher, and that about 14%–17% of the responses are bounded by the initial bids.

The results of the maximum likelihood estimation are shown in Tables ES5–ES7 for storm track, wind speed, and precipitation, respectively. The results show some heterogeneity of magnitude and significance across attributes, but two results are worthy of attention. First, the ORDER coefficient is always significant and negative, which suggests respondents drawing from a common household budget exhibit decreasing marginal willingness to pay from multiple improvements. In other words, multiple and sequential improvements, while still valuable, generate less of a perceived benefit.

Second, the RATE coefficient is not significant for any specification in any of the attributes. This lack of statistical significance could raise concerns as it indicates respondents are not responsive to the degree of improvement shown to them. In the contingent valuation literature, this result is known as failure of the scope test (Arrow et al. 1993). While originally conceived as a (quasi)necessary condition for the validity of contingent valuation studies, Smith and Osborne (1996) and Heberlein et al. (2005) point out that failure of scope tests can occur for reasons due to the contextual nature of the good/service in question, which can also be consistent with psychological and economic theory. The more problematic possibilities are that respondents are not understanding the implications of the improvement rates, or that they are not paying attention to the questions when prompted (Giguere et al. 2020), or that respondents simply do not satisfy the assumptions of rational behavior required for proper elicitation (Hammitt and Herrera-Araujo 2018).

The survey is designed to maximize the chance of respondents fully understanding the implications of further hurricane improvements. Specifically, we provide respondents with detailed explanations of the forecast attributes, the improvement on record, the practical implications, and model-based projections that they could contrast when deciding their response. Providing this level of detail sacrifices our ability to ask more questions, but we deem this trade-off necessary to gather informed responses. In addition, all responses underwent a quality check to ensure thoughtful answers.

We also implement a latent class analysis à la Hammitt and Herrera-Araujo (2018) to explore the presence of unobserved classes that might be affecting the coefficients for RATE in the main analysis. This estimation is provided in the supplemental material and it incorporates additional follow-up questions in our survey to establish class membership. The results suggest that unobserved classes may be behind the lack of significance for the coefficient, but the aggregate estimates are still consistent with the ones provided in the main analysis. In light of these results and the steps taken in the survey to ensure informed responses, we interpret this result as evidence that respondents value improvement and the rate of improvement separately. This result is consistent with previous findings in the literature (Heberlein et al. 2005).

Proceeding with the analysis, we take the maximum likelihood estimates and project the average WTP by multiplying the statistically significant coefficients of the estimation by the

relevant sample averages. These calculations are shown in Table ES8. Extrapolations are then performed using the fully specified model. The implications and interpretation of these results are in the main text, but below we cover other potential sources of concern.

Contingent valuation studies are susceptible to several potential sources of bias. One possibility is hypothetical bias, which is the lack of consistency between stated and revealed preferences. In the survey, we include an oath-type truth-telling commitment question asking respondents to swear that they will provide thoughtful and honest answers in the survey. This approach relies on Jacquemet et al. (2013), who find that utilizing oath as a truth-telling commitment can be effective in eliciting true preferences for nonmarket goods and it can substantially reduce hypothetical bias. Respondents are also asked to consider their budget constraints and alternative use of that money. Earlier contingent valuation studies show that these reminders can also help further mitigate hypothetical bias (Penn and Hu 2019).

Another issue relates to consequentiality, or the lack of trust that responses will be considered by the relevant authorities. Previous results suggest that truthful preference revelation in repeated dichotomous choice studies, such as ours, is possible when participants view their decisions as having a weak chance of influencing policy (Vossler et al. 2012). We include survey questions that capture the consequentiality aspect of the contingent valuation study, and they consistently come up as a significant and positive factor associated with WTP (Vossler and Watson 2013; Carson et al. 2014).

Finally, we have to address the potential bias generated by working with a nonrepresentative sample. For this purpose, we follow Mozumder et al. (2011) and use representative values from the U.S. American Community Survey (2018) instead of sample means to estimate WTPs. These results are shown in Fig. ES1 and indicate no meaningful differences between both estimations. Based on this result, we assert that our results are not significantly biased by the lack of representativeness of the sample. Additional robustness checks, including using different elicitation methods, income levels, and additional controls are presented in the supplemental material and attest for the robustness of our results to different specifications and assumptions.

## References

- Adams, B., and K. Judd, 2016: 2030 agenda and the SDGs: Indicator framework, monitoring and reporting. *Global Policy Watch*, No. 10, Global Policy Forum, New York, NY, 5 pp., [www.2030agenda.de/sites/default/files/GPW10\\_2016\\_03\\_18.pdf](http://www.2030agenda.de/sites/default/files/GPW10_2016_03_18.pdf).
- Ahsan, M. N., A. Khatun, M. S. Islam, K. Vink, M. Oohara, and B. S. Fakhruddin, 2020: Preferences for improved early warning services among coastal communities at risk in cyclone prone south-west region of Bangladesh. *Prog. Disaster Sci.*, **5**, 100065, <https://doi.org/10.1016/j.pdisas.2020.100065>.
- Anaman, K. A., S. C. Lellyett, L. Drake, R. J. Leigh, A. Henderson-Sellers, P. F. Noar, P. J. Sullivan, and D. J. Thampapillai, 1998: Benefits of meteorological services: Evidence from recent research in Australia. *Meteor. Appl.*, **5**, 103–115, <https://doi.org/10.1017/S1350482798000668>.
- Arrow, K., R. Solow, P. R. Portney, E. E. Leamer, R. Radner, and H. Schuman, 1993: Report of the NOAA panel on contingent valuation. *Fed. Regist.*, **58**, 4601–4614.
- Boyle, K. J., M. P. Welsh, and R. C. Bishop, 1988: Validation of empirical measures of welfare change. *Land Econ.*, **64**, 94–98, <https://doi.org/10.2307/3146613>.
- Broad, K., A. Leiserowitz, J. Weinkle, and M. Steketeetee, 2007: Misinterpretations of the “cone of uncertainty” in Florida during the 2004 hurricane season. *Bull. Amer. Meteor. Soc.*, **88**, 651–668, <https://doi.org/10.1175/BAMS-88-5-651>.
- Carson, R. T., and W. M. Hanneman, 2005: Contingent valuation. *Handbook of Environmental Economics*, Vol. 2., K. G. Mäler and J. R. Vincent, Eds., Elsevier, 821–936, [https://doi.org/10.1016/S1574-0099\(05\)02017-6](https://doi.org/10.1016/S1574-0099(05)02017-6).

- , T. Groves, and J. A. List, 2014: Consequentiality: A theoretical and experimental exploration of a single binary choice. *J. Assoc. Environ. Resour. Econ.*, **1**, 171–207, <https://doi.org/10.1086/676450>.
- DeMaria, M., and J. Kaplan, 1994: A Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic basin. *Wea. Forecasting*, **9**, 209–220, [https://doi.org/10.1175/1520-0434\(1994\)009<0209:ASHIPS>2.0.CO;2](https://doi.org/10.1175/1520-0434(1994)009<0209:ASHIPS>2.0.CO;2).
- , J. A. Knaff, R. Knabb, C. Lauer, C. R. Sampson, and R. T. DeMaria, 2009: A new method for estimating tropical cyclone wind speed probabilities. *Wea. Forecasting*, **24**, 1573–1591, <https://doi.org/10.1175/2009WAF2222286.1>.
- Drupp, M. A., M. C. Freeman, B. Groom, and F. Nesje, 2018: Discounting disentangled. *Amer. Econ. J. Econ. Policy*, **10**, 109–134, <https://doi.org/10.1257/pol.20160240>.
- Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436**, 686–688, <https://doi.org/10.1038/nature03906>.
- Ewing, B. T., J. B. Kruse, and D. Sutter, 2007: Hurricanes and economic research: An introduction to the Hurricane Katrina symposium. *South. Econ. J.*, **74**, 315–325.
- Gaddis, E. B., B. Miles, S. Morse, and D. Lewis, 2007: Full-cost accounting of coastal disasters in the United States: Implications for planning and preparedness. *Ecol. Econ.*, **63**, 307–318, <https://doi.org/10.1016/j.ecolecon.2007.01.015>.
- Gall, R., J. Franklin, F. Marks, E. N. Rappaport, and F. Toepfer, 2013: The Hurricane Forecast Improvement Project. *Bull. Amer. Meteor. Soc.*, **94**, 329–343, <https://doi.org/10.1175/BAMS-D-12-00071.1>.
- Giguere, C., C. Moore, and J. C. Whitehead, 2020: Valuing hemlock woolly adelgid control in public forests: Scope effects with attribute nonattendance. *Land Econ.*, **96**, 25–42, <https://doi.org/10.3368/le.96.1.25>.
- Gladwin, H., J. K. Lazo, B. H. Morrow, W. G. Peacock, and H. E. Willoughby, 2007: Social science research needs for the hurricane forecast and warning system. *Nat. Hazards Rev.*, **8**, 87–95, [https://doi.org/10.1061/\(ASCE\)1527-6988\(2007\)8:3\(87\)](https://doi.org/10.1061/(ASCE)1527-6988(2007)8:3(87)).
- Hammit, J. K., and D. Herrera-Araujo, 2018: Peeling back the onion: Using latent class analysis to uncover heterogeneous responses to stated preference surveys. *J. Environ. Econ. Manage.*, **87**, 165–189, <https://doi.org/10.1016/j.jeem.2017.06.006>.
- Hanemann, M., J. Loomis, and B. Kanninen, 1991: Statistical efficiency of double-bounded dichotomous choice contingent valuation. *Amer. J. Agric. Econ.*, **73**, 1255–1263, <https://doi.org/10.2307/1242453>.
- Heberlein, T. A., M. A. Wilson, R. C. Bishop, and N. C. Schaeffer, 2005: Rethinking the scope test as a criterion for validity in contingent valuation. *J. Environ. Econ. Manage.*, **50**, 1–22, <https://doi.org/10.1016/j.jeem.2004.09.005>.
- Jacquemet, N., R.-V. Joule, S. Luchini, and J. F. Shogren, 2013: Preference elicitation under oath. *J. Environ. Econ. Manage.*, **65**, 110–132, <https://doi.org/10.1016/j.jeem.2012.05.004>.
- Kanninen, B. J., 1993: Optimal experimental design for double-bounded dichotomous choice contingent valuation. *Land Econ.*, **69**, 138–146, <https://doi.org/10.2307/3146514>.
- Landsea, C. W., and J. P. Cangialosi, 2018: Have we reached the limits of predictability for tropical cyclone track forecasting? *Bull. Amer. Meteor. Soc.*, **99**, 2237–2243, <https://doi.org/10.1175/BAMS-D-17-0136.1>.
- Lazo, J. K., and D. M. Waldman, 2011: Valuing improved hurricane forecasts. *Econ. Lett.*, **111**, 43–46, <https://doi.org/10.1016/j.econlet.2010.12.012>.
- , R. E. Morss, and J. L. Demuth, 2009: 300 billion served: Sources, perceptions, uses, and values of weather forecasts. *Bull. Amer. Meteor. Soc.*, **90**, 785–798, <https://doi.org/10.1175/2008BAMS2604.1>.
- Letson, D., D. S. Sutter, and J. K. Lazo, 2007: Economic value of hurricane forecasts: An overview and research needs. *Nat. Hazards Rev.*, **8**, 78–86, [https://doi.org/10.1061/\(ASCE\)1527-6988\(2007\)8:3\(78\)](https://doi.org/10.1061/(ASCE)1527-6988(2007)8:3(78)).
- Lewbel, A., D. McFadden, and O. Linton, 2011: Estimating features of a distribution from binomial data. *J. Econ.*, **162**, 170–188, <https://doi.org/10.1016/j.jeconom.2010.11.006>.
- Lonfat, M., R. Rogers, T. Marchok, and F. D. Marks Jr., 2007: A parametric model for predicting hurricane rainfall. *Mon. Wea. Rev.*, **135**, 3086–3097, <https://doi.org/10.1175/MWR3433.1>.
- Marks, F. D., and Coauthors, 1998: Landfalling tropical cyclones: Forecast problems and associated research opportunities. *Bull. Amer. Meteor. Soc.*, **79**, 305–323, [https://doi.org/10.1175/1520-0477\(1998\)079<0305:LTCFPA>2.0.CO;2](https://doi.org/10.1175/1520-0477(1998)079<0305:LTCFPA>2.0.CO;2).
- , B. D. McNoldy, M.-C. Ko, and A. B. Schumacher, 2020: Development of a probabilistic tropical cyclone rainfall model: P-rain. *Tropical Meteorology and Tropical Cyclones Symp.*, Boston, MA, Amer. Meteor. Soc., 367310, <https://ams.confex.com/ams/2020Annual/webprogram/Paper367310.html>.
- Martinez, A. B., 2020: Forecast accuracy matters for hurricane damage. *Econometrics*, **8**, 18, <https://doi.org/10.3390/econometrics8020018>.
- Mozumder, P., W. F. Vásquez, and A. Marathe, 2011: Consumers' preference for renewable energy in the southwest USA. *Energy Econ.*, **33**, 1119–1126, <https://doi.org/10.1016/j.eneco.2011.08.003>.
- , A. G. Chowdhury, W. F. Vásquez, and E. Flugman, 2015: Household preferences for a hurricane mitigation fund in Florida. *Nat. Hazards Rev.*, **16**, 04014031, [https://doi.org/10.1061/\(ASCE\)NH.1527-6996.0000170](https://doi.org/10.1061/(ASCE)NH.1527-6996.0000170).
- Murnane, R., and J. Elsner, 2012: Maximum wind speeds and us hurricane losses. *Geophys. Res. Lett.*, **39**, L16707, <https://doi.org/10.1029/2012GL052740>.
- Nguyen, T. C., J. Robinson, S. Kaneko, and S. Komatsu, 2013: Estimating the value of economic benefits associated with adaptation to climate change in a developing country: A case study of improvements in tropical cyclone warning services. *Ecol. Econ.*, **86**, 117–128, <https://doi.org/10.1016/j.ecolecon.2012.11.009>.
- OFCEM, 2020: The Federal Weather Enterprise: Fiscal year 2020 budget and coordination report. Tech. Rep. FCM-R38-2020, 38 pp., [www.icams-portal.gov/publications/fedrep/2021\\_fedrep.pdf](http://www.icams-portal.gov/publications/fedrep/2021_fedrep.pdf).
- Penn, J., and W. Hu, 2019: Cheap talk efficacy under potential and actual hypothetical bias: A meta-analysis. *J. Environ. Econ. Manage.*, **96**, 22–35, <https://doi.org/10.1016/j.jeem.2019.02.005>.
- Quiring, S. M., A. B. Schumacher, and S. D. Guikema, 2014: Incorporating hurricane forecast uncertainty into a decision-support application for power outage modeling. *Bull. Amer. Meteor. Soc.*, **95**, 47–58, <https://doi.org/10.1175/BAMS-D-12-00012.1>.
- Regnier, E., 2008: Public evacuation decisions and hurricane track uncertainty. *Manage. Sci.*, **54**, 16–28, <https://doi.org/10.1287/mnsc.1070.0764>.
- Smith, V. K., and L. L. Osborne, 1996: Do contingent valuation estimates pass a “scope” test? A meta-analysis. *J. Environ. Econ. Manage.*, **31**, 287–301, <https://doi.org/10.1006/jeem.1996.0045>.
- Stewart, S., and R. Berg, 2019: National Hurricane Center Tropical Cyclone Report: Hurricane Florence (AL062018). NOAA/National Weather Service, 98 pp., [www.nhc.noaa.gov/data/tcr/AL062018\\_Florence.pdf](http://www.nhc.noaa.gov/data/tcr/AL062018_Florence.pdf).
- Trumbo, C. W., L. Peek, M. A. Meyer, H. L. Marlatt, E. Gruntfest, B. D. McNoldy, and W. H. Schubert, 2016: A cognitive-affective scale for hurricane risk perception. *Risk Anal.*, **36**, 2233–2246, <https://doi.org/10.1111/risa.12575>.
- Vossler, C. A., and S. B. Watson, 2013: Understanding the consequences of consequentiality: Testing the validity of stated preferences in the field. *J. Econ. Behav. Organ.*, **86**, 137–147, <https://doi.org/10.1016/j.jebo.2012.12.007>.
- , M. Doyon, and D. Rondeau, 2012: Truth in consequentiality: Theory and field evidence on discrete choice experiments. *Amer. Econ. J. Microecon.*, **4**, 145–171, <https://doi.org/10.1257/mic.4.4.145>.
- Weinkle, J., C. Landsea, D. Collins, R. Musulin, R. P. Crompton, P. J. Klotzbach, and R. Pielke, 2018: Normalized hurricane damage in the continental United States 1900–2017. *Nat. Sustainability*, **1**, 808–813, <https://doi.org/10.1038/s41893-018-0165-2>.