A Spatial Analysis of Decisions Made in Response to Simulated Tornado Warnings in the United States

TAYLOR A. DEWINTER-MACIAG[a,b] AND RENEE A. MCPHERSON[a,c]

[a] *Department of Geography and Environmental Sustainability, University of Oklahoma, Norman, Oklahoma*
[b] *Cooperative Institute for Severe and High Impact Weather Research and Operations, University of Oklahoma, Norman, Oklahoma*
[c] *USGS South Central Climate Adaptation and Science Center, Norman, Oklahoma*

(Manuscript received 15 November 2022, in final form 10 August 2023, accepted 21 September 2023)

ABSTRACT: Although decision-making in response to tornado warnings is well researched, most studies do not examine whether individual responses to these warnings vary across different geographical locations and demographic groups. This gap is addressed by using data from a decision experiment that places participants virtually in a simulated tornado warning and asks them to minimize the costs of their decisions. The authors examine the following: 1) what demographic attributes may contribute to choices to minimize costs to protect assets at a specific location in a tornado warning, 2) whether there is a spatial component to how these attributes influence decision-making, and 3) if there are specific U.S. regions where individuals do not make protective decisions that minimize their overall cost. Multilevel regression analysis and poststratification are applied to data from the simulated decision experiment to estimate which demographic attributes and National Weather Service County Warning Areas are most associated with the costliest protective decisions. The results are then analyzed using spatial autocorrelation to identify spatial patterns. Results indicate that sex, race, and ethnicity are important factors that influence protection decisions. Findings also show that people across the southern portions of the United States tend to make more costly protective decisions, as defined in this work.

SIGNIFICANCE STATEMENT: Tornadoes, although rare, threaten both life and property. Studies have shown that certain demographic groups are more negatively impacted by disasters than others and are at higher risk of severe weather hazards. We ask if there are demographic characteristics or geographic locations in common among people who are more prone to making protection decisions during tornado warnings to minimize economic costs. Results can help warning providers, such as the National Weather Service, direct resources and education to specific types of decision-makers or locations to improve sheltering decisions.

KEYWORDS: Social Science; Decision making; Societal impacts

1. Introduction

Annually, the United States averages more than 1000 tornadoes (NCEI 2020). Tornado outbreaks on a large spatial scale or those with violent tornadoes are rare but account for most deaths and injuries (Ćwik et al. 2021; Simmons and Sutter 2011). For example, the 25 April 2011 Super Outbreak across the southern United States, the EF5 tornado on 22 May 2011 in Joplin, Missouri, and the EF5 tornado on 20 May 2013 in Moore, Oklahoma, caused 316, 158, and 47 direct fatalities, respectively (NOAA 2011a,b, 2014). These events also can be costly, with the former outbreak assessed at $13.7 billion (in 2023 dollars; NCEI 2023). Postevent assessments by the National Weather Service (NWS) documented that although these events had good forecasts with adequate lead times, residents were reluctant to personalize the threat and seek appropriate shelter (NOAA 2011a,b, 2014). As a result, the National Oceanic and Atmospheric Administration (NOAA) and its NWS continue to seek a better understanding of how people make decisions to protect themselves and their property in the face of tornado events (National Weather Service 2019; Uccellini and Ten Hoeve 2019).

In response to this need, researchers have studied decision-making under uncertainty as related to tornadic events (e.g., Slovic 1987; Nagele and Trainor 2012; Joslyn and LeClerc 2012; Durage et al. 2016), yet most studies do not examine how living in a different NWS County Warning Area (CWA) may affect responses. Because tornado climatology varies across the United States, it is possible that people's perceptions of tornado risk and their resulting decision-making processes may vary as well. Any differences need to be recognized so that local NWS forecasters, media, and emergency responders can focus their messaging more effectively. The goal of this research is to enhance the current knowledge regarding if and how tornado warning decision-making varies across the United States and if certain demographic groups respond differently to a warning.

We briefly summarize some of the relevant literature about decision-making during natural hazards in section 2, concluding the section with our research hypotheses. Section 3 documents the data and methods applied, and section 4 overviews the study results. Sections 5 and 6 discuss the relevance of the results and summarize the study, respectively.

Denotes content that is immediately available upon publication as open access.

Corresponding author: Taylor DeWinter-Maciag, tdewinter@ou.edu

DOI: 10.1175/WCAS-D-22-0133.1

2. Background

Risk and decision-making during disasters has been researched extensively resulting, in part, through several theoretical frameworks (e.g., Mileti and Sorensen 1990; Lindell and Perry 2012). In these frameworks, individuals are exposed to a threat (e.g., tornado), gather information, and analyze their situation to determine their risk before they respond to the threat. An important step for achieving an appropriate response, such as taking protective action, is for the individual to believe the threat is real and to personalize the risk (i.e., believe the threat could happen *to* them). To determine their own risk, a person may use social interactions, individual attitudes, environmental cues, cultural knowledge, or their sense of place (Kasperson et al. 1988; Masuda and Garvin 2006). Many of these factors are related to where a person lives and, thus, can vary geographically.

Prior research has analyzed geographic differences in tornado hazard exposure or tornado climatology (e.g., Boruff et al. 2003; Ashley 2007; Brotzge et al. 2011; Ashley and Strader 2016), tornado risk (e.g., Strader et al. 2021), weather salience (Stewart et al. 2012), and social vulnerability (e.g., Cutter et al. 2003; Cutter and Finch 2008; Emrich and Cutter 2011). For example, Boruff et al. (2003) assessed the frequency and location of tornadoes across the contiguous United States that resulted in a human fatality or injury or in reported economic damage (from 1950 to 1999). They found that tornado hazard density was highest east of the 100th meridian except for New England and portions of Appalachia. In a spatial and temporal analysis of tornado fatalities, Ashley (2007) documented that a region encompassing the lower Arkansas, Tennessee, and lower Mississippi River valleys had the largest concentration of tornadoes associated with fatalities from 1950 to 2004. However, the region was not the center of the greatest overall frequency of all tornadoes or significant (F2+) tornadoes; instead, that area was located over much of Oklahoma and parts of southern Kansas and northern Texas. The mid-South region (i.e., most of Alabama and Mississippi, eastern portions of Arkansas and Louisiana, southern Tennessee, and western Georgia) had the highest tornado disaster potential, or annual chance an EF1+ tornado path would intersect with developed lands, according to research by Ashley and Strader (2016).

Strader et al. (2021) examine tornado risk, exposure, and vulnerability combinations by NWS CWA. They found substantial differences among many CWAs as related to the offices' false alarms, unwarned tornadoes, and lead time for tornado warnings as well as tornado climatology (e.g., nonfatal and fatal tornado frequencies, EF1+ path density, and fatalities per capita). Similarly, Harrison and Karstens (2017) demonstrated that the mean and maximum numbers of tornado warnings issued per convective day varied by CWA, with higher values from CWAs in the lower Mississippi River valley. Using data from 2000 to 2004, Brotzge et al. (2011) documented geographical variation of false alarm ratios across four regions of the contiguous United States—West, Plains, Midwest/East, and Southeast—with the Midwest/East and Southeast having significantly higher false alarm ratios than the West or Plains. On the other hand, Chamberlain et al. (2023) found that geographic location did not vary strongly among the NWS Central, Southern, and Eastern regions as related to the probability of detection of the first, middle, or last tornado in a tornado outbreak. Neither Brotzge et al. (2011) nor Chamberlain et al. (2023) examined these warning performance statistics on the CWA scale, however.

Strader et al. (2021) also documented differences among the CWAs in four social vulnerability themes: 1) socioeconomic status (e.g., persons living in poverty, unemployment, and income per capita), 2) household composition and disability (e.g., persons over the age of 65 or under the age of 17, single-parent households, disabled persons), 3) minority status and language (e.g., persons who are non-White or speak English "less than well"), and 4) housing type and transportation (e.g., mobile and manufactured housing, multiunit structures, households with no vehicle). Because of these differences across CWAs, Strader et al. (2021) recommended that individual NWS forecast offices develop their own climatologies of tornado risk and assessments of exposure and vulnerability as well as enhance training modules according to this CWA-specific knowledge. Finally, they encouraged additional research that connects environmental factors to potential impacts on people and the tornado warning process.

Studies also have compared geographic distributions of risk reception, perception, and response to tornadoes based on tornado climatology. For example, Ripberger et al. (2020) and Allan et al. (2020) analyzed survey data from the Severe Weather and Society Survey (WxSurvey, administered through the University of Oklahoma's Institute for Public Policy Research and Analysis) to test whether survey responses related to the respondent's location. Initiated in 2017, WxSurvey asks recurring questions annually to measure individuals' reception, comprehension, and responses to tornado warnings and other natural hazards (tropical weather, floods, etc.) (Ripberger et al. 2019). Ripberger et al. (2020) found that, on average, reception, comprehension, and response to tornadoes were highest in the central United States, where exposure to and experience with tornadoes are generally high. Allan et al. (2020) also analyzed weather hazards and found that risk perception and response correlated with hazard climatology. Tornado risk perception was also highest in the central United States, matching the areas with highest exposure to tornadoes.

Stewart et al. (2012) surveyed 1465 participants across the United States to compare the degree to which people are psychologically attuned to and affected by weather and weather changes, defined as "weather salience." Participants who lived in temperate (mesothermal) or cold/continental (microthermal) climates had higher weather salience than those from dry (arid/semiarid) climates. Those living in dry climates were comparatively less attentive to weather information than those living in the other two climate regions when it might result in a delay, cancellation, or holiday from work or school. Participants in dry climates also sought weather forecasts for fewer time periods than those in temperate or cold/continental climates.

Studies also demonstrate that emotional connection to a location, or place attachment, can influence decision-making.

An individual's attachment to a place, such as their home, church, or favorite walking trail, influences how they perceive risk in that place (Masuda and Garvin 2006). For example, Klockow et al. (2014) interviewed residents of Alabama and Mississippi after the April 2011 tornado outbreak, documenting that local knowledge of place (e.g., nearby river, tree-cleared construction area) by the residents influenced their risk perception and resulting actions. Peppler et al. (2018) noted that residents of central Oklahoma perceived their tornado risk to be different than others' risk based on their knowledge of the locations of a river, higher elevations, Native American burial grounds, a major interstate highway, prior tornado paths, and the urban heat island. These and other studies (e.g., Masuda and Garvin 2006; Donner et al. 2012; Paul et al. 2015; Jauernic and Van Den Broeke 2016; Sanders et al. 2020) have highlighted that individuals incorporate their understanding of place with tornado warning information to evaluate their risk and make decisions.

Research hypotheses

Although these studies have highlighted important geospatial patterns related to tornadoes, their impacts, and people's perceptions of tornado risk, they have not examined decision-making in response to tornado warnings across different regions of the United States at the scale of NWS CWAs. Based on this prior research and gap in knowledge, this study asks the following question: How do protective decisions that are prompted by a tornado warning vary spatially by NWS County Warning Area? A CWA is the multicounty jurisdiction (usually covering only a part of a state or states) of a single NWS Weather Forecast Office. NWS forecasters from a single office warn for tornadoes and other hazards in their CWA according to their local data, knowledge and experience of the warning issuer, office policies, and other input. Hence, the processes that result in a tornado warning may vary by CWA. Thus, the public's response decisions may vary across CWAs.

We pose the following hypotheses:

- There will be differences in the decisions made by individuals from different demographic groups as related to protecting an asset during a tornado warning.
- These differences will contribute to warning-related decision-making that varies geographically across the contiguous United States.
- There will be clusters of NWS CWAs where warning-related decision-making will be more similar than other clusters.

Null results will be important to inform the NWS that nationally consistent warning graphics, similar to those used in this study, should result in geographically and demographically consistent decision responses. Significant results will highlight geographic locations or demographic groups that the NWS can prioritize in their tornado warning outreach activities, providing more equitable services across the nation. In addition, if there are clusters of CWAs that are more alike in decision response than other CWAs, there is an opportunity for them to work together or to share successful outreach materials with each other.

3. Data and methods

a. Simulated tornado-warning experiment

Although there are a variety of ways to answer these questions, this study applied data from a nationwide experiment designed to study "potential influences on subjective estimates of threat within the geospatial context of a tornado warning" (Klockow-McClain et al. 2020; see this citation for a full description of the experimental context). We chose to use this dataset because it spanned the contiguous United States, was census-balanced by several demographic attributes (i.e., education, household income, age, and sex), had a large sample size ($n = 5564$), and was quality assured, consistent, and well documented. It also focused on the decision context of a tornado warning and was readily accessible to the author team. Limitations of the dataset, particularly as related to the hypotheses, are addressed in section 5d.

A web-based survey provided a simulated decision context whereby each participant was randomly and evenly assigned one of six tornado warning displays for any given trial in the experiments (Fig. 1). The six displays comprised two deterministic warnings (i.e., one with a 30-min lead time and one with a 1-h lead time) and four probabilistic warnings (i.e., 1-h lead time with either no color or sequential, divergent, or spectral color themes). In the deterministic forecasts, visually resembling realistic warnings issued by the NWS, the forecast probability of being affected by a tornado was 100% somewhere within the polygon and zero percent outside of the polygon. For the probabilistic warnings, the corresponding probability for one of the four sections of the polygon was written next to that section (Fig. 1). For any given warning graphic appearing on the participant's screen during a trial, they were randomly assigned a location A–D (dots in Fig. 1) where they were hypothetically in charge of sheltering aircraft.

During each trial, the participant would choose whether to protect/shelter the aircraft at their location or to not do anything. After the choice was made, the computer calculated whether a tornado occurred based on the probability at the participant's assigned location (using random sampling without replacement over multiple trials). If the participant chose to protect their aircraft, they would obtain a hit (H) when the computer generated a tornado and a false alarm (F) when no tornado occurred. If they chose not to protect, they gained a miss (M) if a tornado occurred and a null (N) when there was not a tornado. Associated costs were assigned accordingly ($H = -\$3{,}000$, $F = -\$3{,}000$, $M = -\$6{,}000$, $N = \$0$). This experimental design was a modification of a traditional cost–loss decision problem (Thompson 1952) that has been applied to other studies of weather forecasting (e.g., Murphy 1976; Katz and Murphy 1997; Howard et al. 2021).

b. Data

1) SIMULATED TORNADO RESPONSE

Each participant engaged in 96 total decision trials (i.e., through computer-generated distributions of graphics types, airport locations, and tornado probabilities). The nationwide survey included 5564 participants, with about 100 respondents

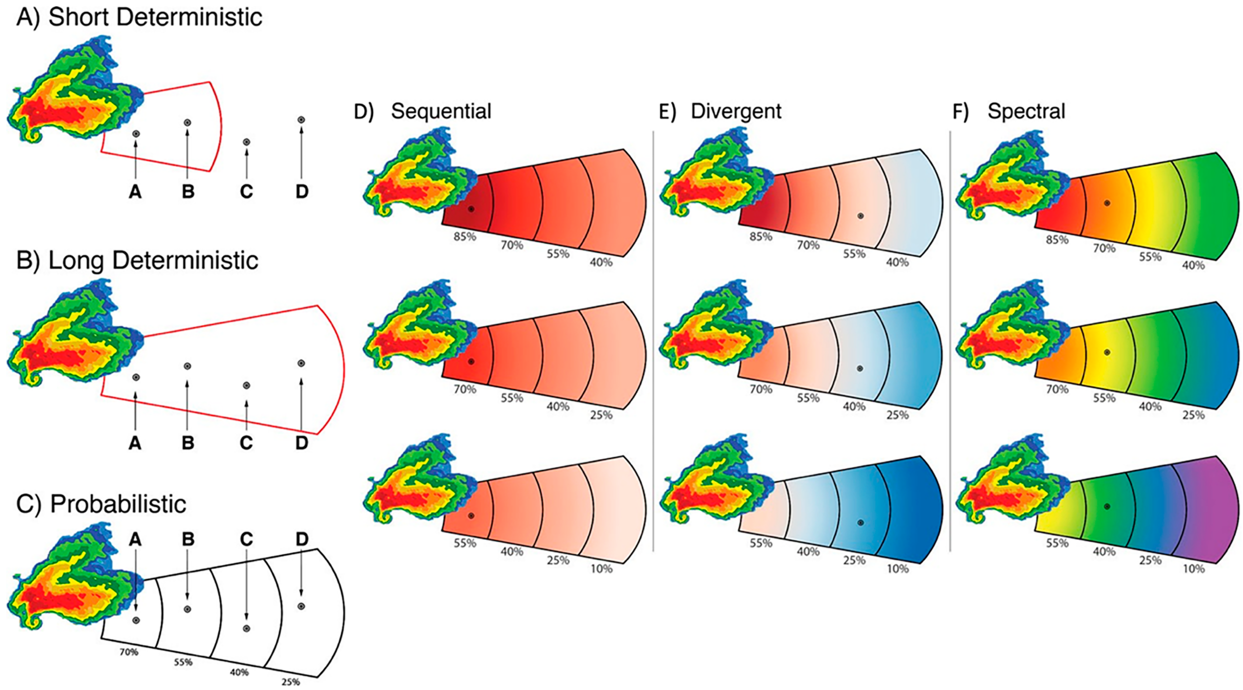

FIG. 1. Cartographic designs of tornado warnings from Klockow-McClain et al. (2020) showing two deterministic and four probabilistic (including no color and three color schemes) graphics. The current study uses this nationwide decision experiment.

from each state (Klockow-McClain et al. 2020), although only 4461 participants completed all 96 trials. We removed those participants who did not provide their geographic location as well as those from CWAs in Alaska and Hawaii (outside our study domain), resulting in 4331 participants. Demographic and socioeconomic information included sex, age, race, ethnicity, household income, employment status, education, and the state and county of residence. This information was optional for all participants. Most participants chose not to include their income, employment status, or education; thus, to maintain an adequate sample size, we omitted these variables from the following analysis.

We assumed no carryover effects from one trial to the next; thus, each trial was considered independent. For any given trial that incurred a tornado, it was costlier to not protect (*M*; loss of $6,000) than to protect (*H*; cost of $3,000). For a trial that did not incur a tornado, it was costlier to protect (*F*; cost of $3,000) than to not protect (*N*; $0). Hence, the costliest decisions were those that resulted in *M* and *F*. The independent variable for our study was the proportion of costliest decisions (PCDs) over the 96 trials, or PCD = $(M + F)/96$.

We sorted participant PCD values by CWAs because the sample size was not large enough for county- or climate division-level groupings. CWAs are more similar in size across the contiguous United States than states, and tornado warnings are issued within CWA boundaries rather than state boundaries. Also, prior literature on warnings and other forecasts has focused on CWA as the geographical unit of data aggregation and research (e.g., Harrison and Karstens 2017; Henderson et al. 2020; Strader et al. 2021).

2) TORNADO FREQUENCY AND CENSUS DATA

To determine if tornado exposure was related to decision-making during the simulation, we used the mean number of tornadoes per year by CWA from NOAA's Storm Events Database. Figure 2 depicts the event frequency in percentiles

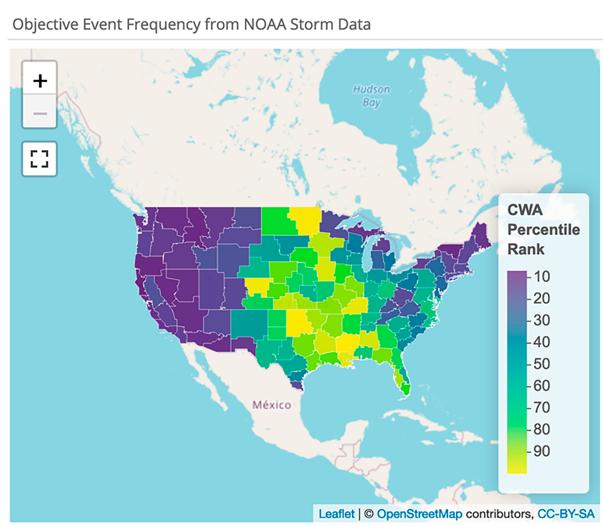

FIG. 2. Mean numbers of tornadoes per CWA by percentile rank. Blues and purples (greens and yellows) indicate lower (higher) percentiles. Data from the NOAA Storm Event Database (NCEI 2022). Image generated in WxDash, an online tool of the University of Oklahoma's Institute for Public Policy Research and Analysis (https://crcm.shinyapps.io/WxDash/#section-risk-perceptions).

to rank each CWA compared to others. Note that the CWAs in the central and southern United States have the highest percentile rank for tornado frequency.

Finally, U.S. census data aggregated by CWA were provided by Dr. J. Ripberger [2020 personal communication, used in Ripberger et al. (2020)] from the University of Oklahoma's Institute for Public Policy Research and Analysis (OU IPPRA). The census data originated from the U.S. Census Bureau website for the years 2010–18—the latest years with complete census data records available at the time of this study. The dataset included county resident population estimates by age, sex, race, and Hispanic ethnicity.

c. *Methods*

We used multilevel regression analysis and poststratification (MRP) 1) to analyze the demographic of the participants who were associated with higher PCD during the tornado warning experiment and 2) to determine whether these participants were more likely to be located within certain CWAs. We applied spatial autocorrelation to examine if there were spatial patterns where PCDs were most prevalent. Finally, two-sample Student's *t* tests were performed to determine whether the demographic groups identified as having higher PCDs were more likely to overestimate or underestimate their risk.

1) MRP

Multilevel regression analysis and poststratification is a type of small-area estimation (SAE) that has been applied to downscale data from larger to smaller population characteristics (Gelman et al. 1997; Lax and Phillips 2009; Ripberger et al. 2020). SAE "is a statistical technique used to produce statistically reliable estimates for smaller geographic areas than those for which the original surveys were designed" (Zhang et al. 2015). SAE techniques have been applied to predict voting patterns in relatively small regions (e.g., states or voting districts) using data from, for example, national surveys of political opinion. MRP has been applied in election forecasting, voting patterns, and other areas in political sciences primarily (Hanretty 2020). In the past several years, its use has been expanded to study how people respond to extreme weather or climate change (e.g., Howe et al. 2015; Zhang et al. 2018; Howe et al. 2019; Allan et al. 2020; Ripberger et al. 2020; Howe et al. 2023).

MRP uses individual characteristics and survey responses based on location to estimate what the broader response of the public may be in geographic units (Buttice and Highton 2013). Spatial characteristics of the dataset are retained in cases where the region's sample size is sufficiently large. In this case, MRP models responses in the region as a function of both demographic and geographic variables (Lax and Phillips 2009), using data from the broader sample. To do so, MRP applies a multilevel regression to create a preference estimate for each type of person in the broader dataset and then weights those preferences by population frequency within the smaller region (Buttice and Highton 2013). Hence, MRP uses data from the full experiment to predict how different types of people respond to the surveys and applies that information, as needed, in predefined geographic units with an insufficient sample size.

For our study, the geographic units are CWAs, the independent variables are the demographic characteristics of the participants, and the dependent variable is their PCD over the 96 trials, or $(M + F)/96$. Hence, the participant's PCD varies as a function of their demographic attributes (sex, age group, ethnicity, and race) and their CWA as follows:

$$y_i = \beta^0 + \alpha^{\text{sex}}_{j[i]} + \alpha^{\text{age}}_{k[i]} + \alpha^{\text{ethnicity}}_{l[i]} + \alpha^{\text{race}}_{m[i]} + \alpha^{\text{CWA}}_{s[i]}, \qquad (1)$$

where

$$\alpha^{\text{sex}}_j \sim N(0, \sigma^2_{\text{sex}}), \quad j = 0 \text{ or } 1,$$

$$\alpha^{\text{age}}_k \sim N(0, \sigma^2_{\text{age}}), \quad k = 1, 2, \text{ or } 3,$$

$$\alpha^{\text{ethnicity}}_l \sim N(0, \sigma^2_{\text{ethnicity}}), \quad l = 0 \text{ or } 1,$$

$$\alpha^{\text{race}}_m \sim N(0, \sigma^2_{\text{race}}), \quad m = 1, 2, \text{ or } 3, \quad \text{and}$$

$$\alpha^{\text{CWA}}_s \sim N(\text{tornado exposure}_s, \sigma^2_{\text{CWA}}), \quad s = 1, \ldots, 116.$$

In Eq. (1), β_0 is the intercept and α is the offset (Buttice and Highton 2013), where each α classifies the characteristics of the participants in a given number of categories (i.e., j, k, l, m, n, and s). For instance, α^{sex}_j represents sex in two categories in the decision experiment: female (level 0) and male (level 1). Age groups use three categories from the U.S. census: ages 18–34 (level 1), ages 35–59 (level 2), and ages 60–110 (level 3). The youngest and oldest participants were 18 and 82 years old, respectively. Ethnicity is either non-Hispanic (0) or Hispanic (1) and race is White (1), African American/Black (2), or other (i.e., not White and not African American/Black) (3). CWA has 116 categories that are associated with Albuquerque, New Mexico (ABQ) (1), to Las Vegas, Nevada (VEF) (116), numbered alphabetically, for CWAs in the lower 48 states.

The regression model assessed how much influence each of the independent variables had on the PCDs, and the model was applied to predict the PCDs made by specific types of people. The multilevel regression model predicts outcomes based only on the participant's responses, their CWA (i.e., geographic location), and their demographic characteristics. To better represent the population of the entire CWA, poststratification includes the demographics in the U.S. census for each CWA to make the prediction, generating weighted values (θ) for each demographic–geographic combination (r) (Buttice and Highton 2013; Allan et al. 2020; Ripberger et al. 2020). The frequencies (N) and the weights (θ) are used to calculate the MRP estimates for each CWA:

$$Y^{\text{MRP}}_{\text{CWA}} = \frac{\sum\limits_{r \in \text{CWA}} N_r \theta_r}{\sum\limits_{r \in \text{CWA}}}, \qquad (2)$$

where θ_r are the weighted predictions from the multilevel regression model output for each demographic–geographic combination and N_r are the population frequencies (from U.S. census data) for each demographic–geographic combination.

The results from the MRP analysis will output the estimated mean PCD for each CWA. Thus, MRP compares results from each CWA as individual units and ignores any potential between-neighbor (i.e., neighboring CWAs) correlations (Xu 2014). MRP addresses sampling limitations found in many internet-based surveys that do not have truly randomized samples (e.g., from nonresponse bias) or do not have sufficient sampling of various demographic groups at subnational scales (e.g., Downes et al. 2018).

2) Spatial autocorrelation

We apply spatial autocorrelation to identify if there are between-neighbor correlations or spatial patterns in the dataset. Spatial autocorrelation compares a datum from one spatial location, such as a CWA, to the datum from its "nearest neighbor" (i.e., adjacent CWA) to determine if they are related to each other. A variety of phenomena, such as epidemiology for disease patterns, economic geography for crime patterns, and ecology for migration, have been assessed using spatial autocorrelation (Ord and Getis 1995, 2001). Here, we apply spatial autocorrelation using the Moran's I and Getis–Ord statistics to identify patterns, if any, across CWAs.

The Moran's I statistic is perhaps the most popular method to test spatial autocorrelation (Ord and Getis 1995; Anselin 1995). We constructed a spatial weights matrix using Queen nearest neighbors and then applied the matrix in the Moran's I calculation [Eq. (3)] to test whether the CWA-averaged PCD in a CWA was spatially correlated with any of its nearest neighbors. The equation is as follows:

$$I = \frac{N\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \overline{X})(x_j - \overline{X})}{\left(\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\right)\sum_{i=1}^{n}(x_i - \overline{X})^2}, \tag{3}$$

where N is the number of observations, in this case, the number of CWAs; $\overline{X}$ is the mean of the PCD scores across all CWAs; x_i is the average of the PCDs for all individual respondents within a particular CWA; x_j is the average of the PCDs for all individual respondents at a different CWA; and w_{ij} is the spatial weight from CWA$_i$ to CWA$_j$ (Cao 2014).

To determine if there were CWAs where participants had higher PCDs than surrounding CWAs, we calculated the Getis–Ord statistic to identify spatial clusters, or hotspots, of high values of PCDs. The term G_i measures the degree of association across all values at all locations (Getis and Ord 2010) using only values in its neighborhood, as follows:

$$G_i = \frac{\sum_{j\neq i} w_{ij}x_j}{\sum_{j\neq i} x_j}, \tag{4}$$

where w_{ij} is the spatial weight from CWA$_i$ to CWA$_j$ and x_j is the value of a neighboring CWA. A positive (negative) G_i denotes a spatial cluster of high (low) values, that is, a hot (cold) spot (Anselin 2021).

3) Student's t test

We conducted two-sampled t tests [W. S. Gosset writing as Student (1908)] to ascertain whether differences in PCDs by demographic group resulted from an overestimation (more false alarms) or underestimation (more misses) of their risk. We tested whether the following pairs of populations were statistically different: 1) male ($n = 2475$) versus female ($n = 1856$), 2) non-Hispanic ($n = 3934$) versus Hispanic ($n = 397$), 3) White ($n = 3460$) versus African American/Black ($n = 372$), and 4) White ($n = 3460$) versus other races ($n = 499$). Here, p values ≤ 0.05 indicate significance.

4. Results

The first research question asks if there are demographic attributes that contribute to the proportion of costliest decisions. To answer this question, the first step in MRP is creating a multilevel model that will show if there are significant differences among demographic groups, indicating there are certain types of people who are associated with higher PCDs. The second question determines if these attributes contribute to decision-making that varies geographically across the United States. The poststratification step in MRP is used to map predicted proportions of costliest decision scores across the United States according to census data. The map highlights where there are CWAs with populations that are associated with higher PCDs. The final question asks if there are regions of the United States that have similar proportions of costliest decisions. To answer this question, spatial autocorrelation of the PCD scores is used to determine if there are broader regions, or hotspots, in the United States where people may be more prone to making costlier sheltering decisions.

a. *MRP results*

The multilevel regression model was developed to determine if there were significant differences among demographic groups to address our first hypothesis that there will be differences in the decisions made by individuals from different demographic groups. In this case, we pooled data for all participants, regardless of their CWA and the graphic type that were assigned (see Fig. 1). Table 1 displays the summary statistics: the intercept (expected outcome), an estimate of the amount each demographic group varied from the intercept in their likelihood to have higher PCDs, the standard error of this estimate (measured using the mean absolute deviation), and the t value from a two-sample t test. Positive (negative) estimate values reveal that the PCDs increased (decreased), and larger absolute values of the estimates highlight large differences compared to another category in the same classification (e.g., female compared to male in the sex classification). Values near zero indicate little difference between categories. Significance was defined as the 95% confidence interval (i.e., $|t| > 1.96$).

For all participants regardless of the graphic type, demographics, and location, the mean PCD (i.e., combined misses and false alarms over their 96 trials) was 0.397. Males tended to make fewer costly decisions than did females, and this difference was significant ($-0.009\,81$, $t = -5.292$; Table 1). For

TABLE 1. Summary statistics for the multilevel regression analysis; an asterisk (*) indicates 95% CI → $|t| > 1.96$ is significant.

Mean change in the ratio of the proportion of costliest decisions to all decisions from the intercept			
	Estimated difference from intercept = 0.397 06	Standard error	t value
Sex: Group 1 (female) to group 2 (male)	−0.009 81	0.001 85	−5.292*
Age: Group 1 (18–34) to group 2 (35–59)	0.001 36	0.002 07	0.659
Age: Group 1 (18–34) to group 3 (60–110)	−0.004 37	0.002 59	−1.684
Ethnicity: Group 1 (non-Hispanic) to group 2 (Hispanic)	0.008 72	0.003 25	2.685*
Race: Group 1 (White) to group 2 (African American/Black)	0.019 00	0.003 30	5.757*
Race: Group 1 (White) to group 3 (other)	0.007 94	0.002 97	2.670*
Tornado exposure	0.000 28	0.000 92	0.310

age groups, however, differences were not significant. Those participants who were 35–59 years old had higher PCDs (0.001 36, $t = 0.659$), in general, than did those ages 18–34. Conversely, participants who were 60–110 years old had lower (−0.004 37, $t = -1.684$) PCDs than did participants ages 18–34. Results for ethnicity and race both yielded statistically significant results. Participants who identified as Hispanic had more combined misses and false alarms than those of non-Hispanic ethnicity (0.008 72, $t = 2.685$). Similarly, participants who identified as either African American/Black (0.019 00, $t = 5.757$) or other (0.007 94, $t = 2.670$) had higher PCDs than their White counterparts. Finally, we examined whether tornado exposure affected decision-making in this experiment, finding that it had little impact on the PCDs (0.000 28, $t = 0.310$). Thus, our results show that individuals who are female, Hispanic, African American/Black, or other races had higher proportions of costliest decisions, confirming our first hypothesis.

The results from the MRP analysis were then poststratified into CWAs to address our second hypothesis (warning-related decision-making will vary geographically). Figure 3 displays the predicted PCD scores for the multilevel model, highlighting locations in dark red where there are higher PCD scores. Our results show that there appears to be areas of geographic differences in participants' costliest protection decisions. Mainly, these areas crossed southern portions of the United States, including Southern California, the south-central United States, the southeastern United States, and the Atlantic coast. These differences appeared relatively minor, however—a CWA minimum of 0.391 and CWA maximum of 0.401—and not statistically significant in comparison with the median ($p = 0.087 89$). The differences amount to about one decision during the 96 trials. This result was unexpected, as prior research had shown that geographic location affected risk perception, reception, comprehension,

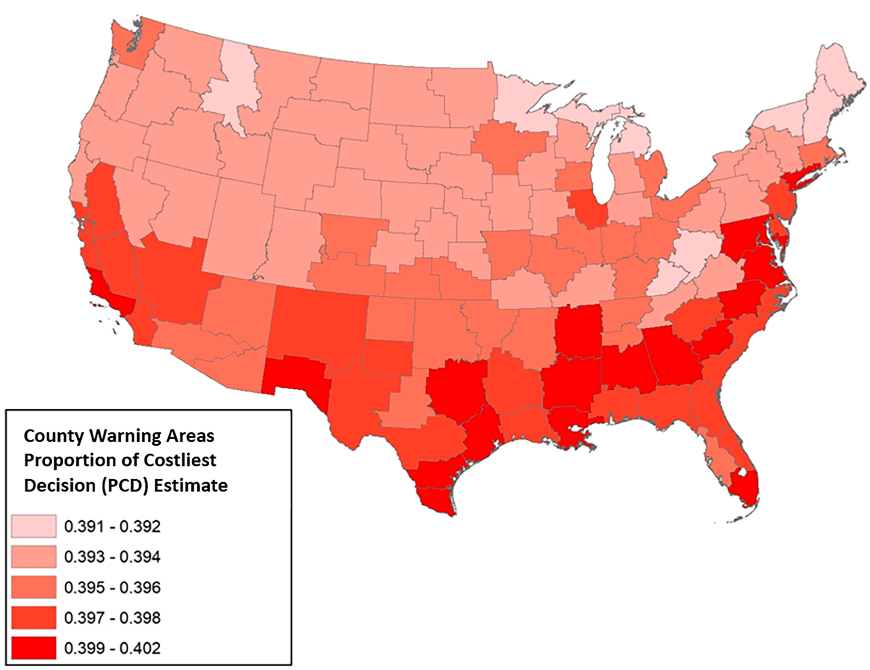

FIG. 3. Predicted PCD estimates for the multilevel model applying poststratification. Darker reds represent higher predicted PCD scores.

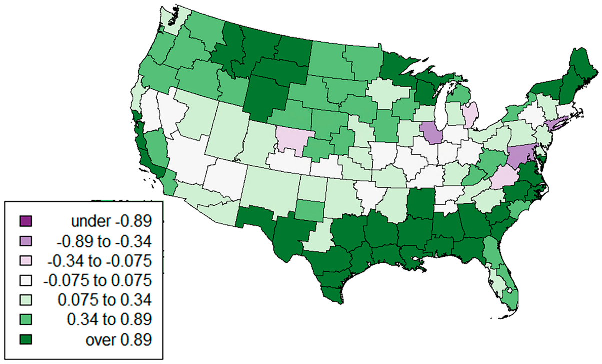

FIG. 4. Local Moran's I statistic for model predictions. Greens indicate positive values of the local Moran's I (similarity) and purples indicate negative values of the local Moran's I (dissimilarity). If I were positive (negative), then it is surrounded by CWAs with similar (different) values. The closer I is to 1 or -1 indicates the magnitude of the relationship, just like R^2 indicates the strength of a linear relationship. Darker green (purple) CWAs are those more similar (dissimilar) to their neighboring CWAs.

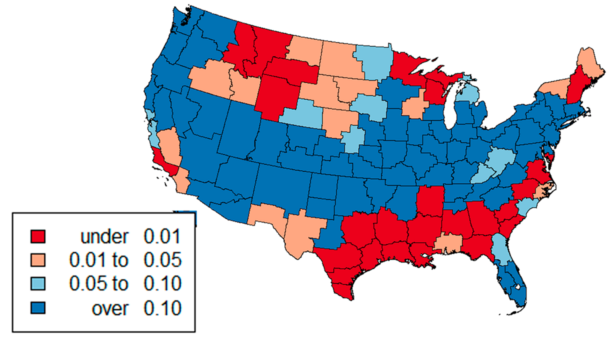

FIG. 5. The p values for local Moran's I at the 95% confidence interval. Red-shaded CWAs indicate those that are significantly related to their neighboring CWAs. Blue-shaded CWAs do not pass significance testing.

and response to tornadoes (e.g., Allan et al. 2020; Ripberger et al. 2020; Klockow et al. 2014; Peppler et al. 2018). Thus, while our second hypothesis appeared to be supported visually, a deeper look revealed that estimated tornado warning decision-making (in terms of sheltering an asset) for any single CWA does not vary significantly from other CWAs. In other words, there is not an individual CWA that "stands out" as having higher PCDs. We next examine if there are clusters of higher PCDs by comparing each CWA with their neighboring CWAs, rather than comparing them all individually.

It is important to note that the geographic patterns shown in Fig. 3 are not based solely on the demographic variables, but the combination of the demographic variables (sex, age, ethnicity, and race) and the location of the CWA for each participant. Because the MRP analysis revealed that sex, ethnicity, and race had more influence on higher PCDs, these demographic factors are weighted more heavily than age and tornado exposure and account for more of the variation.

b. *Spatial autocorrelation results*

Last, we examined spatial variations in the data in a different manner to determine if there are clusters of CWAs where warning-related decision-making is more similar (our third hypothesis). Using the Moran's I statistic on the poststratified multilevel model results, we found four prominent regions that were spatially similar (darker green shades in Fig. 4): 1) the Intermountain West and far northwest Great Plains, 2) the Lake Superior region, 3) northern New England, and 4) much of the southern United States. Figure 5 displays the p values for the Moran's I statistic, indicating that CWAs with positive spatial autocorrelation are statistically significant (red colors). These results indicate that within each of the four "regions," the people in one of the CWAs (e.g., Duluth, Minnesota, CWA) make similar decisions as their neighboring CWAs (e.g., Marquette and Green Bay, Wisconsin, CWAs). Each distinct region, however, is not necessarily similar to another region (e.g., people in Green Bay may not make the same decisions as those in Atlanta, Georgia). Calculating G_i identified a cluster for high values of the predicted PCDs in Southern California and another cluster in the south-central and southeastern United States (Fig. 6). These "hotspots" colocate with the predicted PCD estimates (Fig. 3).

c. *t-test results*

To assess the differences by demographic group (sex, ethnicity, and race) found in section 4a, we conducted t tests on the full participant sample to see if these groups were more prone to underestimate (more misses) or overestimate (more false alarms) their risk. In terms of sex, differences in the percentage of misses for males versus females were not statistically significant ($p = 0.226$); however, differences in false alarms—0.240 for males and 0.247 for females—were significant ($p = 0.013$). This result indicates that female participants in our study tended to overestimate their risk. Similarly, in the case of ethnicity, there were not statistically significant differences in the misses for Hispanics versus non-Hispanics ($p = 0.819$). Although the average percentage of false alarms for Hispanics (0.251) appeared quite different than those for

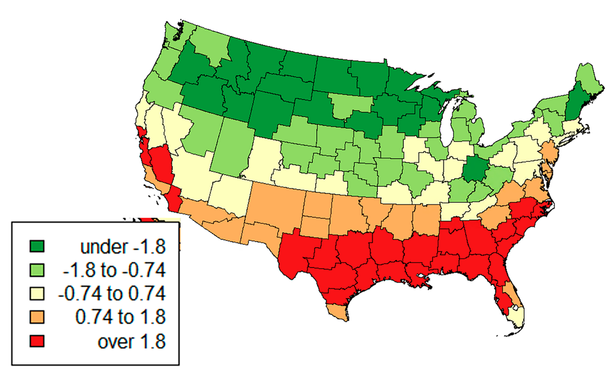

FIG. 6. The G_i statistic indicating hotspots and cold spots of the PCDs. Reds indicate hotspots, greens indicate cold spots, and yellows indicate little or no correlation with their neighboring CWAs. Thus, reds are "hotspots" of PCDs.

non-Hispanics (0.243), the p value of 0.060 was slightly too high to be deemed significant. Still, Hispanics also tended to overestimate their risk.

Interestingly, when examining race, the cause of different PCD values was not more false alarms, but more misses. Proportions of misses for White participants versus African American/Black participants were 0.149 and 0.162, respectively, and were significant ($p = 0.0002$). Proportions of misses for White versus other race participants were 0.149 and 0.156, respectively, and significant ($p = 0.029$). For false alarms, however, the differences were not significant for White versus African American/Black participants ($p = 0.155$) nor for White versus other race participants ($p = 0.476$), indicating that there was no significant difference between race groups in their false alarm proportions. Hence, participants identifying as African American/Black or other tended to underestimate their risk.

5. Discussion

a. Demographics

The results of this study show that female, Hispanic, African American/Black, or other race participants were more likely than others in their demographic classification (i.e., sex, ethnicity, race) to have higher PCDs. Previous studies have identified these demographic groups as having high social vulnerability (e.g., Cutter et al. 2003; Cutter and Finch 2008; Emrich and Cutter 2011). Hence, these groups may be more adversely impacted by tornadoes than their male, non-Hispanic, or White counterparts regardless of other demographic characteristics or environmental conditions.

Our results are consistent with those identified within Klockow-McClain et al. (2020) using the same dataset. Those authors noted that the higher average number of PCDs for these demographic groups could result from them having higher risk aversion, therefore sheltering more frequently and incurring more false alarms.

Risk aversion helps explain why some of the demographic groups we identified have higher PCDs. Our results show that females have a higher overall proportion of false alarms compared to males, consistent with other studies (e.g., Fothergill 1996; Comstock and Mallonee 2005; Nagele and Trainor 2012; Ripberger et al. 2020). For instance, females are more likely to personalize a tornado warning and take shelter (Fothergill 1996). The female participants in our experiment protected the assets more frequently regardless of cost, therefore resulting in a higher proportion of costly decisions.

Klockow-McClain et al. (2020) also found that Hispanic and African American/Black populations had a significantly higher average proportion of protection decisions. Our results indicate that while Hispanic participants were more risk averse than non-Hispanic participants, African American/Black and other race participants had higher overall proportions of missed protections. Other studies also have mixed results regarding risk aversion with race and ethnicity (e.g., Riad 1997; Nagele and Trainor 2012; West and Orr 2007; Trainor et al. 2015). For example, West and Orr (2007) found that those who identified as a racial or ethnic minority felt they were more vulnerable to hurricanes but were less likely than White individuals to evacuate due to lack of resources or situational factors (lack of vehicle, taking care of children/elderly, etc.). On the other hand, Trainor et al. (2015) could not find race to be a significant factor in sheltering decisions during tornado warnings. Because our study was a virtual decision experiment, Hispanic, African American/Black, or other race individuals who felt at risk were able to shelter or not to shelter assets without worrying about situational factors. Although West and Orr (2007) examined the perception that Rhode Island voters had about their residence's vulnerability to a major hurricane, their findings appear relevant to our study. They found that experiencing different situations resulted in a variety of responses that women and minorities would take as compared to White males; hence, these studies highlighting different responses or perceptions of women and minorities as compared to White males likely are due to the different contexts for each study. For our decision experiment, participants essentially had endless resources (because we did not "cap" their spending), so they had different constraints than those of other studies. Thus, it is not surprising that Hispanic, African American/Black, or other race individuals had higher PCDs for this study, as they did not have real-life barriers that may have affected their decision-making.

The higher PCDs from female, Hispanic, African American/Black, or other race participants also may be related to a lack of knowledge or experience with tornadoes or the misunderstanding of the warning message. Although sheltering regardless of a tornado occurring is the recommendation of the NWS, it is still important to make sure all populations understand tornado warning messaging, so they can make cost-effective sheltering decision, especially for those who chose not to protect at all. Strader et al. (2021) found differences among CWAs in tornado risk, exposure, and vulnerability that led them to recommend that NWS staff within each CWA may need to give their environment consideration and certain populations in their CWA special attention to ensure that they have adequate information and education to make appropriate warning-related decisions. For example, in CWAs with high Hispanic populations, warning information can be provided in Spanish to ensure that everyone has access to life-saving information, regardless of language (Trujillo-Falcón et al. 2021). Our results align with these prior studies and recommendations.

b. Geography

Our results highlight that participants in Southern California, the south-central United States, and the southeastern United States had higher proportions of costliest decisions than others across the contiguous United States (Figs. 3, 4, and 6), although these differences were small. Analyses using MRP showed the differences not to be significant, a somewhat surprising result because previous studies demonstrated that there are geographic differences with tornado warning risk perception, response, and comprehension (e.g., Allan et al. 2020; Ripberger et al. 2020). Our spatial autocorrelation analysis, however, did show significant results, and both analyses distinguished

Southern California and the south-central and southeastern United States as having populations making the costliest decisions than other locations across the United States. These regions have populations that are socially vulnerable (e.g., Cutter et al. 2003; Cutter and Finch 2008; Emrich and Cutter 2011; Strader et al. 2021), and the south-central and southeastern United States are more exposed to tornado hazards (Fig. 2) (e.g., Boruff et al. 2003; Harrison and Karstens 2017), have higher numbers of tornado fatalities (e.g., Ashley 2007; Ashley and Strader 2016), and have higher false alarm rates (Brotzge et al. 2011; Strader et al. 2021).

The south-central and southeastern regions collocated with regions of social vulnerability, tornado exposure, tornado fatalities, and false alarms are interesting and important to note. For instance, Strader et al. (2021) found that CWAs in the southeastern United States have higher concentrations of tornado fatalities, higher frequency of tornado watches and warnings, longer tornado warning lead times, unwarned tornado reports, and higher false alarms compared to other tornado-prone regions. Thus, people may be choosing to shelter assets regardless of cost because of their susceptibility to harm or damage from tornadoes. The south-central United States also has higher tornado frequency but lower false alarm rates (Boruff et al. 2003; Ashley 2007; Brotzge et al. 2011; Ashley and Strader 2016). It is possible, then, that forecasters have more experience issuing tornado warnings, resulting in lower overall false alarm rates. However, people in the south-central United States are still exposed to a higher frequency of tornadoes and, like their counterparts in the southeastern United States, may also choose to shelter assets regardless of cost rather than suffer potential damage.

Why might our results be different than other studies? First, Lindell and Perry (2012) explained that people's responses within an experimental environment, such as ours, may not reflect real life. Experiments do not account for situational factors, such as the environment or a person's intent. A person may intend to shelter, for example, but their current situation may prevent them from seeking adequate shelter. Likewise, our results may differ from other studies because there was no place attachment in the virtual experiment. Recall that place links a particular location to ones' experiences and worldviews, shaping their behaviors and decisions. This virtual experiment of Klockow-McClain et al. (2020) simulated tornado warnings in an arbitrary, virtual place that participants had no prior experience or perception of. Although they may have instilled some of their own experiences into their decisions, the lack of place attachment may have caused them to rely only on the information provided on the web page. Our results, therefore, might not reflect actual differences in decision-making based on demographics or CWA location.

c. *MRP versus spatial autocorrelation*

Both the MRP and spatial autocorrelation analyses identified that participants in the southern portions of the United States were more prone to making the costliest decisions than other regions (Figs. 3 and 6), but the results were only significant from the spatial autocorrelation. Recall that multilevel models ignore potential between-neighbor correlations, as they treat each CWA as its own unit rather than comparing it to its neighboring CWA (Xu 2014). Therefore, when assessing the MRP results for spatial variation, each *individual* CWA was compared with each other CWA and differences between each pair were found to be insignificant. Spatial autocorrelation, on the other hand, compares neighboring CWAs for spatial patterns. In that case, CWAs in the southern United States had significantly similar PCDs to their neighbors; there is a geographic pattern, or "hotspot," of the proportion of costliest decisions in the southern United States. Thus, we recommend that between-neighbor correlations and broader patterns be evaluated using spatial autocorrelation.

d. *Limitations*

There were several limitations of the dataset, methods, and geographic scale in this study. Although we used a high-quality and high-resolution dataset from a standardized survey, the experiment design may have been a weakness for our study. We noted that the virtual space where participants made decisions may have disassociated them from how they would make decisions in real life, such as by a lack of situational factors like environmental cues or interaction with family or friends (Lindell and Perry 2012). Also, the large number of trials ($n = 96$) that were needed to thoroughly test the cartographic designs in the original experiment of Klockow-McClain et al. (2020) may have resulted in too much complexity (e.g., too many independent variables) to answer our three hypotheses.

A great benefit of MRP for areas with small sample sizes is that it can perform cross tabulations. If there are few members of a certain demographic group in a CWA, for example, MRP can apply the same demographic group from another CWA in the dataset to account for the small sample (Lax and Phillips 2009). This process may have affected our results negatively, however. Because our intent was to examine the effect of geographic location, the application of information from another CWA may have produced unrepresentative results of the actual populations. The impact of the MRP technique should be diminished where there were adequate sample sizes of demographic groups. As Klockow-McClain et al. (2020) oversampled for Hispanic and African American/Black participants, this limitation likely did not have much impact on our results.

In addition, the geographic scale of CWAs may have been a limitation to our study. We selected CWAs because our results could then be directly useful for NWS forecasters in their already-defined jurisdiction. Also, based on the dataset, there was insufficient sampling within counties for MRP to work effectively. It is possible, however, that the CWA scale was too large to capture the nuances of decision-making. For instance, Peppler et al. (2018) found that residents of central Oklahoma had drastically different views of risk and response to tornadoes, even across neighboring counties. Future work using higher-resolution data could help determine if there is a more desired scale for this analysis to work. Still, NWS forecasters are faced with the reality of providing warnings, forecasts, and decision-support services to all people across their CWA. Hence, CWA-scale analyses are relevant to NWS staff and policymakers. We recommend that future work investigate decision-making at smaller scales to provide forecasters with localized information about the populations they serve.

6. Conclusions

This study provides an overview of geographic variations of tornado-warning decision-making as related to protecting an asset. Three hypotheses were posed in this study: 1) there will be differences in the decisions made by individuals from different demographic groups as related to protecting an asset during a tornado warning, 2) these differences will contribute to warning-related decision-making that varies geographically across the contiguous United States, and 3) there will be clusters of NWS CWAs where warning-related decision-making will be more similar than other clusters. Using MRP, four demographic groups were identified as being more prone to making the costliest decisions: female, Hispanic, African American/Black, and other races, confirming our first hypothesis. However, the differences between White and non-White populations or among CWAs were relatively small or, in some cases, statistically insignificant. Poststratification and spatial autocorrelation of the multilevel model results revealed that people located in CWAs in the southern portions of the United States have higher proportions of costliest decisions. The MRP results were not deemed significant, countering our second hypothesis. However, results from the spatial autocorrelation identified clusters of similarly high proportions of costliest decisions, supporting our third hypothesis. These CWAs were associated with areas previously shown in the literature to have high social vulnerability, high tornado exposure, tornado fatalities, and false alarm rates (in the south-central and southeastern United States). The subtle spatial variations in decision-making found in this study highlight that although decision-making is a complex, place-dependent, and individual process, there are large-scale spatial patterns that can help inform warning providers where extra efforts to communicate risk are needed.

The results of our study are relevant for decision-makers and risk communicators for several reasons. First, the demographic groups identified are also groups that previous studies indicate have higher social vulnerability. This result is important for risk communicators, such as NWS forecasters, broadcast meteorologists, or local emergency managers, to know so they can help the populations they serve. Second, the south-central and southeast United States were consistently identified as a region with higher PCDs and are important because this region has more tornado exposure, tornado fatalities, false alarm rates, and social vulnerability among its populations. Risk communicators in these areas should also be aware that their populations may not make wrong decisions, but, rather, are more risk averse. The nuance between missed protections and risk aversion should be pursued in future work, as it is important to encourage those who are risk averse to continue their sheltering behaviors while searching for better ways to communicate with those who chose not to protect.

Acknowledgments. The corresponding author would like to thank her coauthor for her support, wisdom, and guidance through this research project and graduate school. Many thanks to Dr. Kim Klockow-McClain for allowing usage of her dataset for this project, for her help in understanding her data collection and methods, and for serving on the corresponding author's master's thesis committee. Thanks to Dr. Jennifer Koch for her help and support for this research and for serving on the corresponding author's master's thesis committee. Dr. Joe Ripberger was incredibly helpful in his help understanding MRP, editing code, and providing census and NOAA tornado frequency data aggregated to CWA scale, cutting out a tremendous amount of busy work. Special thanks also to Dr. Kodi Berry for funding and support for this project. Funding was provided by the NOAA Office of Oceanic and Atmospheric Research under the NOAA–University of Oklahoma Cooperative Agreement NA16OAR4320115, the South Central Climate Adaptation Science Center under USO-122799700, National Aeronautics and Space Administration under Grant NNX15AK02H for the NASA Oklahoma Space Grant Consortium, and the University of Oklahoma's Department of Geography and Environmental Sustainability. Financial support for the submission of the manuscript was provided by the Office of the Vice President for Research and Partnerships and the Office of the Provost at the University of Oklahoma.

Data availability statement. Due to its proprietary nature, the virtual decision-making experiment data cannot be made publicly available. Further information about the dataset is included in Klockow-McClain et al. (2020). Census data are available through the United States Census Bureau at https://www.census.gov/programs-surveys/decennial-census/decade.2010.html. Storm frequency data are available through NOAA's Storm Events Database at https://www.ncdc.noaa.gov/stormevents/.

REFERENCES

Allan, J. N., J. T. Ripberger, W. Wehde, M. Krocak, C. L. Silva, and H. C. Jenkins-Smith, 2020: Geographic distributions of extreme weather risk perceptions in the United States. *Risk Anal.*, **40**, 2498–2508, https://doi.org/10.1111/risa.13569.

Anselin, L., 1995: Local indicators of spatial association—LISA. *Geogr. Anal.*, **27**, 93–115, https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.

——, 2021: Week 6b: Other local spatial autocorrelation statistics (Introduction to Spatial Data Science). The Center for Spatial Data Science, University of Chicago, accessed 14 July 2021, https://www.youtube.com/watch?v=h7J26lDheFQ.

Ashley, W. S., 2007: Spatial and temporal analysis of tornado fatalities in the United States: 1880–2005. *Wea. Forecasting*, **22**, 1214–1228, https://doi.org/10.1175/2007WAF2007004.1.

——, and S. M. Strader, 2016: Recipe for disaster: How the dynamic ingredients of risk and exposure are changing the tornado disaster landscape. *Bull. Amer. Meteor. Soc.*, **97**, 767–786, https://doi.org/10.1175/BAMS-D-15-00150.1.

Boruff, B., J. Easoz, S. Jones, H. Landry, J. Mitchem, and S. Cutter, 2003: Tornado hazards in the United States. *Climate Res.*, **24**, 103–117, https://doi.org/10.3354/cr024103.

Brotzge, J., S. Erickson, and H. Brooks, 2011: A 5-yr climatology of tornado false alarms. *Wea. Forecasting*, **26**, 534–544, https://doi.org/10.1175/WAF-D-10-05004.1.

Buttice, M. K., and B. Highton, 2013: How does multilevel regression and poststratification perform with conventional national surveys? *Polit. Anal.*, **21**, 449–467, https://doi.org/10.1093/pan/mpt017.

Cao, G., 2014: Spatial analysis and modeling (GIST 4302/5302). Dept. of Geosciences, Texas Tech University, 61 pp., https://www.depts.ttu.edu/geospatial/center/gist4302/documents/lectures/Spring%202014/lecture6.pdf.

Chamberlain, J. U., M. D. Flournoy, M. J. Krocak, H. E. Brooks, and A. K. Anderson-Frey, 2023: Analyzing tornado warning performance during individual storm life cycles. *Wea. Forecasting*, **38**, 773–785, https://doi.org/10.1175/WAF-D-22-0142.1.

Comstock, R. D., and S. Mallonee, 2005: Comparing reactions to two severe tornadoes in one Oklahoma community. *Disasters*, **29**, 277–287, https://doi.org/10.1111/j.0361-3666.2005.00291.x.

Cutter, S. L., and C. Finch, 2008: Temporal and spatial changes in social vulnerability to natural hazards. *Proc. Natl. Acad. Sci. USA*, **105**, 2301–2306, https://doi.org/10.1073/pnas.0710375105.

——, B. J. Boruff, and W. L. Shirley, 2003: Social vulnerability to environmental hazards. *Soc. Sci. Quart.*, **84**, 242–261, https://doi.org/10.1111/1540-6237.8402002.

Ćwik, P., R. A. McPherson, and H. E. Brooks, 2021: What is a tornado outbreak?: Perspectives through time. *Bull. Amer. Meteor. Soc.*, **102**, E817–E835, https://doi.org/10.1175/BAMS-D-20-0076.1.

Donner, W. R., H. Rodriguez, and W. Diaz, 2012: Tornado warnings in three southern states: A qualitative analysis of public response patterns. *J. Homeland Secur. Emerg. Manage.*, **9**, 5, https://doi.org/10.1515/1547-7355.1955.

Downes, M., L. C. Gurrin, D. R. English, J. Pirkis, D. Currier, M. J. Spittal, and J. B. Carlin, 2018: Multilevel regression and poststratification: A modeling approach to estimating population quantities from highly selected survey samples. *Amer. J. Epidemiol.*, **187**, 1780–1790, https://doi.org/10.1093/aje/kwy070.

Durage, S. W., S. C. Wirasinghe, and J. Y. Ruwanpura, 2016: Decision analysis for tornado warning and evacuation. *Nat. Hazards Rev.*, **17**, 04015014, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000195.

Emrich, C. T., and S. L. Cutter, 2011: Social vulnerability to climate-sensitive hazards in the southern United States. *Wea. Climate Soc.*, **3**, 193–208, https://doi.org/10.1175/2011WCAS1092.1.

Fothergill, A., 1996: Gender, risk, and disaster. *Int. J. Mass Emerg. Disasters*, **14**, 33–56, https://doi.org/10.1177/028072709601400103.

Gelman, A., T. C. Little, and M. S. D. Witter, 1997: Poststratification into many categories using hierarchical logistic regression. *Surv. Methodol.*, **23**, 127–135.

Getis, A., and J. K. Ord, 2010: The analysis of spatial association by use of distance statistics. *Perspectives on Spatial Data Analysis*, Springer, 127–145.

Hanretty, C., 2020: An introduction to multilevel regression and post-stratification for estimating constituency opinion. *Polit. Stud. Rev.*, **18**, 630–645, https://doi.org/10.1177/1478929919864773.

Harrison, D. R., and C. D. Karstens, 2017: A climatology of operational storm-based warnings: A geospatial analysis. *Wea. Forecasting*, **32**, 47–60, https://doi.org/10.1175/WAF-D-15-0146.1.

Henderson, J., E. R. Nielsen, G. R. Herman, and R. S. Schumacher, 2020: A hazard multiple: Overlapping tornado and flash flood warnings in a National Weather Service forecast office in the southeastern United States. *Wea. Forecasting*, **35**, 1459–1481, https://doi.org/10.1175/WAF-D-19-0216.1.

Howard, S. P., K. E. Klockow-McClain, and K. M. Simmons, 2021: Firm behavior in the face of severe weather: Economic analysis between probabilistic and deterministic warnings. *Wea. Forecasting*, **36**, 757–767, https://doi.org/10.1175/WAF-D-20-0107.1.

Howe, P. D., M. Mildenberger, J. R. Marlon, and A. Leiserowitz, 2015: Geographic variation in opinions on climate change at state and local scales in the USA. *Nat. Climate Change*, **5**, 596–603, https://doi.org/10.1038/nclimate2583.

——, J. R. Marlon, X. Wang, and A. Leiserowitz, 2019: Public perceptions of the health risks of extreme heat across US states, counties, and neighborhoods. *Proc. Natl. Acad. Sci. USA*, **116**, 6743–6748, https://doi.org/10.1073/pnas.1813145116.

——, O. V. Wilhelmi, M. H. Hayden, and C. O'Lenick, 2023: Geographic and demographic variation in worry about extreme heat and COVID-19 risk in summer 2020. *Appl. Geogr.*, **152**, 102876, https://doi.org/10.1016/j.apgeog.2023.102876.

Jauernic, S. T., and M. S. Van Den Broeke, 2016: Perceptions of tornadoes, tornado risk, and tornado safety actions and their effects on warning response among Nebraska undergraduates. *Nat. Hazards*, **80**, 329–350, https://doi.org/10.1007/s11069-015-1970-9.

Joslyn, S. L., and J. E. LeClerc, 2012: Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *J. Exp. Psychol. Appl.*, **18**, 126–140, https://doi.org/10.1037/a0025185.

Kasperson, R. E., O. Renn, P. Slovic, H. S. Brown, J. Emel, R. Goble, J. X. Kasperson, and S. Ratick, 1988: The social amplification of risk: A conceptual framework. *Risk Anal.*, **8**, 177–187, https://doi.org/10.1111/j.1539-6924.1988.tb01168.x.

Katz, R. W., and A. H. Murphy, Eds., 1997: *Economic Value of Weather and Climate Forecasts*. Cambridge University Press, 222 pp.

Klockow, K. E., R. A. Peppler, and R. A. McPherson, 2014: Tornado folk science in Alabama and Mississippi in the 27 April 2011 tornado outbreak. *GeoJournal*, **79**, 791–804, https://doi.org/10.1007/s10708-013-9518-6.

Klockow-McClain, K. E., R. A. McPherson, and R. P. Thomas, 2020: Cartographic design for improved decision making: Trade-offs in uncertainty visualization for tornado threats. *Ann. Assoc. Amer. Geogr.*, **110**, 314–333, https://doi.org/10.1080/24694452.2019.1602467.

Lax, J. R., and J. H. Phillips, 2009: How should we estimate public opinion in the states? *Amer. J. Polit. Sci.*, **53**, 107–121, https://doi.org/10.1111/j.1540-5907.2008.00360.x.

Lindell, M. K., and R. W. Perry, 2012: The protective action decision model: Theoretical modifications and additional evidence. *Risk Anal.*, **32**, 616–632, https://doi.org/10.1111/j.1539-6924.2011.01647.x.

Masuda, J. R., and T. Garvin, 2006: Place, culture, and the social amplification of risk. *Risk Anal.*, **26**, 437–454, https://doi.org/10.1111/j.1539-6924.2006.00749.x.

Mileti, D. S., and J. H. Sorensen, 1990: Communication of emergency public warnings: A social science perspective and state-of-the-art assessment. ORNL Tech Rep. ORNL-6609, 162 pp., https://www.osti.gov/servlets/purl/6137387.

Murphy, A. H., 1976: Decision-making models in the cost-loss ratio situation and measures of the value of probability forecasts. *Mon. Wea. Rev.*, **104**, 1058–1065, https://doi.org/10.1175/1520-0493(1976)104<1058:DMMITC>2.0.CO;2.

Nagele, D. E., and J. E. Trainor, 2012: Geographic specificity, tornadoes, and protective action. *Wea. Climate Soc.*, **4**, 145–155, https://doi.org/10.1175/WCAS-D-11-00047.1.

National Weather Service, 2019: Building a Weather-Ready Nation 2019–2022 strategic plan. NOAA, 28 pp., https://www.weather.

gov/media/wrn/NWS_Weather-Ready-Nation_Strategic_Plan_2019-2022.pdf.

NCEI, 2020: U.S. tornado climatology. NOAA, accessed 10 April 2020, https://www.ncei.noaa.gov/access/monitoring/monthly-report/tornadoes/202013.

——, 2022: Storm Events Database. NOAA, accessed 20 October 2022, https://www.ncdc.noaa.gov/stormevents/.

——, 2023: U.S. billion-dollar weather and climate disasters, 1980–present. NOAA, accessed 25 April 2023, https://doi.org/10.25921/STKW-7W73.

NOAA, 2011a: The historic tornadoes of April 2011. U.S. Department of Commerce, 76 pp., https://www.weather.gov/media/publications/assessments/historic_tornadoes.pdf.

——, 2011b: Joplin, Missouri, tornado—May 22, 2011. U.S. Department of Commerce, 41 pp., https://www.weather.gov/media/publications/assessments/Joplin_tornado.pdf.

——, 2014: May 2013 Oklahoma tornadoes and flash flooding. U.S. Department of Commerce, 63 pp., https://www.weather.gov/media/publications/assessments/13oklahoma_tornadoes.pdf.

Ord, J. K., and A. Getis, 1995: Local spatial autocorrelation statistics: Distributional issues and an application. *Geogr. Anal.*, **27**, 286–306, https://doi.org/10.1111/j.1538-4632.1995.tb00912.x.

——, and ——, 2001: Testing for local spatial autocorrelation in the presence of global autocorrelation. *J. Reg. Sci.*, **41**, 411–432, https://doi.org/10.1111/0022-4146.00224.

Paul, B. K., M. Stimers, and M. Caldas, 2015: Predictors of compliance with tornado warnings issued in Joplin, Missouri, in 2011. *Disasters*, **39**, 108–124, https://doi.org/10.1111/disa.12087.

Peppler, R. A., K. E. Klockow, and R. D. Smith, 2018: Hazardscapes: Perceptions of tornado risk and the role of place attachment in central Oklahoma. *Explorations in Place Attachment*, J. S. Smith, Ed., Routledge, 33–45.

Riad, J. K., 1997: Hurricane threat and evacuation intentions: An analysis of risk perception, preparedness, social influence, and resources. Ph.D. dissertation, Georgia State University, 24 pp., https://www.proquest.com/docview/304341552?pq-origsite=gscholar&fromopenview=true.

Ripberger, J. T., M. J. Krocak, W. W. Wehde, J. N. Allan, C. Silva, and H. Jenkins-Smith, 2019: Measuring tornado warning reception, comprehension, and response in the United States. *Wea. Climate Soc.*, **11**, 863–880, https://doi.org/10.1175/WCAS-D-19-0015.1.

——, C. L. Silva, H. C. Jenkins-Smith, J. Allan, M. Crock, W. Wehde, and S. Ernst, 2020: Exploring community differences in tornado warning reception, comprehension, and response across the United States. *Bull. Amer. Meteor. Soc.*, **101**, E936–E948, https://doi.org/10.1175/BAMS-D-19-0064.1.

Sanders, S., T. Adams, and E. Joseph, 2020: Severe weather forecasts and public perceptions: An analysis of the 2011 super outbreak in Tuscaloosa, Alabama. *Wea. Climate Soc.*, **12**, 473–485, https://doi.org/10.1175/WCAS-D-18-0090.1.

Simmons, K. M., and D. Sutter, 2011: *Economic and Societal Impacts of Tornadoes*. Amer. Meteor. Soc., 282 pp.

Slovic, P., 1987: Perception of risk. *Science*, **236**, 280–285, https://doi.org/10.1126/science.3563507.

Stewart, A. E., J. K. Lazo, R. E. Morss, and J. L. Demuth, 2012: The relationship of weather salience with the perceptions and uses of weather information in a nationwide sample of the United States. *Wea. Climate Soc.*, **4**, 172–189, https://doi.org/10.1175/WCAS-D-11-00033.1.

Strader, S. M., A. M. Haberlie, and A. G. Loitz, 2021: Assessment of NWS county warning area tornado risk, exposure, and vulnerability. *Wea. Climate Soc.*, **13**, 189–209, https://doi.org/10.1175/WCAS-D-20-0107.1.

Student, 1908: The probable error of a mean. *Biometrika*, **4**, 1–25, https://doi.org/10.2307/2331554.

Thompson, J. C., 1952: On the operational deficiences in categorical weather forecasts. *Bull. Amer. Meteor. Soc.*, **33**, 223–226, https://doi.org/10.1175/1520-0477-33.6.223.

Trainor, J. E., D. Nagele, B. Philips, and B. Scott, 2015: Tornadoes, social science, and the false alarm effect. *Wea. Climate Soc.*, **7**, 333–352, https://doi.org/10.1175/WCAS-D-14-00052.1.

Trujillo-Falcón, J. E., O. Bermúdez, K. Negrón-Hernández, J. Lipski, E. Leitman, and K. Berry, 2021: Hazardous weather communication en Español: Challenges, current resources, and future practices. *Bull. Amer. Meteor. Soc.*, **102**, E765–E773, https://doi.org/10.1175/BAMS-D-20-0249.1.

Uccellini, L. W., and J. E. Ten Hoeve, 2019: Evolving the National Weather Service to build a weather-ready nation: Connecting observations, forecasts, and warnings to decision-makers through impact-based decision support services. *Bull. Amer. Meteor. Soc.*, **100**, 1923–1942, https://doi.org/10.1175/BAMS-D-18-0159.1.

West, D. M., and M. Orr, 2007: Race, gender, and communications in natural disasters. *Policy Stud. J.*, **35**, 569–586, https://doi.org/10.1111/j.1541-0072.2007.00237.x.

Xu, H., 2014: Comparing spatial and multilevel regression models for binary outcomes in neighborhood studies. *Sociol. Methodol.*, **44**, 229–272, https://doi.org/10.1177/0081175013490188.

Zhang, B., S. van der Linden, M. Mildenberger, J. R. Marlon, P. D. Howe, and A. Leiserowitz, 2018: Experimental effects of climate messages vary geographically. *Nat. Climate Change*, **8**, 370–374, https://doi.org/10.1038/s41558-018-0122-0.

Zhang, X., J. B. Holt, S. Yun, H. Lu, K. J. Greenlund, and J. B. Croft, 2015: Validation of multilevel regression and poststratification methodology for small area estimation of health indicators from the behavioral risk factor surveillance system. *Amer. J. Epidemiol.*, **182**, 127–137, https://doi.org/10.1093/aje/kwv002.