Household Evacuation Decision Making and the Benefits of Improved Hurricane Forecasting: Developing a Framework for Assessment

JEFFREY K. LAZO

Societal Impacts Program, National Center for Atmospheric Research, Boulder, Colorado

DONALD M. WALDMAN

Department of Economics, University of Colorado, Boulder, Colorado

BETTY HEARN MORROW

SocResearch, Miami, Florida

JENNIFER A. THACHER

Department of Economics, University of New Mexico, Albuquerque, New Mexico

(Manuscript received 4 June 2009, in final form 9 September 2009)

ABSTRACT

Hurricane warnings are the primary sources of information that enable the public to assess the risk and develop responses to threats from hurricanes. These warnings have significantly reduced the number of hurricane-related fatalities in the last several decades. Further investment in the science and implementation of the warning system is a primary mission of the National Weather Service and its partners. It is important that the weather community understand the public's preferences and values for such investments; yet, there is little empirical information on the use of forecasts in evacuation decision making, the economic value of current forecasts, or the potential use or value for improvements in hurricane forecasts. Such information is needed to evaluate whether improved forecast provision and dissemination offer more benefit to society than alternative public investments.

Fundamental aspects of households' perceptions of hurricane forecasts and warnings and their potential uses of and values for improved hurricane forecast information are examined. The study was designed in part to examine the viability of survey research methods for exploring evacuation decision making and for eliciting values for improved hurricane forecasts and warnings. First, aspects that affect households' stated likelihood of evacuation are explored, because informing such decisions is one of the primary purposes of hurricane forecasts and warnings. Then, stated-choice valuation methods are used to analyze choices between potential forecast-improvement programs and the accuracy of existing forecasts. From this, the willingness to pay (WTP) for improved forecasts is derived from survey respondents.

1. Introduction

Hurricane warnings and forecasts are issued by the National Weather Service (NWS)—mainly through the National Hurricane Center (NHC)—and disseminated through the media and by emergency managers. These warnings and forecasts are the primary sources of information that allow the public to assess the risk of and decide how to respond to threats from landfalling hurricanes (Gladwin et al. 2007). The hurricane forecast and warning system is credited with significantly reducing the number of hurricane-related fatalities in the twentieth century (Willoughby 2002), and investment in the science and implementation for this system is a primary mission of the NWS and its partners. Significant additional investments are planned in response to the recent impacts of landfalling hurricanes, even though there is little empirical information on the economic value of these investments or the potential value of improvements in hurricane forecasts.

Corresponding author address: Jeffrey K. Lazo, Societal Impacts Program, National Center for Atmospheric Research, Box 3000-RAL, Boulder, CO 80307.
E-mail: lazo@ucar.edu

DOI: 10.1175/2009WAF2222310.1

For instance, in response to the devastation of Hurricanes Katrina, Rita, and Wilma, the National Oceanic and Atmospheric Administration (NOAA) established the Hurricane Forecast Improvement Project (HFIP)—a 10-yr plan to improve 1–5-day tropical cyclone forecasts (NOAA 2008). Spending for HFIP in fiscal year 2009 alone was proposed at $19 million [approximately $6 million of this is for research with the balance slated for development, transition, and operations (F. Toepfer, NOAA, 2009, personal communication)]. Because households are the ultimate end users of hurricane forecasts, it is important to understand the public's preferences and values for current forecasts and for improvements in forecasts. Estimates of the value of forecasts can be used to evaluate whether improved forecast accuracy and dissemination offer more benefit to society than alternative public investments such as infrastructure or improved forecasts of other hazards (Letson et al. 2007).

To date there has been little quantitative assessment of the general public's preferences and values for hurricane forecasts and warnings. In a narrow sense, we use the term ''hurricane forecasts and warnings'' to indicate the official warning products produced and disseminated by the NHC. Although a key component of hurricane information, these official products are only part of the broader process by which public officials, the media, and others combine and transform this information into the forecasts, warnings, evacuation orders, or other hurricane-related information ultimately used by other stakeholders, especially the public, in decision making.

Considerable research interest has focused on hurricane evacuation, but not forecast and warning use. Much of the evacuation work is descriptive and, at times, contradictory (Baker 1993; Gladwin and Peacock 1997; Dow and Cutter 1998; Yin and Newman 1999; Dash and Gladwin 2007). Early research focused on the effects of warnings, but other factors such as household composition, living in an evacuation zone, and length of residence have consistently been associated with evacuation decisions (Whitehead et al. 2000; Whitehead 2005). Smaller households, those with children present, those living in evacuation zones, and newer residents who have never been through a hurricane are more likely to evacuate (Dash and Gladwin 2007). Most directly related to forecast and warning use is work examining the relationship between risk perception and evacuation behavior (Dow and Cutter 1998; Dash and Gladwin 2007). Risk is a social construction in which most people use their knowledge and experience to interpret their level of safety (Morrow 2009). Although some coastal residents rely primarily on an evacuation order for decision making (Gladwin et al. 2001; Zhang et al. 2007), many follow the storm via various information sources, consider the safety of their home and family, and then weigh their evacuation options. There is some indication that people pay close attention to information about the approaching hurricane, but they may not always understand the nuances of the warning system, such as how to interpret watches and warnings (Morrow and Gladwin 2005).

To build on this research and begin filling these knowledge gaps, this research examines fundamental aspects of nonmeteorologists' perceptions of hurricane forecasts and warnings and their values for improved hurricane forecast information. This work is part of longer-term efforts to explore and understand these perceptions and values across all hurricane-vulnerable populations. We focus here on work to date that includes data from a survey implemented in Miami, Florida, in September 2008. The current results show the viability of these methods for exploring evacuation decision making and values for improved hurricane forecasts and warnings, but the empirical results cannot be generalized to the entire hurricane-vulnerable population. Future work, based on results from these efforts, will involve larger, more geographically diverse samples developed using sampling approaches that should yield results that can be aggregated to the relevant hurricane-vulnerable populations.

We present results specifically about households' evacuation decision making and the economic value of improving hurricane forecasts. We explore aspects affecting households' stated likelihood of evacuation, because informing such decisions is a primary purpose of hurricane forecasts and warnings. Using responses from several other questions in the survey, we model the stated likelihood of evacuation as a function of forecasted hurricane intensity, perceived risk from wind and storm surge, experience with hurricanes, perceived accuracy of forecasts, barriers to evacuation, and sociodemographic characteristics.

We then use stated-choice valuation methods to analyze choices between potential forecast-improvement programs and the accuracy of existing forecasts and to derive the willingness to pay (WTP) for improved forecasts from survey respondents. Stated-choice methods are used in economic benefit estimation when markets do not exist for the goods or services being valued. In this case, little or no market information on households' value for hurricane forecast and warning information exists. These methods use a hypothetical context in a survey format, with questions designed as choices between alternatives that include differences in goods and services as well as in costs. The alternatives that a subject prefers reveal information about his or her underlying values for the goods and services in those alternatives.

An important aspect of this approach is evaluating differences in preferences and values across different segments of society. We use our valuation approach to look at values for improved forecasts for different subgroups. Such differences need further exploration in light of the critical issue of vulnerable populations and the differential impacts of hurricanes on these populations, as evidenced by Hurricane Katrina's impacts on New Orleans, Louisiana (Phillips and Morrow 2007). Although potentially important, all of the findings reported here suggest the need for further research across broader geographic populations with larger samples to better understand how and why these factors affect people's preparedness.

2. Survey development, sampling, and implementation

We collected data from 80 households using an in-person survey in Miami. To ensure that the survey will eventually be applicable across all hurricane-vulnerable geographic areas, the draft survey instrument was developed through an extensive series of focus groups in Miami, in New Orleans, and in Charleston, South Carolina; we also conducted one-on-one cognitive interviews (Lazo 2002, 2004). We analyzed two versions of the draft final survey for reading level using the Flesch Reading Ease test, and changed wording as necessary for clarity. Verbal protocol analysis (Ericsson and Simon 1993) was used in six one-on-one pretests of the instrument in August 2008. Each session was taped as the respondents read the questions aloud and "thought aloud" as they answered each question. After each pretest, a facilitator discussed the survey with each respondent to elicit additional suggestions, thus providing a retrospective report. Based on analysis of the results, suggestions for instrument revision were undertaken and improvements made throughout the pretesting process.

The survey elicits information on respondents' experiences with hurricanes, perceived and actual vulnerability to hurricane risks and impacts, preparation for hurricanes, factors affecting evacuation decision making, perceptions of hurricane forecasts and warnings, preferences and values for current and improved forecasts and warnings, and general sociodemographic information. Not all aspects or results are reported here.

The survey was implemented at National Opinion Research Services (NORS) in Miami on 4 September 2008. The only restrictions on recruitment for the sample was that respondents be at least 18 years of age and live within 30 mi of the coast. Because a predetermined sample size of 80 was the target for survey implementation, and as we are not attempting to generalize to the general population, we do not calculate a response rate per se. NORS performed the manual data entry and quality control assessment.

The sample was composed of 16 Caucasians, 16 African Americans, 47 Spanish/Hispanic/Latino, and one "other." Half of the respondents were male and half were female. Half of the sample had a college degree; this is a higher educational-attainment level than the general population of Miami–Dade County, where 22% reported at least a bachelor's degree in 2000 (EIDS 2008).[1] The average age was 43 and ranged from 18 to 70. The average income was approximately $61,000, and the median was between $40,000 and $50,000. This compared well with the general Dade County population, which had a median household income of $46,931 in 2000 (EIDS 2008). Because housing characteristics are an important factor in hurricane vulnerability, we also collected data on respondents' housing. Only five respondents live in mobile homes and none live in boats or recreation vehicles—this was comparable to the incidence of mobile home occupancy in the general Miami–Fort Lauderdale metropolitan area population.[2] Those who rented and those who owned their residence were split 39% and 61%, respectively, reasonably close to the split reported in the 2002 American Housing Survey (AHS) of roughly 35% and 65% (U.S. Department of Housing and Urban Development–U.S. Department of Commerce 2003). For those occupying houses, 67% were owners, and for those occupying apartments, 29% were owners. Although we used a small, nonrandom sample, most of the sociodemographic characteristics are reasonably similar to the Miami–Fort Lauderdale metropolitan area population, and we would not expect our results to be substantively different from future, larger random samples.

3. General results

Prior to discussing results specific to evacuation decision making and valuation, several general results give additional context to the responses. Eighty-nine percent of respondents indicated that they had been personally

[1] In general, we compare our sample to the general population of Miami–Dade County because this was the source of our sample and also the level of data available from U.S. Census Bureau statistics for comparison. Some sociodemographic information is available for the Miami–Fort Lauderdale metropolitan area and thus, when available, we compare our data to this information.

[2] Table 1-1 in the AHS (available online at http://www.census.gov/prod/2003pubs/h170-02-28.pdf) indicates 53 000 "manufactured/mobile home tiedowns" out of 1 638 700 total housing units. Manufactured/mobile homes are thus about 3.2% of the total housing units in the metropolitan statistical area, so our sample may not be too far off at 6.25%.

affected by a hurricane in the past. Most (~60%) had experience with Hurricane Andrew, and of those, most accurately recalled the year of Andrew (1992). Given that Andrew was 16 years prior to the date of the survey, this is a high level of accurate recall, indicating the impact of that hurricane on respondents. More than 50% of respondents indicated that they had not received ''information from any public officials regarding what you should do or where you should go in the event of a possible future hurricane.'' For those who did indicate receiving information, the mayor and other (nonspecific) government entities were mentioned several times as sources of this information.

Seven of 10 respondents have taken some action to prepare for a hurricane. A simple statistical analysis of the prepare/do not prepare response revealed that the only significant predictor of taking action in preparation for a hurricane was previous experience with a hurricane: those with personal experience with a prior hurricane were less likely to have taken action to prepare for a hurricane. Forty-five of the 80 respondents (56%) indicated that they had adequate hurricane shutters installed at their residence; those who owned their residence were more likely to have shutters. In addition, those who indicated that they perceived a higher likelihood of tornadoes associated with a hurricane were more likely to have shutters. With these two factors accounted for, the likelihood of having shutters was independent of age, income, and length of time in residence.

4. Factors affecting evacuation decision making

For each of the five hurricane categories used in the Saffir–Simpson hurricane scale (SSHS), respondents were asked, ''How likely is it that you would evacuate if you were to receive a hurricane warning for your area?'' Individuals indicated their likelihood of evacuating on a five- point Likert scale ranging from ''not at all likely'' (1) to ''extremely likely'' (5). Figure 1 shows the mean response for each of the five levels on the SSHS.

The intensity of the potential hurricane obviously plays a role in individuals' decision making with respect to evacuation, but a broad range of other factors constrain or encourage an individual's or household's evacuation decision (Whitehead et al. 2000; Dash and Gladwin 2007). We explored these issues by analyzing responses to the evacuation-decision question based on perceived risk from wind and storm surge, experience with hurricanes and perceived accuracy of forecasts, stated potential barriers to evacuation, and sociodemographic characteristics. Because the response options of the dependent variable were an ordered, five-point Likert scale, we analyze responses using ordinal logistic regression in SAS using Proc Genmod (O'Connell 2005). Unlike ordinary least squares regression, logistic regression does not assume a linear relationship between independent and dependent variables (it is a form of nonlinear regression analysis), does not require normally distributed variables, and does not assume constant error variance. Ordinal logistic regression allows for the fact that there are multiple classes of the dependent variable that can be ranked (i.e., from not at all likely to extremely likely). Further, because each respondent provided up to five answers to this question (one for each hurricane intensity level), to account for potential intrasubject correlation, we used the method of generalized estimating equations (GEE; Allison 1999; Ballinger 2004).

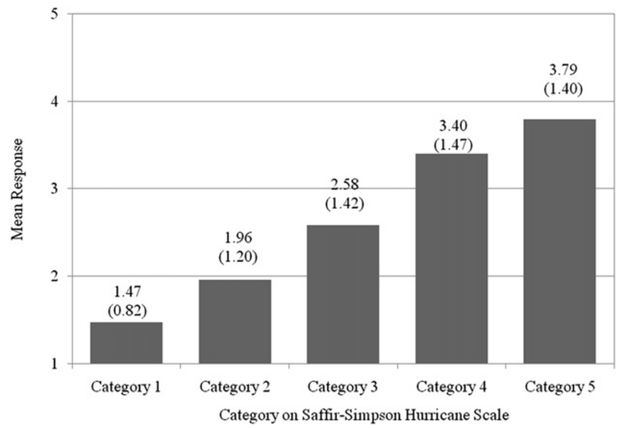


FIG. 1. Mean likelihood (std dev) of evacuation by SSHS category, from 1, ''not at all likely,'' to 5, ''extremely likely''; $n = 80$.

Table 1 provides the regression results, descriptions of the independent variables and the response options, and interpretation of the estimation results. Not all respondents provided answers for all five levels of hurricane intensity, and thus we have 383 observations across the 80 respondents.

Explanatory variables are discussed in groups of 1) hurricane intensity, 2) risk perceptions, 3) experience with hurricanes and perceived accuracy of hurricane warnings, 4) barriers to evacuation, and 5) sociodemographic characteristics. Estimates of the intercepts are also given, but because they do not provide any substantive interpretation, they are not discussed further. We treated independent variables elicited on Likert scales (e.g., 1–5) as ordinal categorical variables for the purposes of estimation (with one level omitted to allow for estimation). Dummy variables, such as those for potential barriers to evacuation, are coded as yes = 1 and no = 0. The only continuous independent variables are length in residence, age, household size (treated as continuous although technically only integers), and income. For a small number of missing responses (0.5%)

on independent variables, the mean of the responses from other respondents was substituted. Table 1 also shows parameter estimates and standard errors. Parameter estimates significant at the 10% level or better are marked with asterisks.

Parameter estimates show the influence of hurricane strength (as indicated by the SSHS) on respondents' stated likelihood of evacuation. These were estimated for each level of the SSHS, with category 5 as the omitted variable. All parameter estimates on SSHS levels are highly significant and negative, indicating that compared to a category 5 hurricane, lower levels of hurricane intensity are less likely to induce evacuation. This result is consistent with the previous literature (Lindell and Prater 2007).

The next set of explanatory variables concerned the perceived risk of potential damage to their residence. Potential for wind damage did not influence evacuation decision making, whereas potential flood damage marginally did so (significance level < 13%). This could be interpreted as an encouraging outcome, as emergency managers generally suggest that people remain locally in a strong structure, either their own home or an appropriate building, if they are in no danger from flooding, but recommend evacuation when faced with the threat of surge or inland flooding (e.g., see the information online at http://www.nhc.noaa.gov/HAW2/english/disaster_prevention.shtml).

Prior experience with hurricanes did not affect intentions to evacuate. Individuals who perceived hurricane forecasts to be less accurate, however (i.e., responded with a 2 on the 1–5 scale; no one responded with a 1), indicated a lower likelihood of evacuation (significant at the 11% level). Combined with the finding that an attitude of "I do not trust the accuracy of hurricane forecasts enough to be willing to evacuate" was not a barrier to evacuation decision making indicates the positive role of hurricane forecasts and warnings in evacuation decision making. Another interpretation of this response is that people who do not trust the accuracy of forecasts are less likely to evacuate given a hurricane warning.

When respondents were asked about factors that may represent barriers to their evacuation decisions, the only barrier that was nearly significant (at the 12.9% level) in explaining their stated intent to evacuate was that they did not want to leave their property unprotected. Factors such as too much traffic, having pets, and being in poor health did not influence their stated intentions. This requires further research and comparison with findings from other studies (e.g., Whitehead et al. 2000), including studies with larger sample sizes and studies that compare stated ex ante intentions and actual behavior ex post.

Sociodemographics played a significant role on stated evacuation intentions. Older respondents, those with higher education levels, and those employed full time (significant at the 11% level) were more likely to evacuate than their counterparts, whereas those longer in their residence and of higher income indicated a lower likelihood of evacuation. Household size, home ownership, race, primary language, and living in mobile homes were not significant in explaining stated evacuation likelihoods. Because only 5 of the 80 respondents indicated living in mobile homes, this portion of the sample may have been too small to affect the results.

5. Valuation

a. Stated preference approach

To explore values for potential improvements in hurricane forecasts, we used a set of stated-preference choice questions (also called stated choice) in the survey. In this approach, individuals are presented with a set of alternatives and asked to choose their preferred alternative. The preferred alternatives reveal information about the underlying values for the goods and services in those alternatives. Choice questions evolved from conjoint analysis, a method used extensively in marketing and transportation research and now commonly used in nonmarket valuation studies in environmental and health economics (Adamowicz et al. 1998; Ben-Akiva and Lerman 1985; Louviere et al. 2001).

The basic idea of stated choice is that a commodity (i.e., a good or service) is composed of a set of attributes. Commodities vary in the level of each attribute. Choosing between two commodities that vary in the attribute levels thus reveals information about respondents' preferences for those attributes. Including cost as one of the attributes reveals information about the marginal value of money and allows us to convert the preferences for the attributes into marginal values for those attributes.

In this study, the commodities are hurricane forecasting programs. Respondents were asked to choose between two hurricane forecast improvement programs in each of eight scenarios. Each program was described by specific levels of four hurricane forecast attributes—time of expected landfall, maximum wind speed, projected location of landfall, and expected storm surge—as well as increased annual cost to the household. Table 2 shows these attributes, their baseline levels, and their possible levels under the improvement programs. Note that the "baseline" or "current" cost to households is the current level of taxes; that is, the implicit "increase in annual cost to your household" at baseline is zero dollars. The dollar values indicated in Table 2 are only for improvements above baseline and represent the three

TABLE 1. Ordinal logistic regression of likelihood of evacuation.[a]

Variable (measurement units)	Estimated (std error) significance	Interpretation of results
Hurricane intensity based on SSHS (1–5): information provided on associated wind speed and storm surge		
Category 1	−4.46 (0.45) ***	Increased intensity increases likelihood of evacuation; parameters are interpreted with respect to a category 5 hurricane
Category 2	−3.15 (0.35) ***	
Category 3	−2.06 (0.26) ***	
Category 4	−0.70 (0.13) ***	
Risk perception scales: indication of expected damage on respondent residence from a major hurricane		
Risk of wind damage to current residence from a major hurricane (from 1, no wind damage, to 5, extreme wind damage)		
Wind risk 1	0.80 (1.41)	No impact of perceived risk of damage from wind caused by hurricane; parameters are interpreted with respect to a wind risk of ''extreme wind damage''
Wind risk 2	−0.76 (1.15)	
Wind risk 3	−0.74 (0.75)	
Wind risk 4	−0.21 (0.80)	
Risk of flooding or storm surge damage to current residence from a major hurricane (from 1, no flooding or storm surge damage, to 5, extreme flooding or storm surge damage)		As a group, flood risk has a 13% level of significance (type 3 GEE analysis score statistics χ^2 = 7.19; DF, 4), but individually it is not significant; this implies that the higher a perceived risk of damage from flooding caused by hurricane, the more likely the inclination to evacuate; parameters are interpreted with respect to a flood risk of ''extreme flooding or storm surge damage''
Flood risk 1	−1.69 (1.11)	
Flood risk 2	−1.33 (1.07)	
Flood risk 3	−0.80 (0.90)	
Flood risk 4	0.62 (0.93)	
Experience with hurricanes and perceptions of hurricane forecasts		
Personally affected by a past hurricane (1, yes; 0, no)	−0.37 (0.70)	Having been previously affected by a hurricane did not significantly affect likelihood of evacuation
Perceived accuracy of hurricane forecasts (from 1, not at all accurate, to 5, extremely accurate; no respondents indicated a 1)		At the 11% level of significance, compared to a low level of perceived accuracy (choice 2) of hurricane forecasts, respondents perceiving high accuracy are significantly more likely to evacuate
Accurate 2	−3.66 (2.32)	
Accurate 3	−1.65 (1.65)	
Accurate 4	−0.73 (1.60)	
Barriers to evacuation: reasons a respondent may not evacuate (1, yes; 0, no)		
I believe my home is safe from wind or flooding	−0.21 (0.60)	Perception that house is safe from wind or flooding did not significantly affect likelihood of evacuation
I do not know where I am supposed to go	0.32 (0.52)	Not knowing where to evacuate to did not significantly affect likelihood of evacuation
I have a pet or pets and that makes it difficult to find a place to go	−0.23 (0.40)	Having pets did not significantly affect likelihood of evacuation
I do not trust the accuracy of hurricane forecasts enough to be willing to evacuate	0.08 (0.57)	Not trusting accuracy of hurricane forecasts did not significantly affect likelihood of evacuation
There would be so much traffic that I would not be able to get somewhere safe in time	0.40 (0.43)	Concern of too much traffic on evacuation routes did not significantly affect likelihood of evacuation
I do not have transportation in order to evacuate	−0.52 (0.98)	Not having transportation available did not significantly affect likelihood of evacuation
I do not want to leave my home or business unprotected	−0.65 (0.48)	Concern about protecting property did not significantly affect likelihood of evacuation; this has a 12.9% level of significance with a type 3 GEE analysis score statistics (χ^2 = 2.31; DF, 1)
I am not in good enough health to leave my house and go somewhere else	0.47 (0.91)	Being in poor health did not significantly affect likelihood of evacuation

TABLE 1. (*Continued*)

Variable (measurement units)	Estimated (std error) significance	Interpretation of results
	Sociodemographic characteristics	
Live in mobile home (1, yes; 0, no)	0.38 (0.85)	Living in a mobile home did not significantly affect likelihood of evacuation
Own their residence (1, yes; 0, no)	0.34 (0.64)	Homeownership did not significantly affect likelihood of evacuation
Length in current residence (yr)	−0.06 (0.03) *	Longer in residence less likely to evacuate
Age (yr)	0.03 (0.02) *	Older people more likely to evacuate
Household size (No. of people living in household)	−0.06 (0.14)	Household size did not significantly affect likelihood of evacuation
White (1, Caucasian; 0, other)	0.31 (0.57)	Race did not significantly affect likelihood of evacuation
English speaking as primary language (1, English; 0, other)	0.68 (0.51)	Primary language did not significantly affect likelihood of evacuation
Education (yr)	0.17 (0.07) **	People with more education are more likely to evacuate
Gender (male, 1; female, 2)	−0.07 (0.44)	Gender did not significantly affect likelihood of evacuation
Employed (full time or part time) (1, yes; 0, no)	0.87 (0.54)	At the 11% level of significance, those with full- or part-time employment are more likely to evacuate
Income (thousands of 2008$)	−0.02 (0.01) **	Higher income less likely to evacuate
	Intercept terms	
Intercept 1	−0.70 (2.86)	Estimates of the ''intercepts'' between the five levels of likelihood
Intercept 2	0.30 (2.89)	
Intercept 3	1.30 (2.88)	
Intercept 4	2.70 (2.88)	

[a] Here, $n = 400$: 80 subjects × five responses each.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

levels of cost indicated for scenarios involving improvements. These dollar values were not based on any specific real-world expectation of costs but rather were set at levels expected to provide information on respondents' reactions to believable cost levels.

Any specific program was a combination of different levels (baseline, intermediate improvement, and maximum improvement) of the five attributes (including cost). Each program thus represented an improvement in accuracy over current (baseline) forecasts at some cost to the household. For each choice scenario, respondents indicated their preference between two potential improvement programs, for instance, program A and program B, and then whether they would prefer the status quo (i.e., keeping all levels at their baseline level) over their A–B choice. Programs were labeled A–B in the first scenario, C–D in the second, and so forth. A typical choice appeared as shown in Fig. 2, where the programs were labeled C and D. For example, in this question, an individual could choose program C with improved accuracy at predicting the time and location of expected landfall and the expected storm surge for an additional $12 per year or program D with improved ability to forecast the maximum wind speed and the projected landfall for $24 per year. Suppose the individual says they prefer program C. In the follow-up question, the individual is asked whether they value the improved accuracy of predicting the time and location of expected landfall and the expected storm surge enough to pay an additional $12 per year or whether they are content with current forecasting abilities. The program labels had no normative or ordinal content. For simplicity we henceforth refer to all of the choices between two potential improvement programs as A–B choices.

b. *Choice set design*

By having many different scenarios with different programs offered, we can use statistical methods based on a theoretical model of decision making (i.e., random utility behavioral model, discussed below) to estimate respondent preferences and values for the different attributes. A critical component of this approach is designing the set of scenarios and programs to permit and optimize statistical analysis. The process of setting up these scenarios and choosing the levels of the attributes in each scenario is called choice set design.

TABLE 2. Attributes of hurricane forecasts and warnings and the levels of these attributes used to create choice set alternatives.*

Attributes	Levels		
	Baseline	Intermediate improvement	Max improvement
Time of expected landfall	8 h 48 h in advance	6 h	4 h
Max wind speed	20 mi h^{-1} 48 h in advance	15 mi h^{-1}	10 mi h^{-1}
Projected location of landfall	100 mi 48 h in advance	80 mi 48 h in advance	65 mi 48 h in advance
Expected storm surge	±8 ft MSL 48 h in advance	6 ft MSL	4 ft MSL
Increase in annual cost to your household	No additional cost if all attributes at baseline	\$12, \$24, \$48: if any attribute is improved, then an increase in cost is included in the scenario	

* One level for each attribute and a cost to household composed a single alternative that was then compared to a different alternative.

A number of constraints affected the final form of this design: the sample of approximately 80 respondents, six choice questions each, and two generic alternatives for each question. Given these constraints, we created a nonlinear main-effects design; this does not allow for the estimation of any interaction effects among attributes.[3] The design also allows for the estimation of quadratic effects. The design was created from a fractional-factorial orthogonal array of 36 runs with three blocks, using the %choiceff macro in SAS (Huber and Zwerina 1996; Kuhfeld 2005).[4] Thus, the initial choice set design consisted of 18 questions of two alternatives each, which were then randomly divided into three survey versions.

After the initial choice set design was created, we made two modifications. First, we added two additional simple choice questions as the first two choice questions (for a total of eight choice questions). These two initial choice questions offered to respondents were designed by hand and were intended to ''teach'' respondents how the choice questions worked. Thus, these first two choice questions were purposefully simple and restricted the number of attributes that varied. In addition, the A–B choice (but not the follow-up question) was asked in our survey for the first choice question. These simple choice questions also included a price level of zero.

Second, we changed five of the prices in the choice sets in the original design to make the scenarios more believable (e.g., so an alternative with all attribute levels better than the attribute levels in the alternative did not cost less).

[3] Future work with larger samples will allow us to use more complex choice set designs and to estimate interactions between attributes.

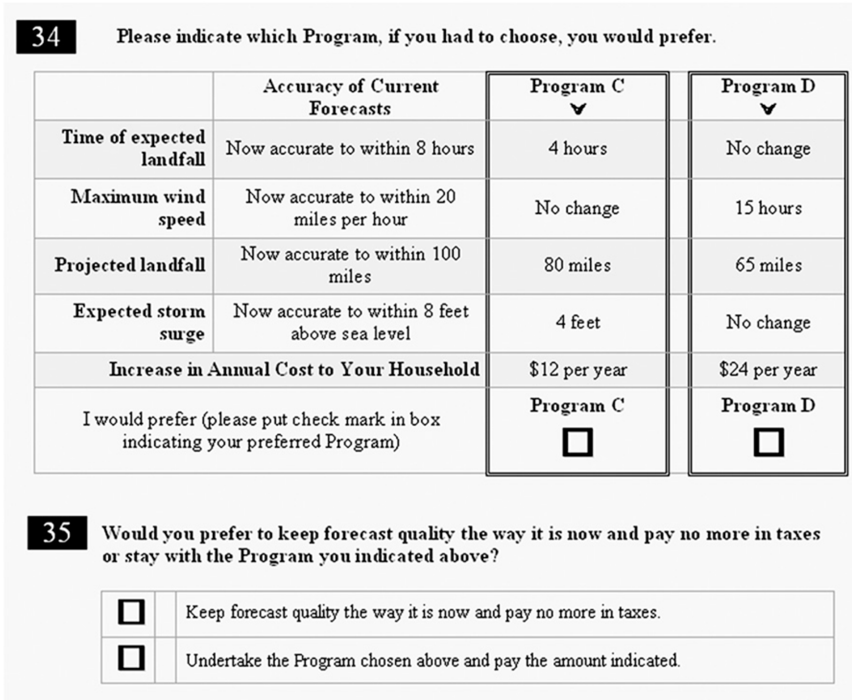
[4] For purposes of the %choiceff macro, we assumed all beta parameters (i.e., marginal valuations) to be zero, set a convergence criterion of 0.00005, set 200 random starts, and allowed for up to 20 internal iterations. The best design was selected on the basis of the D-efficiency criterion; the D-efficiency score was 67.55.

Because the sample was split roughly into three subgroups, each of which completed one of the three survey versions, we assessed if there were significant differences in the responses or attitudes of these subgroups that may have affected data quality. After the series of choice questions, we asked respondents to state their confidence in their response to these questions; there was no significant difference in the level of confidence between the three versions (Kruskal–Wallis test, $\chi^2 = 0.34$; degrees of freedom (DF) 2; probability (Pr) $> \chi^2 = 0.84$). We also asked respondents to indicate how important hurricane forecasts are to them at this time. Again, there was no significant difference between subgroups (Kruskal–Wallis test, $\chi^2 = 1.38$; DF 2; Pr $> \chi^2 = 0.50$). We thus believe that there were no significant effects of the different versions on respondents other than the content of the choice questions.

c. *The valuation model and econometric methodology*

The random utility behavioral model was assumed for econometric modeling of the choice question responses (see McFadden 1976; Manski 1977). In this approach, total utility is assumed to be the sum of the marginal utility derived from the characteristics or attributes that make up a good, in this case a hurricane forecast. When asked to choose between two alternatives differing only in the levels of the attributes and potential costs, individuals are assumed to choose the alternative providing the greatest utility, including the disutility of the costs. By asking many individuals to make several choices over many different combinations of alternatives, the marginal utility of the different attributes is implicitly revealed in these choices. We can use statistical analysis to ''back out'' the contribution of each attribute to the total utility (Ben-Akiva and Lerman 1985; Louviere et al. 2001). The utility of a choice is modeled as a linear combination of the choice attributes and a random error:

$$U_{ij} = \boldsymbol{\beta}'\mathbf{x}_{ij} + \varepsilon_{ij}, \quad i = \mathrm{A, B}; \quad j = 1, \ldots, 8, \qquad (1)$$



34 **Please indicate which Program, if you had to choose, you would prefer.**

	Accuracy of Current Forecasts	Program C	Program D
Time of expected landfall	Now accurate to within 8 hours	4 hours	No change
Maximum wind speed	Now accurate to within 20 miles per hour	No change	15 hours
Projected landfall	Now accurate to within 100 miles	80 miles	65 miles
Expected storm surge	Now accurate to within 8 feet above sea level	4 feet	No change
Increase in Annual Cost to Your Household		$12 per year	$24 per year
I would prefer (please put check mark in box indicating your preferred Program)		**Program C** ☐	**Program D** ☐

35 **Would you prefer to keep forecast quality the way it is now and pay no more in taxes or stay with the Program you indicated above?**

☐	Keep forecast quality the way it is now and pay no more in taxes.
☐	Undertake the Program chosen above and pay the amount indicated.

FIG. 2. Choice questions: 34 and 35, version 1.

where the elements of the vector $\boldsymbol{\beta}$ are the marginal utilities[5] of the attributes in the vector $\mathbf{x}$,

$$\mathbf{x} = \begin{pmatrix} \text{Time of landfall accuracy} \\ \text{Maximum wind speed accuracy} \\ \text{Project location of landfall accuracy} \\ \text{Expected storm surge accuracy} \\ \text{Annual household cost} \end{pmatrix},$$

and ε is a random disturbance.

Estimation is by bivariate probit (Greene 2003, 710–713) combining responses to the A–B choice and the follow-up question. In addition, the individual, rather than the choice occasion, is considered the observation. Under this assumption, the individual evaluates his or her utility of the status quo only once, as compared to evaluating his or her utility of the status quo at each choice occasion. To implement this approach, estimation is by maximum likelihood employing quadrature methods (Savage and Waldman 2008).

Although the usual practice in analyzing discrete choice experiments is to include all choice occasions in the estimation, the results in Table 3 include only choice occasions 2–8 because the A–B choice but not the follow-up question was asked in the first choice occasion, and the first choice occasion is sometimes omitted to allow the respondent to become familiar with the exercise. As evidence of this, other research has found it takes respondents considerably longer to choose in the first choice occasion than in subsequent choice occasions (D. M. Waldman 2008, unpublished manuscript).

As shown in Table 3, all slope coefficients (marginal utilities) are negative and highly statistically significant, as noted by the t ratios, except for the coefficient on the wind speed, which is negative and marginally statistically significant (t ratio = 1.08). The negative sign is expected because an increase in the value of any attribute represents a decrease in the accuracy of that forecast component. For example, a projected landfall location accurate to within 100 mi is inferior to one accurate to within 80 mi.

To clarify the meaning of the estimated coefficients, the third column in Table 3 reports the WTP for a one-unit improvement in the attribute. The marginal WTP for an improvement in the attribute is the ratio of each

[5] Because U is the total utility defined as the sum of the utility from the different attributes (x values), the $\boldsymbol{\beta}$ values measure the change in total utility caused by a one unit change in any given x. The $\boldsymbol{\beta}$s can thus be interpreted as the marginal utility of the attributes. For the cost attribute, the associated $\boldsymbol{\beta}$ measures the marginal utility of money and is expected to be negative because increased cost implies decreased utility (or disutility).

TABLE 3. Modeling results from a bivariate probit model with quadrature.*

Forecast attribute	Coefficient	t ratio	Marginal WTP (std error of marginal WTP estimate)
Time of expected landfall	−0.067	−3.57	\$2.18 (1.35)
Max wind speed	−0.008	−1.08	\$0.26 (0.16)
Projected location of landfall	−0.007	−3.22	\$0.23 (0.61)
Expected storm surge	−0.062	−3.50	\$2.04 (1.21)
Increase in annual cost to household	−0.030	−11.27	
Mean log likelihood −1.1971			

* Coefficient estimates represent the marginal utility of the attribute. Marginal WTP is the ratio of each marginal utility coefficient to the marginal disutility of cost. Choice sets 2–8 only ($n = 560$).

marginal utility coefficient to the marginal disutility of cost:

$$\mathrm{WTP}_k = \frac{\beta_k}{\beta_{\mathrm{cost}}},$$
$$k = \text{landfall time accuracy}, \ldots, \text{surge accuracy}.$$

To derive a benefit estimate for a specific potential program, we can use the model results and the proposed levels of improvement for the potential program and aggregate values across the four attributes. For example, we can derive a total household benefit estimate for a program that would improve all levels to the intermediate level offered in the survey. Starting with landfall time, forecasts are considered accurate to within 8 h at the time of this study. The improvement from the baseline to the intermediate level in the survey is a forecast improvement from 8 to 6 h. The estimate of \$2.18 for the WTP per hour for landfall time means that an individual would be willing to pay approximately 2 × \$2.18, or \$4.36, for this improvement. Similarly, for an improvement from the current forecast accuracy to the intermediate level for wind speed (a 5 mi h^{-1} difference), an individual would be willing to pay about 5 × \$0.26 = \$1.30, for the 20-mi improvement in estimated landfall site about 20 × \$0.23 = \$4.60, and for the 2-ft difference in storm surge about 2 × \$2.04 = \$4.08. As shown in Table 4, the total WTP for this average overall superior forecast (from baseline to intermediate levels on all attributes) is the sum: \$14.34 per household per year.

To explore differences in WTP for different populations, we partitioned the dataset based on income, gender, and education and reestimated the bivariate probit with quadrature estimation models for each of these selected partitions. Sixty percent of respondents reported income less than \$60,000, exactly 50% are male and 50% are female, and exactly half were college graduates and half were not. For each partition, we estimated the model and calculated a marginal WTP for the attributes as described above for the full model. Table 5 presents the regression results for these subsamples. As in Table 4, we derived WTP for improvement from the baseline to an intermediate program for each of these selected partitions. Those with higher income, females, and those without a college degree are willing to pay more for better hurricane forecasts (differences of \$4.05, \$0.71, and \$6.73, respectively).

6. Conclusions and future research

The stated mission of the National Hurricane Center (NHC) is "to save lives, mitigate property loss, and improve economic efficiency by issuing the best watches, warnings, forecasts and analyses of hazardous tropical weather, and by increasing understanding of these

TABLE 4. Example WTP calculations for improvements from baseline to intermediate level on all forecast attributes based on choice sets 2–8.*

Attribute	Baseline	Intermediate improvement	Difference	Marginal WTP (b_k/b_{cost}; \$)	WTP (\$)
Time of expected landfall	8 h 48 h in advance	6 h	2 h	2.18	4.36
Max wind speed	20 mi h^{-1} 48 h in advance	15 mi h^{-1}	5 mi h^{-1}	0.26	1.30
Projected location of landfall	100 mi 48 h in advance	80 mi 48 h in advance	20 mi	0.23	4.60
Expected storm surge	±8 ft of height above sea level 48 h in advance	6 ft of height above sea level	2 ft	2.04	4.08
Total WTP					14.34

* Derived via a bivariate probit with quadrature model.

TABLE 5. Modeling based on selected subsamples and WTP for intermediate program.*

	Income < \$60,000, $n = 336$			Income ≥ \$60,000, $n = 224$		
	Coef	*t* ratio	WTP (\$)	Coef	*t* ratio	WTP (\$)
Landfall time	−0.081	−3.219	2.29	−0.102	−2.420	3.07
Wind speed	−0.009	−0.952	0.26	−0.008	−0.506	0.25
Location	−0.009	−3.061	0.26	−0.01	−2.106	0.29
Surge	−0.065	−2.717	1.83	−0.093	−2.282	2.80
Cost	−0.035	−9.298		−0.033	−6.400	
Mean log likelihood	−1.15939			−1.24338		
WTP for intermediate program (\$)			14.74			18.79
	Males, $n = 280$			Females, $n = 280$		
	Coef	*t* ratio	WTP (\$)	Coef	*t* ratio	WTP (\$)
Landfall time	−0.053	−2.064	1.86	−0.083	−3.024	2.53
Wind speed	−0.008	−0.720	0.27	−0.008	−0.793	0.24
Location	−0.007	−2.278	0.25	−0.007	−2.310	0.22
Surge	−0.057	−2.182	1.98	−0.066	−2.696	2.04
Cost	−0.029	−7.367		−0.033	−8.489	
Mean log likelihood	−1.2157			−1.17603		
WTP for intermediate program (\$)			14.03			14.74
	College graduate, $n = 280$			Not a college graduate,** $n = 280$		
	Coef	*t* ratio	WTP (\$)	Coef	*t* ratio	WTP (\$)
Landfall time	−0.049	−1.885	1.69	−0.083		2.61
Wind speed	−0.008	−0.776	0.27	−0.008		0.24
Location	−0.002	−0.799	0.08	−0.012		0.38
Surge	−0.067	−2.600	2.30	−0.058		1.82
Cost	−0.029	−7.440		−0.032		
Mean log likelihood	−1.19119			−1.19177		
WTP for intermediate program (\$)			10.93			17.66

* Derived via a bivariate probit with quadrature; choice sets 2–8 only.

** In the estimation for the not college graduate subsample, the likelihood converged but would not invert, so that no *t* ratios could be calculated.

hazards'' (full text available online at http://www.nhc.noaa.gov/mission.shtml). Given the significant impacts on society of hurricanes and an apparent disconnect between the quality of forecasts and the societal responses in events such as Hurricanes Katrina and Rita, understanding behavior with respect to hurricanes and forecasts and warnings is critical to meeting societal goals such as those elucidated by the NHC. We developed and implemented a survey of households in Miami, Florida, to explore behavior with respect to hurricane risks and preferences for forecast and warning information.

Evacuation decision making is modeled as a function of the level of hurricane threat as represented by the SSHS and a range of risk perceptions, behavioral constraints, and sociodemographic characteristics. Individuals perceiving a higher level of potential flood damage, a higher level of perceived accuracy of hurricane forecasts, and older respondents were all more likely to evacuate than other respondents. Those not willing to leave their property unprotected, longer in their residence, and of higher income indicated a lower likelihood of evacuation. In general, responses are consistent with a priori expectations and will be fertile ground for future research.

We used a discrete choice valuation approach to assess the value of improving hurricane forecasts. Respondents showed a significant WTP for improved information on several components of hurricane forecasts, including projected timing and location of landfall and the likely magnitude of the storm surge and wind speed—factors that determine the likely magnitude of the impacts of a hurricane. We cannot draw direct conclusions on the relative values of improvements in the different attributes because each is measured on a different scale. As illustrated, we can, however, use the marginal values for improving the attributes of hurricane forecasts to derive values for specific forecast improvement programs that would improve these attributes to certain levels. This approach can be used to derive values for any potential program combining improvements in the range of improvements considered: for example, leaving some attributes at baseline, improving some to an intermediate level, and improving others to the maximum level. Not surprisingly, we found differences in preferences and values between different segments of society. Such differences need further exploration in light of the critical issue of vulnerable populations and the differential impacts of hurricanes on these populations (Phillips and Morrow 2007).

In an evacuation decision making model (Table 1), we did not find a significant relationship between wind risk perception and likelihood of evacuation. This is consistent with findings reported in Table 3 that the marginal utility of improvements in accuracy of maximum wind speed was not significant. This is a reasonable finding given that wind speed is generally not considered as important a concern for evacuation as storm surge. Flood risk perceptions were marginally significant (12% level as report in Table 1) in driving evacuation intentions and had significant value in the WTP modeling. The results on income indicate higher-income respondents were less likely to evacuate (Table 1), whereas higher income respondents also have a higher WTP for improved forecasts (Table 5). Economic theory would suggest higher WTP for higher-income respondents, but we do not have a priori expectations about the relationship between evacuation likelihood and income. For instance, higher-income respondents may live in less hurricane-vulnerable properties or be more reluctant to leave their property unprotected.[6] Future research should look at these interrelationships in greater detail.

Although the current results show the viability of these methods for exploring evacuation decision making and values for improved hurricane forecasts and warnings, the empirical results cannot be generalized to the entire hurricane-vulnerable population. Using results from future larger random samples, such derived value estimates will be amenable to use in benefit–cost analyses of potential programs.

Our future work on the communication and value of hurricane forecasts will expand the current work to a larger, more geographically diverse, stratified random sample. We will focus on a similar but slightly modified range of forecast attributes of specific relevance to develop benefit estimates for evaluating programs such as NOAA's Hurricane Forecast Improvement Project (NOAA 2008). Even with such focused valuation efforts, there remains a critical need for "a long-term, multidisciplinary, institutional approach" for social science research to address the larger range of issues on all aspects of the hurricane forecast and warning system (Gladwin et al. 2009).

[6] We thank a reviewer for insightful comments on the consistencies between the evacuation modeling and WTP analysis.

Acknowledgments. The authors thank Julie Demuth, Kim Klockow, Rebecca Morss, Andrea Schumacher, and three anonymous referees for helpful comments on the article. Any remaining errors are those of the authors.

REFERENCES

Adamowicz, W., P. Boxall, M. Williams, and J. Louviere, 1998: Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation. *Amer. J. Agric. Econ.,* **80,** 64–75.

Allison, P. D., 1999: *Logistic Regression Using the SAS System: Theory and Application*. SAS Publishing, 308 pp.

Baker, E. J., 1993: Empirical studies of public response to tornado and hurricane warnings in the United States. *Prediction and Perception of Natural Hazards,* J. Nemec, J. M. Nigg, and F. Siccardi, Eds., Kluwer, 65–74.

Ballinger, G. A., 2004: Using generalized estimating equations for longitudinal data analysis. *Organ. Res. Methods,* **7,** 127–150.

Ben-Akiva, M., and S. Lerman, 1985: *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press, 416 pp.

Dash, N., and H. Gladwin, 2007: Evacuation decision making and behavioral responses: Individual and household. *Nat. Hazards Rev.,* **8** (3), 69–77.

Dow, K., and S. L. Cutter, 1998: Crying wolf: Repeat responses to hurricane evacuation orders. *Coast. Manage.,* **26,** 237–252.

EIDS, cited 2008: Miami-Dade County (FL), December 2008. Economic Intelligence Data Service. [Available online at https://edis.commerce.state.nc.us/docs/countyProfile/FL/12086.pdf.]

Ericsson, K. A., and H. A. Simon, 1993: *Protocol Analysis: Verbal Reports as Data*. The MIT Press, 496 pp.

Gladwin, C. H., H. Gladwin, and W. G. Peacock, 2001: Modeling hurricane evacuation decisions with ethnographic methods. *Int. J. Mass Emergencies Disasters,* **19,** 117–143.

Gladwin, H., and W. G. Peacock, 1997: Warning and evacuation: A night for hard houses. *Hurricane Andrew: Ethnicity, Gender and the Sociology of Disasters,* W. G. Peacock, B. H. Morrow, and H. Gladwin, Eds., Routledge, 52–74.

——, J. K. Lazo, B. Morrow, W. G. Peacock, and H. E. Willoughby, 2007: Social science research needs for the hurricane forecast and warning system. *Nat. Hazards Rev.,* **8** (3), 87–95.

——, ——, ——, ——, and ——, 2009: Social science research needs for the hurricane forecast and warning system. *Bull. Amer. Meteor. Soc.,* **90,** 25–29.

Greene, W. H., 2003: *Econometric Analysis*. 5th ed. Prentice-Hall, 806 pp.

Huber, J., and K. Zwerina, 1996: The importance of utility balance in efficient choice designs. *J. Marketing Res.,* **33,** 307–317.

Kuhfeld, W. F., 2005: Marketing research methods in SAS: Experimental design, choice, conjoint, and graphical techniques. SAS Institute TS-722, 49 pp.

Lazo, J. K., 2002: Assessment of total household benefits of improved hurricane forecasting: First interim report. SC10111, Stratus Consulting Inc., Boulder, CO, 213 pp.

——, 2004: Draft summary on Miami one-on-one cognitive interviews—Hurricane forecast valuation study. Societal Impacts Program, National Center for Atmospheric Research, 24 pp.

Letson, D., D. Sutter, and J. K. Lazo, 2007: The economic value of hurricane forecasts: An overview and research needs. *Nat. Hazards Rev.,* **8** (3), 78–86.

Lindell, M. K., and C. S. Prater, 2007: Critical behavioral assumptions in evacuation time estimate analysis for private vehicles: Examples from hurricane research and planning. *J. Urban Planning Development,* **133,** 18–29.

Louviere, J. J., D. A. Hensher, and J. D. Swait, 2001: *Stated Choice Methods: Analysis and Application*. Cambridge University Press, 418 pp.

Manski, C., 1977: The structure of random utility models. *Theory Decision,* **8,** 229–254.

McFadden, D., 1976: The revealed preferences of a government bureaucracy: Empirical evidence. *Bell J. Econ. Manage. Sci.,* **7,** 55–72.

Morrow, B. H., 2009: Risk behavior and risk communication: Synthesis and expert interviews. SocResearch Rep. for NOAA/Coastal Services Center. [Available online at http://www.csc.noaa.gov/Risk_Behavior_&_Communication_Report.pdf.]

——, and H. Gladwin, 2005: Hurricane Ivan Behavioral Analysis. Report prepared for FEMA and U.S. Army Corps of Engineers, 66 pp. [Available online at http://chps.sam.usace.army.mil/USHESDATA/Assessments/2004Storms/PDFfiles/Ivan%20Final%20Behave%20download.pdf.]

NOAA, cited 2008: Proposed framework for addressing the National Hurricane Research and Forecast Improvement Initiatives: NOAA's Hurricane Forecast Improvement Project. [Available online at http://www.nrc.noaa.gov/plans_docs/HFIP_Plan_073108.pdf.]

O'Connell, A. A., 2005: *Logistic Regression Models for Ordinal Response Variables*. Sage Publications, 120 pp.

Phillips, B. D., and B. H. Morrow, 2007: Social science research needs: Focus on vulnerable populations, forecasting, and warnings. *Nat. Hazards Rev.,* **8** (3), 61–68.

Savage, S., and D. M. Waldman, 2008: Learning and fatigue during choice experiments: A comparison of online and mail survey modes. *J. Appl. Econ.,* **23,** 351–371.

U.S. Department of Housing and Urban Development–U.S. Department of Commerce, 2003: American housing survey for the Miami–Ft. Lauderdale metropolitan area: 2002. Current Housing Reports, Series H170/02-28, U.S. Census Bureau, 232 pp. + appendixes. [Available online at http://www.census.gov/prod/2003pubs/h170-02-28.pdf.]

Whitehead, J. C., 2005: Environmental risk and averting behavior: Predictive validity of jointly estimated revealed and stated behavior data. *J. Environ. Resource Econ.,* **32,** 301–316.

——, B. Edwards, M. Van Willigen, J. R. Maiolo, K. Wilson, and K. T. Smith, 2000: Heading for higher ground: Factors affecting real and hypothetical hurricane evacuation behavior. *Environ. Hazards,* **2,** 133–142.

Willoughby, H. E., 2002: Improvements in observations, models and forecasts. *HURRICANE! Coping with Disaster,* R. H. Simpson, R. A. Anthes, and M. Garstang, Eds., Amer. Geophys. Union, 205–216.

Yin, R., and D. H. Newman, 1999: An intervention analysis of Hurricane Hugo's price effect in South Carolina. *Can. J. For. Res.,* **29,** 779–786.

Zhang, F., and Coauthors, 2007: An in-person survey investigating public perceptions of and response to Hurricane Rita forecasts along the Texas coast. *Wea. Forecasting,* **22,** 1177–1190.