# False Alarms, Tornado Warnings, and Tornado Casualties

KEVIN M. SIMMONS

*Department of Economics, Austin College, Sherman, Texas*

DANIEL SUTTER

*Department of Economics and Finance, University of Texas-Pan American, Edinburg, Texas*

(Manuscript received 7 November 2008, in final form 17 June 2009)

ABSTRACT

This paper extends prior research on the societal value of tornado warnings to the impact of false alarms. Intuition and theory suggest that false alarms will reduce the response to warnings, yet little evidence of a ''false alarm effect'' has been unearthed. This paper exploits differences in the false-alarm ratio across the United States to test for a false-alarm effect in a regression model of tornado casualties from 1986 to 2004. A statistically significant and large false-alarm effect is found: tornadoes that occur in an area with a higher false-alarm ratio kill and injure more people, everything else being constant. The effect is consistent across false-alarm ratios defined over different geographies and time intervals. A one-standard-deviation increase in the false-alarm ratio increases expected fatalities by between 12% and 29% and increases expected injuries by between 14% and 32%. The reduction in the national tornado false-alarm ratio over the period reduced fatalities by 4%–11% and injuries by 4%–13%. The casualty effects of false alarms and warning lead times are approximately equal in magnitude, suggesting that the National Weather Service could not reduce casualties by trading off a higher probability of detection for a higher false-alarm ratio, or vice versa.

## 1. Introduction

A trade-off exists in warning for tornadoes between the probability of detection (POD) and false alarms. In the limit, false alarms could be avoided by never issuing warnings, whereas warning for every thunderstorm will increase the POD to 100%. The National Weather Service (NWS) uses the POD and the false-alarm ratio (FAR) as metrics for tornado warning performance, reflecting the intuition that both measures affect the public's response to warnings. Recently Simmons and Sutter (2008) showed that tornado warnings with lead times up to about 17 min reduce tornado fatalities and injuries, but to date no evidence exists on false alarms and casualties.

We examine the influence of false alarms on casualties using tornadoes in the contiguous United States from 1986 to 2004. Our unit of observation is the individual tornado, but false alarms are nonevents. How precisely false alarms will affect tornado casualties is unclear. We hypothesize that residents will use the recent local FAR in assessing warning credibility, and construct multiple FARs based on different geographies and time horizons. We then include the FAR variable in a regression analysis of tornado casualties. We do not examine the link between false alarms and warning response directly, but since warning response affects casualties and our 19 yr of data include over 21 000 state tornado segments, we seek to detect an impact on response through casualties.

We find a strong relationship between the FAR and both tornado fatalities and injuries. We define FARs based on three different geographies [states, NWS Weather Forecast Offices (WFO), and television markets] and two time horizons (the past year and past two years). The results are remarkably consistent across the various definitions of FAR: a one-standard-deviation (1 std dev) increase in the FAR increases expected fatalities by between 12% and 29% and expected injuries by between 14% and 32%. The reduction in the national FAR between 1986 and 2004, although modest, reduced expected fatalities by 4%–11% and expected injuries by 4%–13%, based on the definition of FAR. We also evaluate the trade-off between FAR and POD depicted by Brooks (2004) on casualties. The point estimates of

*Corresponding author address:* Daniel Sutter, Department of Economics and Finance, University of Texas-Pan American, Edinburg, TX 78539-2999.
E-mail: dssutter@utpa.edu

DOI: 10.1175/2009WCAS1005.1

the impact of warnings and false alarms on casualties nearly offset each other, so we find no evidence that the NWS could reduce casualties by trading off a higher POD for higher FAR, or vice versa.

This paper is organized as follows: The next section discusses the intuition behind the false-alarm or cry-wolf effect, as well as evidence on false alarms and warning response to different natural hazards. Defining false-alarm variables for use in an analysis of casualties from tornadoes is challenging because false alarms are nonevents. Section 3 defines our false-alarm ratios and present preliminary evidence on the influence of false alarms on casualties. Section 4 defines the other variables employed in the analysis and discusses the econometric specification, and section 5 presents the regression results. Section 6 considers two extensions of our analysis, to consider the decline in tornado casualties attributable to the reduction in the false-alarm ratio since 1986, and the casualty trade-off between the probability of detection and false alarms in warning for tornadoes. The final section contains conclusions.

## 2. Theory of response to tornado warnings and the role of false alarms

The value of a weather forecast or warning depends on the accuracy of the forecast. A tornado warning is a binary forecast, and so a $2 \times 2$ matrix can be applied to attempt to estimate the value of the forecast (Katz and Murphy 1997). The error probabilities of the warning—the probability that a tornado occurs without a warning being issued (a missed event), and the probability that a tornado does not occur when a warning is issued (a false alarm)—drive the value of the warning. The first error probability is $1 - \mathrm{POD}$, and the second error is the FAR.[1] The expected-utility approach predicts that people will respond to warnings they receive when the net expected utility of response is positive. An increase in either error probability reduces the value of a warning. The net expected value depends on the losses avoided by response (the reduction in likelihood of being killed or injured) and the cost of response (interrupting activities to take cover) in addition to the error probabilities. If the FAR is sufficiently high, the net value of responding to a warning becomes negative.

Intuition and theory suggest that false alarms should reduce warning response, yet evidence on this in field studies is sparse. Breznitz (1984, p. 11) defines the false-alarm effect as the "credibility loss due to a false alarm" and describes the effect as follows: "Each false alarm reduces the credibility of a warning system. The credibility loss following a false alarm episode has serious ramifications to behavior in a variety of response channels." In the framework of the value of forecasts and warnings, the effect occurs when residents judge the FAR to be sufficiently high that warnings are ignored. Evidence of a false-alarm effect in field studies, however, is sparse. Breznitz (1984), for example, provides evidence consistent with the effect, but from a laboratory setting, not for warning responses in the field. Field research found a reduction in credibility for earthquake predictions after a quake predicted by I. Browning on the New Madrid fault for December 1990 failed to occur (Atwood and Major 1998). The relevance of this evidence is questionable, because the warning was issued by a private individual instead of a government agency, and there was no scientific basis for such a targeted prediction. Field evidence against a false-alarm effect is provided by Dow and Cutter's (1998) study of hurricane evacuation. They found no evidence of a lower evacuation rate in South Carolina for Hurricane Fran in 1996, which occurred just weeks after a false-alarm evacuation for Hurricane Bertha, and conclude, "The influence of 'premature evacuations' for Bertha played only a minor role in evacuation decisions for Hurricane Fran in South Carolina" (p. 249). Recently Barnes et al. (2007) observed, "Evidence for the cry-wolf effect in natural hazards research, however, has not been forthcoming." Based on this, researchers may not be surprised if false alarms did not increase tornado casualties.

We test for a false-alarm effect for tornado warnings using NWS warning verification records and casualties from over 20 000 tornadoes across the contiguous United States. We will use the actual warning performance, or the objective FAR as calculated from NWS warning records, in our analysis. But residents' warning response in the expected utility model depends on their subjective perception of warning quality, that is, what they perceive the FAR to be and not the objective FAR. Subjective perceptions may diverge from objective measures, yet researchers do not have access to the subjective estimates of warning quality for the general population. Our approach is valid if subjective perceptions are distributed around the objective probability, with some people overestimating and others underestimating the FAR. Furthermore, our approach really only requires that differences in the objective FAR affect casualties, and not that the average of residents' subjective perceptions of FAR be equal to the objective FAR.

Another limitation with NWS warning verification records is the potential that the warning performance measures may not correlate closely with residents' perceptions

[1] POD = warned tornadoes/total number of tornadoes; FAR = unverified tornadoes/total warnings.

of hazard events. Barnes et al. (2007) discuss several such potential problems and argue that an event that is technically a false alarm in the NWS warning statistics may be regarded by residents as a near miss and thus a credible warning. For instance, false alarms can result from a storm that generated a funnel cloud that did not quite touch down, or a tornado may have occurred just outside of the warning area or just after expiration of the warning. The issuance of tornado warnings based on counties until October 2007 could also artificially inflate the NWS FAR. For example, if four counties were warned for a supercell thunderstorm that produced a strong, dangerous tornado in only one county, the official FAR for this event would be 0.75, approximately equal to the national average FAR, even though residents may consider this to be a well-warned event and increase their subjective evaluation of warning reliability as a consequence. Barnes et al. suggest that such close calls and near misses may not reduce the public's confidence in tornado warnings. Thus a poor correlation between what officially counts as a false alarm and subjective evaluation could also explain a lack of an impact of false alarms on casualties.

We test for an impact of false alarms on tornado casualties, not warning response. Warning response affects casualties but is not the only determinant of casualties. We use a dataset with numerous control variables, because false alarms (or warnings) will only increase (decrease) casualties if all other factors affecting casualties are held constant. The control variables include the Fujita-scale (F scale) rating of the storm, time of day, month, and storm-path characteristics. The use of a large dataset of over 20 000 tornadoes should help to smooth out any variation in warning response. Our national dataset allows us to exploit considerable differences in FAR to detect a possible influence of false alarms.

## 3. False-alarm variable definitions

Theory tells us that warning performance should affect subjective estimates of warning quality, but theory does not inform us of exactly *how* residents form their perceptions. Also, our data points are tornado events. The warning issued for a tornado (if any) thus can be naturally included as a control variable for the individual tornado. But false alarms are nonevents and cannot be matched with specific tornadoes. We hypothesize that residents will use recent, local warnings to try to estimate the FAR. Residents need to use warnings over some period of time to estimate the true FAR yet distant events fade from memory and warning performance varies over time, most notably with the introduction of Doppler weather radar (Simmons and Sutter 2005). Residents should use local warnings as well, meaning warnings that they hear or see, to estimate the quality of warnings. Residents of Kansas are unlikely to use warnings issued for New York, Florida, or California to evaluate the reliability of warnings.

Because what constitutes local and recent is subject to differing interpretations, we construct FARs using three geographies and two time horizons to ensure the robustness of any findings. The geographies are 1) states, 2) NWS WFO County Warning Areas (CWAs), based on the CWAs of the WFOs of the modernized NWS, and 3) television markets as defined by the A. C. Nielsen Company.[2] We apply 1- and 2-yr lags in constructing FARs for each geography and refer to these as 1- and 2-yr state, WFO, and TV FARs. If a tornado occurs in, say, May 2004, the 1-yr window would be all warnings in the relevant geographic area in the months of May 2003–April 2004 and the 2-yr window would be May 2002–April 2004. All tornadoes in the same geography in the same month have the same FAR.

We do not expect that residents actually know the FAR for their state or TV market over the period. Our approach should not be interpreted too literally. We hope that residents' subjective perceptions of warning quality correlate with our objective measures. That is, residents of a state with a 1-yr FAR of 0.95 should be more likely to think that "warnings are always false alarms" than residents of a state with a 1-yr FAR of 0.25. Our approach relies on these impressions correlating with actual recent local warning performance. If more residents take warnings seriously where recent warning performance has been good, the improved warning response should reduce casualties, and our regression models should detect the casualty reduction.

A complication arises when tornadoes occur in a geography where no warnings have been issued in the previous year or two. The FAR is undefined in this case, and it is unclear how residents will treat these warnings. Most of these tornadoes occur in states with low tornado frequency. The number of tornadoes with no warnings in the area over the period is 332 for 1-yr state, 770 for 1-yr WFO, and 1215 for 1-yr TV markets. Enough tornadoes occur with no warnings for most of the definitions that we code the FAR equal to 1 for these tornadoes and include a dummy variable "no warnings" as a control. This allows us to include these tornadoes without restricting how these warnings affect casualties. Too few

[2] We apply current CWA through our sample, even though WFOs were reorganized with the modernization of the NWS in the 1990s. The counties contained in each Nielsen Designated Market Area are reported in the annual editions of the *Broadcasting and Cable Yearbook*.

Table 1. Summary statistics for FAR variables. The FAR variables are defined based on all warnings in the past year or two years in the state, NWS WFO, or television market where the tornado touched down.

| FAR definition | Mean | Median | Std dev |
|---|---|---|---|
| 1-yr state | 0.754 | 0.754 | 0.117 |
| 1-yr WFO | 0.746 | 0.765 | 0.173 |
| 1-yr TV | 0.764 | 0.770 | 0.187 |
| 2-yr state | 0.751 | 0.747 | 0.100 |
| 2-yr WFO | 0.748 | 0.756 | 0.137 |
| 2-yr TV | 0.752 | 0.753 | 0.148 |

tornadoes occur with no warnings in the state over the previous 2 yr to control for with a dummy variable. As a result, the regressions omit these tornadoes.

Our empirical design seeks to exploit variation in the FAR across locations and time to test for a false-alarm effect. This design requires sufficient variation in FARs to affect residents' perceptions of warning quality and response to warnings. The summary statistics exhibit substantial variation, with 1-yr FARs ranging from 0 to 1, and 2-yr FARs ranging from under 0.5 to 1 (Table 1). The mean and median FAR are about 0.75 by each definition. The standard deviations are slightly larger with the 1-yr than 2-yr FARs, as a longer window of observation allows a more precise estimate of the FAR. The distribution of tornadoes across intervals for each of the six FAR definitions confirms the variation (Table 2). For the 1-yr FARs, between 9.5% and 21.5% of tornadoes occur with a FAR in excess of 0.9 while between 7% and 17% have FARs less than 0.6. Considerable variation exists within states as well; 1-yr FARs for most states exhibit a range of 0.5 or greater for most states. Figure 1 displays the state 1-yr FAR for Kansas over the period, which is typical of tornado-prone states, with variation from 0.5 to 0.9. Considerable variation in warning performance also exists across the country. Figure 2 displays a scatterplot of the FAR for the period of analysis for each of the WFOs in our dataset, along with the POD as calculated from our dataset (i.e., state tornadoes as reported in the Storm Prediction Center archive as opposed to county tornadoes). The FAR for all warnings over the period ranges from just over 0.5 to 1.0 for one WFO while the POD ranges from less than 0.1 to 0.8. The distribution of types of tornado outbreaks varies across the country (Trapp et al. 2005); to provide perspective on how this affects the variation in our data, we also report the aggregate FAR and POD over the period by state, for states with at least 100 tornadoes over the period (Table 3).

Table 4 reports preliminary evidence on the relationship between FARs and casualties. The table presents fatalities and injuries per tornado for each of the intervals from Table 2 and reveals a positive relationship between FARs and casualties. Table 3 reports the statistics for the 1-yr FAR definitions, and the pattern is similar for 2-yr definitions. The highest fatalities per tornado and highest or second highest injuries per tornado occur in the 0.9–0.999 interval. The intervals for 0.8 or greater feature the three highest fatality rates for state 1 yr and TV 1 yr, and three of the four highest rates for WFO 1 yr. The difference in rates is also pronounced, with the casualties per tornado for the smallest FAR intervals generally less than half of that for the 0.9–0.999 interval. The relationship across intervals is not monotonic, and casualties per tornado do not control for F-scale ratings or tornado or path characteristics. Nonetheless, this first look at the data suggests the existence of a false-alarm effect.

## 4. Control variables and the regression model

Tornado casualties take on nonnegative integer values and are count data. Ordinary least squares (OLS) estimation does not account for the censoring of the dependent variable at zero with count data (i.e., that casualties cannot take on negative values). Economists typically employ a Poisson regression model for analysis of count data of this type. The Poisson regression model specifies

Table 2. Distribution of tornadoes by FAR interval. The numbers are the percentage of tornadoes in each category, with 1986 tornadoes excluded in the 1-yr variables and 1986–87 tornadoes excluded for the 2-yr variables.

| | Percentage of tornadoes by interval for FAR definition | | | | | |
|---|---|---|---|---|---|---|
| FAR interval | 1-yr state | 1-yr WFO | 1-yr TV | 2-yr state | 2-yr WFO | 2-yr TV |
| No warnings | 1.61 | 3.72 | 5.88 | 0.78 | 1.86 | 2.71 |
| 1 | 3.43 | 9.65 | 14.29 | 1.76 | 4.44 | 7.31 |
| 0.90–0.999 | 6.09 | 7.77 | 6.27 | 3.52 | 6.43 | 5.84 |
| 0.80–0.899 | 25.40 | 22.95 | 18.27 | 24.89 | 23.01 | 21.08 |
| 0.75–0.799 | 16.68 | 12.44 | 12.29 | 18.35 | 17.72 | 16.06 |
| 0.70–0.749 | 16.73 | 10.49 | 9.83 | 22.25 | 12.63 | 11.86 |
| 0.60–0.699 | 23.21 | 15.94 | 17.71 | 22.38 | 20.00 | 21.40 |
| 0.50–0.599 | 5.08 | 10.48 | 8.78 | 5.81 | 10.77 | 10.49 |
| 0–0.499 | 1.77 | 6.56 | 6.69 | 0.53 | 3.15 | 3.26 |

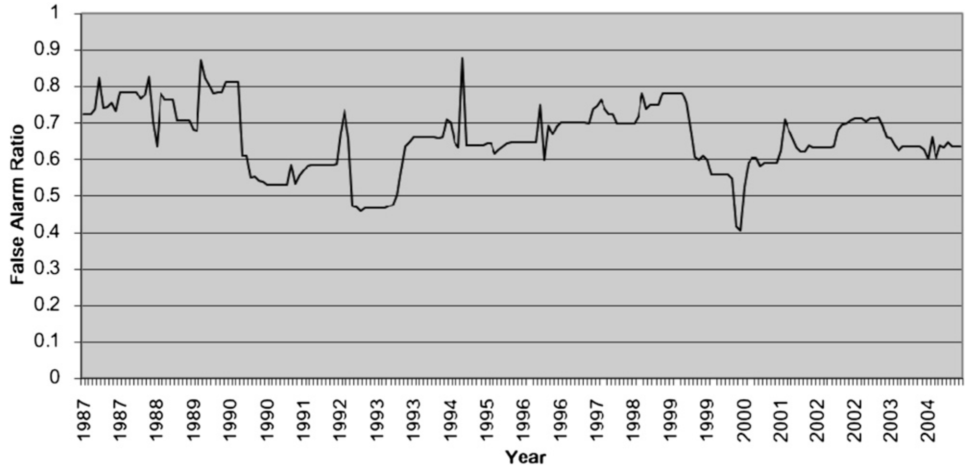


FIG. 1. State-based FAR for Kansas.

that the dependent variable $y_i$, in this application the number of persons killed or injured in a state tornado, is drawn from a Poisson distribution (Greene 2000, 880–886). The probability of a given number of casualties is

$$\text{Prob}(Y_i = y_i) = \exp(-\lambda_i)\lambda_i \frac{y_i}{y_i!}, \quad y_i = 0, 1, 2, \ldots.$$

The parameter $\lambda_i$ of the distribution is assumed to be a log-linear function of the independent variables $\mathbf{x}_i$:

$$\ln\lambda_i = \beta'\mathbf{x}_i.$$

The Poisson model assumes equality of the conditional mean and variance of the dependent variable, and violation of this assumption is known as either overdispersion or underdispersion. The negative binomial generalization of the Poisson model is employed when count data exhibit overdispersion (Greene 2000, 886–888). The negative binomial regression model generalizes the Poisson model by introducing an individual, unobserved effect into the conditional mean, so

$$\ln\mu_i = \beta'\mathbf{x}_i + \varepsilon_i.$$

Thus the distribution of $y_i$ conditioned on $\mathbf{x}_i$ and $\varepsilon_i$ remains Poisson. Tornado injuries but not fatalities exhibit overdispersion, and so we report Poisson models of fatalities and negative binomial models of injuries.

We estimate the following regression models of tornado fatalities and injuries:

$$\text{Casualties} = f(\text{Warning Quality, F Scale, Density, Income, Mobile, Length, Length} \times \text{Density, Season, Day, Evening, Weekend, Year}).$$

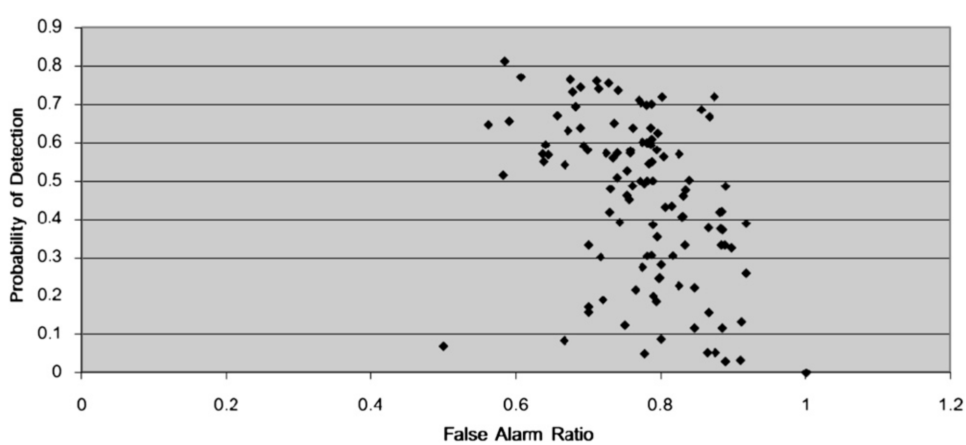


FIG. 2. POD and FAR by NWS WFO.

TABLE 3. POD and FAR by state. States with fewer than 100 tornadoes over the period of 1987–2004 are not included.

| State name | POD | FAR |
|---|---|---|
| Alabama | 0.6624 | 0.8501 |
| Arkansas | 0.6771 | 0.7052 |
| California | 0.1444 | 0.8526 |
| Colorado | 0.5381 | 0.7279 |
| Florida | 0.3474 | 0.7978 |
| Georgia | 0.5000 | 0.8426 |
| Idaho | 0.2273 | 0.8397 |
| Illinois | 0.6230 | 0.7839 |
| Indiana | 0.5023 | 0.7793 |
| Iowa | 0.5104 | 0.6949 |
| Kansas | 0.6920 | 0.6495 |
| Kentucky | 0.4841 | 0.8225 |
| Louisiana | 0.5060 | 0.7802 |
| Maryland | 0.4162 | 0.6820 |
| Michigan | 0.3510 | 0.7988 |
| Minnesota | 0.5661 | 0.7521 |
| Mississippi | 0.6402 | 0.8223 |
| Missouri | 0.6448 | 0.7653 |
| Montana | 0.3095 | 0.8022 |
| Nebraska | 0.5989 | 0.7682 |
| New Mexico | 0.3113 | 0.7838 |
| New York | 0.2745 | 0.8035 |
| North Carolina | 0.4913 | 0.7794 |
| North Dakota | 0.4915 | 0.6042 |
| Ohio | 0.5200 | 0.7963 |
| Oklahoma | 0.7048 | 0.6824 |
| Pennsylvania | 0.4140 | 0.7549 |
| South Carolina | 0.6284 | 0.6133 |
| South Dakota | 0.4992 | 0.7490 |
| Tennessee | 0.6450 | 0.8425 |
| Texas | 0.6417 | 0.7152 |
| Virginia | 0.4845 | 0.7562 |
| Wisconsin | 0.4698 | 0.7960 |
| Wyoming | 0.3926 | 0.7184 |

Our regression analysis includes a number of variables affecting tornado casualties in addition to measures of false alarms. Prior empirical work (Simmons and Sutter 2005, 2008) informs expectations for the variables. The models include year dummy variables to control for factors that vary across the entire nation yet are not captured by our other variables, factors like the emergence of the Internet or cell phones. In addition, tornado casualties might exhibit year-to-year variation if, for instance, a year with few tornadoes lulls residents into a false sense of security the next year. The year variable coefficients are not reported to conserve space in the tables. For a list of all variable names and brief descriptions see Table 5.

We include several different tornado warning variables. Tornado warnings for the period of our data were issued by county whereas our data points are state tornado segments from the Storm Prediction Center tornado archive. Consequently, a tornado that strikes more than one county in a state may have several valid warnings. We apply the warning for the first county in the storm path. Our main warning variables use the lead time for a warning at the start of the tornado; note that for long-path tornadoes the warning at the beginning of the storm path is less representative of the time that residents along the storm path might have to respond. We break the warning lead time down into categories, which allows the marginal effect of lead time to vary. Dummy variables lead 1–5, lead 6–10, lead 11–15, lead 16–20, lead 21–30, and lead 31+ control for the intervals, and respectively equal 1 if the lead time for the warning falls in the corresponding interval and 0 otherwise. For robustness we also use a categorical variable—warning—that equals 1 if a warning was issued for the tornado and 0 otherwise.

We include several tornado characteristic variables, all constructed from the Storm Prediction Center archive. The most important storm characteristic is the tornado's rating on the Fujita scale of tornado damage. The ratings take on integer values from 0 to 5, with 5 being the strongest.[3] We use categorical (or dummy) variables labeled F0, F1, F2, F3, F4, and F5 to represent the F scale. The variable F3, for example, equals 1 if a tornado was rated 3 on the F scale and 0 otherwise. The category dummy variables do not impose any functional relationship between F scale and casualties. More damaging tornadoes should produce more fatalities and injuries, but the categorical variables might indicate that F1 tornadoes are only slightly more deadly than F0 tornadoes, but that F4 tornadoes are much more deadly than F3 tornadoes. The regression model omits the F0 category, and so the coefficients on the included category variables indicate the effect on casualties relative to an F0 tornado.

Other variables include ''length''—the tornado track length in tenths of a mile. Longer track tornadoes have the potential to kill and injure more people, so we expect a positive sign on length. We control for time of day with the dummy variables of day, evening, and night. ''Day'' equals 1 if the tornado occurs between 0600 and 1759, ''evening'' equals 1 if the tornado occurs between 1800 and 2359, and ''night'' equals 1 for tornadoes between 0000 and 0559 (all times local). The regressions omit night, and so the coefficients on day and evening show the effect of tornadoes at these times relative to the late-night hours. Research demonstrates that tornadoes at night are more likely to catch residents unprepared and

[3] Technically the F scale is a measure of damage (Doswell and Burgess 1988), but it is widely reported as a measure of intensity and is the closest available proxy for storm strength.

TABLE 4. Fatalities and injuries per tornado by FAR interval.

| FAR interval | 1-yr state Fatalities | 1-yr state Injuries | 1-yr WFO Fatalities | 1-yr WFO Injuries | 1-yr TV Fatalities | 1-yr TV Injuries |
|---|---|---|---|---|---|---|
| No warning | 0.0181 | 0.961 | 0.0610 | 1.72 | 0.0115 | 0.621 |
| 1 | 0.0490 | 0.956 | 0.0706 | 1.06 | 0.0755 | 1.47 |
| 0.90–0.999 | 0.114 | 2.07 | 0.0828 | 1.62 | 0.129 | 2.11 |
| 0.80–0.899 | 0.0512 | 1.13 | 0.0566 | 0.958 | 0.0455 | 1.03 |
| 0.75–0.799 | 0.0488 | 0.848 | 0.0348 | 0.787 | 0.0299 | 0.571 |
| 0.70–0.749 | 0.0436 | 0.683 | 0.0356 | 0.751 | 0.0408 | 0.737 |
| 0.60–0.699 | 0.0228 | 0.545 | 0.0167 | 0.554 | 0.0279 | 0.468 |
| 0.50–0.599 | 0.0272 | 0.708 | 0.0409 | 0.904 | 0.0375 | 0.814 |
| 0–0.499 | 0.0246 | 0.822 | 0.0332 | 0.462 | 0.0174 | 0.362 |

produce more casualties, so we expect negative signs for day and evening (Ashley 2007; Ashley et al. 2008; Simmons and Sutter 2005, 2008). "Weekend" is a dummy variable equal to 1 if the tornado occurred on a Saturday or Sunday. Previous work has shown that casualties are higher on weekends. Season controls for the month of the year of the tornado and equals 1 for tornadoes during March, April, May, or June and 0 for tornadoes in any other month. Casualties are higher for tornadoes during the off season when controlling for F scale and other factors (Simmons and Sutter 2008).

We include three path characteristic variables. The path characteristic variables were based on U.S. census values for the county struck by the tornado. The values for a county for a given year are estimated by linear interpolation from the decennial censuses. Although tornadoes are small relative to counties, county-based path variables have exhibited strong explanatory power in prior regression analysis (Simmons and Sutter 2005, 2008). For tornadoes that struck more than one county, the storm value averages the county values for the year in question. The first path variable is "density", or persons per square mile for the tornado path.[4] Casualties should increase as population density increases. We also combine density and length because a long-track tornado through a highly populated area may affect casualties differently than an increase in each variable separately. "Mobile" is mobile homes as a percentage of county housing units.[5] The susceptibility of mobile homes to damage from tornadic winds has long been recognized (Golden and Adams 2000; Schmidlin et al. 2001; Brooks and Doswell 2002), and confirmed in previous research. "Income" is median family income in thousands of 1999 dollars.[6] Economists typically find that people spend more to protect themselves against risk as they become wealthier (Viscusi et al. 2000, chapter 19), and cross-national analyses of natural hazard fatalities have confirmed this negative relationship between income and hazard lethality (Anbarci et al. 2005; Kahn 2005). For tornadoes, higher-income individuals might own an in-home tornado shelter and buy National Oceanic and Atmospheric Administration (NOAA) weather radios or other emergency alert systems. Wealthier communities could be more likely to have tornado sirens and more effective emergency management and rescue services. Note, however, that tornadoes appear to be an exception to this pattern; Simmons and Sutter (2005, 2008) have found a strong, positive relationship between income and both fatalities and injuries.

## 5. Regression results

Regression analysis of tornado fatalities reveals that each of the six definitions of a FAR described above is a positive and statistically significant (at better than the 0.01 level) determinant of fatalities (Table 6). Figure 3 graphs the estimated impact of a one-standard-deviation (1 std dev) increase in the different definitions of the FAR on increases in fatalities, calculated using the point estimates from Table 6 and the standard deviations of the FAR variables from Table 1. Figure 3 also displays the upper and lower bounds of the 95% confidence interval (CI) for the 1-std-dev increase in FAR. An increase in the FAR increases expected fatalities by between 12% and 29% across the different definitions. The FAR variables based on two years of tornado warnings have larger effects on fatalities. The variable, "no warnings," which controlled for cases in which no warnings had been issued

[4] The 2001–04 values of density are based on the census's annual estimate of county population.

[5] Mobile homes by county were not reported in 1980 census publications, so the value of this variable for 1986–89 tornadoes is taken from the 1990 census. Mobile home figures from the 2000 census are used for 2001–04 tornadoes.

[6] Nominal dollar values were converted to 1999 dollars using the Consumer Price Index for All Urban Consumers (CPI-U). The change in real per capita personal income in each county since 1999 was used to produce values of this variable for 2000–04.

TABLE 5. List of variable names and descriptions.

| Variable name | Description |
|---|---|
| No warnings | A dummy variable that equals 1 if no tornado warnings were issued in the geography and period in question; the FAR for these tornadoes was set to 1.0 |
| Lead 1–5 | A dummy variable that equals 1 if the tornado had a warning with 1–5-min lead time |
| Lead 6–10 | A dummy variable that equals 1 if the tornado had a warning with 6–10-min lead time |
| Lead 11–15 | A dummy variable that equals 1 if the tornado had a warning with 11–15-min lead time |
| Lead 16–20 | A dummy variable that equals 1 if the tornado had a warning with 16–20-min lead time |
| Lead 21–30 | A dummy variable that equals 1 if the tornado had a warning with 21–30-min lead time |
| Lead 31+ | A dummy variable that equals 1 if the tornado had a warning with a lead time of 31 min or more |
| Warning | A dummy variable that equals 1 if the tornado was warned for |
| Doppler | A dummy variable that equals 1 if the tornado occurred on or after the date of installation of Doppler radar in the local NWS WFO |
| Density | Population density in persons per square mile |
| Income | Median family income in thousands of 1999 dollars |
| Mobiles | Mobile homes as a proportion of housing units |
| Day | A dummy variable that equals 1 for tornadoes occurring between 0600 and 1759 local time |
| Evening | A dummy variable that equals 1 for tornadoes occurring between 1800 and 2359 local time |
| Weekend | A dummy variable that equals 1 for tornadoes that occur on a Saturday or Sunday |
| Season | A dummy variable that equals 1 for tornadoes occurring in the months of March, April, May, or June |
| Length | The length of the tornado path in miles |

in the previous 1 or 2 yr and the FAR was undefined and was included in all specifications except 2-yr state, is significant and negative for the 1-yr state and both TV definitions, with point estimates slightly larger than that of the FAR variable. Recall that tornadoes with no warnings were also coded with FAR = 1. Thus the coefficients on no warnings and the FAR must be combined to infer the impact of tornadoes occurring with no recent, local warnings. Consequently, for the 1-yr state and TV market definitions of FAR, tornadoes with no recent warnings affect fatalities similar to a FAR of 0. On the other hand, no warnings is insignificant with a small point estimate using the 1-yr WFO definition and is positive and significant for 2-yr WFO. With these definitions of FAR, tornadoes with no warnings affect casualties similarly to (or more than) a FAR of 1.[7]

The signs, significance, and magnitude of the control variables are similar to previous results (Simmons and Sutter 2005, 2008). All attain significance at the 0.10 level except some of the lead-time-interval variables. The magnitudes are quite similar across different specifications of the FAR variable. Figure 4 graphs the point estimates of warnings in the various lead-time intervals on fatalities for the 1-yr and 2-yr state FAR variables. In the figure, the effect of a tornado with no lead time is normalized to 100, so the figure displays the impact of warning lead times relative to no lead time. Warnings with lead times up to 11–15 min reduce expected fatalities when compared with no warning, and a 6–10-min lead time reduces fatalities by 46%. Yet the point estimates indicate that lead times greater than 16 min increase expected fatalities relative to a tornado with no warning lead time. The impact is statistically significant and quite large, with a warning with a 16–20-min lead time increasing fatalities by 60%–70% relative to an unwarned tornado. Although anecdotes of increasing risky activities for long–lead time tornadoes exist, such behaviors would have to be quite prevalent to more than offset the effect of some people sheltering. This surprising result is due to a handful of tornadoes with long lead times and large fatality totals.[8] The storm-path variables density, income, and mobiles all increase expected fatalities, consistent with previous results. Timing also matters significantly for tornado fatalities. Figure 5 displays the effect of a tornado occurring during the evening or night periods relative to a day tornado. Expected fatalities from day tornadoes are normalized to equal 100, and the figure displays the increase or decrease for evening and night tornadoes based on the point estimates in Table 6 for each of the FAR definition specifications. Expected fatalities were 67% and 63% lower for tornadoes that occur during the day or evening relative to overnight tornadoes with the 1-yr FAR variables; the effects are just slightly higher for the 2-yr FAR variables. The time-of-day effects provide indirect evidence on the effectiveness of the tornado warning

[7] For robustness, all of the models were estimated with the tornadoes omitted for which no warnings had been issued over the prior window. Omission of these tornadoes did not affect the inferences regarding the FAR or other variables in any meaningful fashion.

[8] For an extended discussion of these results see Simmons and Sutter (2008).

TABLE 6. False alarms and tornado fatalities (Poisson regression model with standard errors in parentheses). Pseudo $R^2$ is 1 − (log likelihood of the full model/log likelihood of a reduced model with only the dependent variable), $R^2_{\mathrm{LRI}} = 1 - L(\tilde{\lambda}_i, y_i)/L(\bar{y}, y_i)$.

| | 1-yr state | 1-yr WFO | 1-yr TV | 2-yr state | 2-yr WFO | 2-yr TV |
|---|---|---|---|---|---|---|
| FAR | 0.971[a] | 1.04[a] | 1.13[a] | 2.53[a] | 1.82[a] | 1.11[a] |
| | (0.329) | (0.240) | (0.214) | (0.431) | (0.331) | (0.219) |
| No warning | −1.16[b] | 0.192 | −1.41[a] | — | 0.418[b] | −1.20[a] |
| | (0.423) | (0.168) | (0.280) | — | (0.201) | (0.282) |
| Lead 1–5 | −0.136 | −0.135 | −0.229[c] | −0.0887 | −0.0453 | −0.120 |
| | (0.133) | (0.134) | (0.133) | (0.136) | (0.136) | (0.134) |
| Lead 6–10 | −0.619[a] | −0.584[a] | −0.668[a] | −0.610[a] | −0.640[a] | −0.733[a] |
| | (0.138) | (0.138) | (0.139) | (0.150) | (0.149) | (0.150) |
| Lead 11–15 | −0.350[b] | −0.337[b] | −0.397[b] | −0.381[b] | −0.358[b] | −0.452[a] |
| | (0.158) | (0.158) | (0.158) | (0.159) | (0.160) | (0.159) |
| Lead 16–20 | 0.494[a] | 0.514[a] | 0.457[a] | 0.571[a] | 0.565[a] | 0.472[a] |
| | (0.126) | (0.126) | (0.126) | (0.129) | (0.129) | (0.128) |
| Lead 21–30 | 0.410[a] | 0.422[a] | 0.394[a] | 0.208[c] | 0.209[c] | 0.150 |
| | (0.111) | (0.111) | (0.111) | (0.119) | (0.120) | (0.119) |
| Lead 31+ | 0.147 | 0.222[c] | 0.198 | 0.206 | 0.220[c] | 0.160 |
| | (0.125) | (0.126) | (0.125) | (0.129) | (0.129) | (0.126) |
| Density | 0.181[a] | 0.196[a] | 0.263[a] | 0.192[a] | 0.192[a] | 0.272[a] |
| | (0.0592) | (0.0612) | (0.0643) | (0.0611) | (0.0620) | (0.0631) |
| Income | 0.0257[a] | 0.0246[a] | 0.256[a] | 0.304[a] | 0.0332[a] | 0.0332[a] |
| | (0.0049) | (0.0050) | (0.0050) | (0.0049) | (0.0048) | (0.0048) |
| Mobiles | 5.79[a] | 5.63[a] | 5.96[a] | 5.93[a] | 6.21[a] | 6.34[a] |
| | (0.523) | (0.523) | (0.521) | (0.544) | (0.540) | (0.541) |
| Day | −1.10[a] | −1.12[a] | −1.09[a] | −1.04[a] | −1.10[a] | −1.05[a] |
| | (0.0867) | (0.0864) | (0.0861) | (0.0682) | (0.0884) | (0.0882) |
| Evening | −0.987[a] | −0.996[a] | −1.00[a] | −1.04[a] | −1.09[a] | −1.06[a] |
| | (0.104) | (0.104) | (0.104) | (0.0882) | (0.110) | (0.110) |
| Weekend | 0.499[a] | 0.540[a] | 0.479[a] | 0.606[a] | 0.617[a] | 0.593[a] |
| | (0.0778) | (0.0788) | (0.0784) | (0.0818) | (0.0824) | (0.0816) |
| Season | −0.170[b] | −0.148[c] | −0.190[b] | −0.215[b] | −0.209[c] | −0.297[b] |
| | (0.0812) | (0.0806) | (0.0813) | (0.0845) | (0.0843) | (0.0841) |
| Length | 0.0007[a] | 0.0008[a] | 0.0008[a] | 0.0010[a] | 0.0011[a] | 0.0011[a] |
| | (0.0002) | (0.0002) | (0.0002) | (0.0002) | (0.0002) | (0.0002) |
| Length × density | 0.0017[a] | 0.0017[a] | 0.0014[b] | 0.0015[b] | 0.0014[b] | 0.0012[b] |
| | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0006) |
| F1 | 2.80[a] | 2.80[a] | 2.80[a] | 2.79[a] | 2.80[a] | 2.80[a] |
| | (0.373) | (0.373) | (0.373) | (0.373) | (0.374) | (0.374) |
| F2 | 4.68[a] | 4.67[a] | 4.69[a] | 4.64[a] | 4.64[a] | 4.64[a] |
| | (0.363) | (0.363) | (0.363) | (0.364) | (0.364) | (0.362) |
| F3 | 6.44[a] | 6.40[a] | 6.44[a] | 6.37[a] | 6.36[a] | 6.40[a] |
| | (0.361) | (0.361) | (0.361) | (0.361) | (0.362) | (0.362) |
| F4 | 7.96[a] | 7.95[a] | 8.00[a] | 7.70[a] | 7.74[a] | 7.76[a] |
| | (0.395) | (0.365) | (0.365) | (0.366) | (0.366) | (0.366) |
| F5 | 10.4[a] | 10.4[a] | 10.4[a] | 10.3[a] | 10.2[a] | 10.3[a] |
| | (0.380) | (0.380) | (0.380) | (0.379) | (0.380) | (0.379) |
| Constant | −7.91[a] | −8.00[a] | −8.08[a] | −10.9[a] | −10.5[a] | −9.74[a] |
| | (0.512) | (0.484) | (0.478) | (0.581) | (0.544) | (0.498) |
| Pseudo $R^2$ | 0.601 | 0.601 | 0.605 | 0.603 | 0.604 | 0.604 |
| Log-likelihood | −1938 | −1934 | −1918 | −1805 | −1800 | −1801 |

[a] Statistical significance at the 0.01 level in a two-tailed test.
[b] Statistical significance at the 0.05 level in a two-tailed test.
[c] Statistical significance at the 0.10 level in a two-tailed test.

process, since plausibly residents will be less likely to receive and respond to a tornado warning at night. Tornadoes are also much more deadly on weekends, with expected fatalities 61%–85% higher relative to weekdays. Tornadoes occurring between March and June are less deadly relative to other months, with a reduction in fatalities of 14%–26% depending on the FAR specification. Expected fatalities increase dramatically with F scale, with an F5 tornado being 30 000 times more lethal than an F0 tornado when controlling

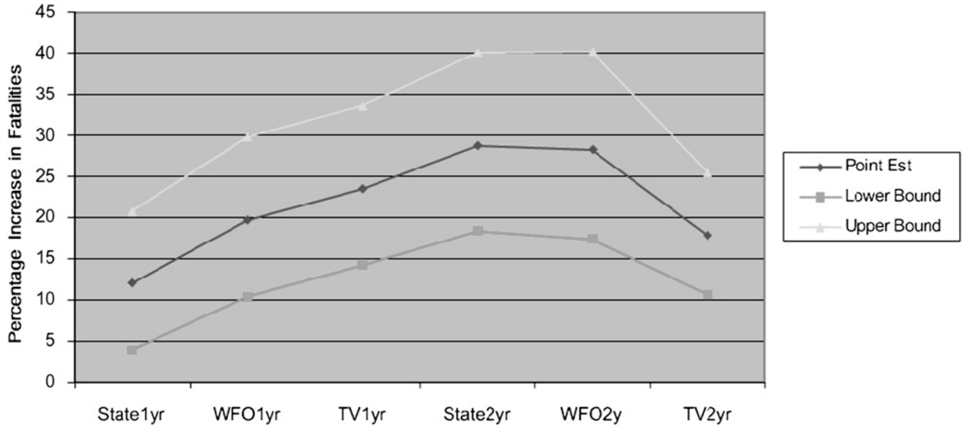


FIG. 3. Effect of a one-std-dev increase in FAR on fatalities.

for other determinants of fatalities. Length and the length × density interaction term also increase fatalities.

The regression models reveal that false alarms also significantly increase injuries (Table 7). Each of the six definitions of the FAR is significant at better than the 0.01 level. Figure 6 displays the effect of a 1-std-dev increase in each of the FAR variables on injuries based on the point estimates, along with the upper and lower bounds of the 95% CI. A higher FAR increases expected injuries by between 13% and 32%, very similar to the range for fatalities, with larger impacts for the WFO FARs than for the state FARs. The variable no warnings, controlling for tornadoes that occur with an undefined local, recent FAR, fails to attain significance for injuries, in contrast to the results for fatalities and rendering inferences about how residents treat rare tornado warnings as problematic. Tornado injuries occur more often than fatalities, and thus fatality analysis contains more "noise," especially given the relatively small number of tornadoes occurring with no recent, local warnings. So in assessing the results for fatalities and injuries, the mixed results for fatalities and the consistently insignificant results for injuries suggest that residents probably treat tornadoes with no previous, recent warnings as similar to high FAR warnings.

As with fatalities, the signs, significance, and magnitude of the effect of the control variables on injuries is similar to previous research (Simmons and Sutter 2005, 2008) and is similar across the definitions of the FAR variable. The control variables affect both fatalities and injuries similarly, with the few exceptions noted below. The main differences between the injury and fatality models are for the warning-lead-time variables. Figure 7 depicts the effect of a warning lead time in each interval relative to a no-lead-time tornado (set at 100) for the 1-yr and 2-yr state specifications. Warnings reduced expected injuries for all lead-time intervals, while lead times under 15 min reduced fatalities. All of the lead-time variables are significant for all FAR specifications except in the 16–20-min interval, which attains significance for 1-yr state. The largest reduction for injuries in most cases occurs for lead 31+, which reduces injuries by about 42%, and this reduction is consistent across the FAR definitions. Reductions in expected injuries of around 40% also

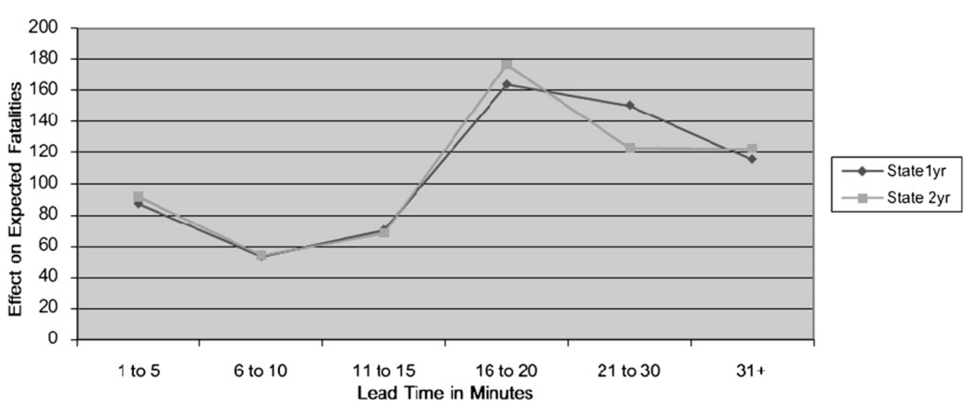


FIG. 4. Warning lead time and tornado fatalities.

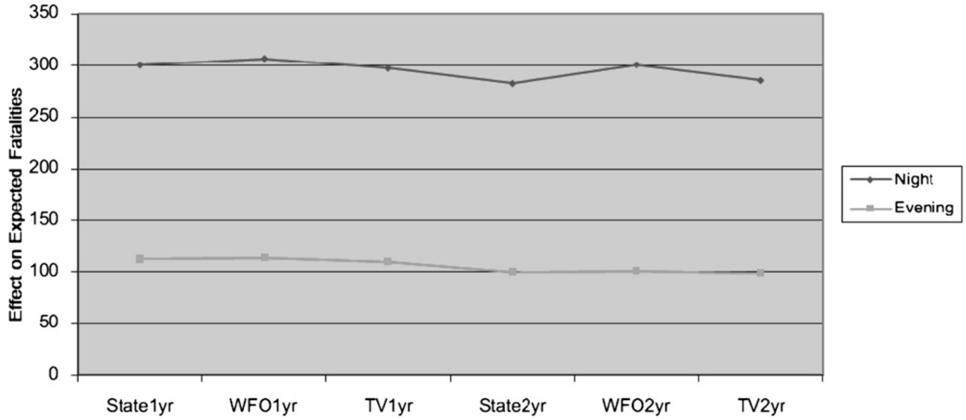


FIG. 5. Time of day and tornado fatalities.

occur in the 11–15-min interval, and so the marginal value of a 31-min lead time relative to an 11–15-min lead time is almost zero. The most notable remaining difference between the injuries and fatalities analysis concerns weekend tornadoes, which increase injuries by less than 10% and fail to attain significance for each FAR definition, whereas tornadoes on weekends significantly increased fatalities. Figure 8 displays the impact of evening and night tornadoes on injuries, relative to a tornado during the day. The magnitude of the effect of several control variables is smaller for injuries than for fatalities, notably time of day (injuries are about 34% and 38% lower during the day or evening relative to night). The effect of F scale on injuries is substantial, with expected injuries 1300 times as great for an F5 tornado than for an F0 tornado, but is less than the effect of F scale on fatalities.

Table 8 presents fatality and injury models for the 1-yr state FAR variables for two alternative measures of warnings—a dummy variable controlling for whether a warning was in effect for a tornado and a second dummy variable "Doppler." Doppler controls for whether the NWS WFO, with warning responsibility for the first county in the tornado path, had weather surveillance Radar 1988-Doppler (WSR-88D) installation on or before the day of the tornado and has the value of 1 if Doppler radar had been installed and 0 otherwise. Only the coefficients for the warning and false-alarm variables are reported; the control variables from the full models were all included and had the same signs and magnitude impacts as in Tables 6 and 7. The FAR variable is significant (at better than the 0.01 level) in each case with a similar magnitude effect on casualties as in Tables 6 and 7. A warning in effect decreases fatalities by 2.7% (but is not significant) and decreases injuries by a statistically significant 30%. Doppler radar reduces fatalities by 34% and injuries by 35% with both effects statistically significant. The estimated impact of Doppler radar on casualties is smaller than the 45% reduction in fatalities and 40% reduction in injuries reported in Simmons and Sutter (2005), in part because the specifications in Table 8 include the FAR, which was reduced following Doppler radar installation. Thus, part of the impact of Doppler radar here is captured in the FAR variables. The FAR variables retain their significance and magnitude, so the false-alarm effect is robust to alternative definitions of the warning variables.

## 6. Extensions

Our results provide strong evidence of a false-alarm effect, meaning that a higher false-alarm ratio increases casualties. This does not establish a false-alarm "problem," because to us the term problem suggests a higher-than-optimal FAR. The tornado warning process must trade off a higher FAR for a higher POD, and that both elements of the trade-off are lethal does not imply a problem. We extend the analysis here by examining the reduction in casualties attributable to a lowered FAR over the sample period and the casualties trade-off between the POD and FAR.

### a. *Casualty reduction attributable to FAR improvements*

A higher FAR increases expected tornado fatalities and injuries. The FAR on tornado warnings has improved over time (Brooks 2004; Simmons and Sutter 2005). We quantify here the casualty reduction attributable to the observed reduction in FAR.

We begin by calculating the mean FAR for all tornadoes in a year for each definition of the FAR. So, for example, for 1-yr state, the mean FAR for 1990 would be the mean of the 1-yr state FAR for all tornadoes during 1990. Because of the lag used in constructing our FAR variables, the mean 1-yr state FAR for 1990 will depend on warnings issued prior to 1990. We use the mean FAR

TABLE 7. False alarms and tornado injuries (Poisson regression model with standard errors in parentheses).

| | 1-yr state | 1-yr WFO | 1-yr TV | 2-yr state | 2-yr WFO | 2-yr TV |
|---|---|---|---|---|---|---|
| FAR | 1.11[a] | 1.60[a] | 0.872[a] | 1.79[a] | 2.01[a] | 0.877[a] |
| | (0.266) | (0.213) | (0.164) | (0.335) | (0.259) | (0.166) |
| No warning | −0.203 | −0.0621 | −0.165 | — | −0.188 | −0.151 |
| | (0.251) | (0.163) | (0.136) | — | (0.220) | (0.138) |
| Lead 1–5 | −0.455[a] | −0.449[a] | −0.434[a] | −0.451[a] | −0.499[a] | −0.419[a] |
| | (0.124) | (0.124) | (0.125) | (0.124) | (0.124) | (0.125) |
| Lead 6–10 | −0.340[a] | −0.268[a] | −0.302[b] | −0.350[a] | −0.311[a] | −0.320[a] |
| | (0.120) | (0.119) | (0.120) | (0.121) | (0.121) | (0.122) |
| Lead 11–15 | −0.524[a] | −0.535[a] | −0.550[a] | −0.511[a] | −0.513[a] | −0.535[a] |
| | (0.127) | (0.119) | (0.128) | (0.127) | (0.127) | (0.128) |
| Lead 16–20 | −0.228[c] | −0.221 | −0.222 | −0.197 | −0.195 | −0.215 |
| | (0.138) | (0.138) | (0.138) | (0.139) | (0.139) | (0.139) |
| Lead 21–30 | −0.314[b] | −0.326[b] | −0.333[a] | −0.303[b] | −0.291[b] | −0.345[a] |
| | (0.120) | (0.120) | (0.120) | (0.120) | (0.120) | (0.121) |
| Lead 31+ | −0.549[a] | −0.539[a] | −0.551[a] | −0.522[a] | −0.509[a] | −0.544[a] |
| | (0.125) | (0.125) | (0.125) | (0.125) | (0.125) | (0.125) |
| Density | 0.623[a] | 0.634[a] | 0.649[a] | 0.608[a] | 0.623[a] | 0.649[a] |
| | (0.0923) | (0.0926) | (0.0931) | (0.0926) | (0.0934) | (0.0945) |
| Income | 0.0160[a] | 0.0162[a] | 0.0167[a] | 0.0151[a] | 0.0168[a] | 0.0168[a] |
| | (0.0043) | (0.0043) | (0.0043) | (0.0043) | (0.0043) | (0.0043) |
| Mobiles | 4.33[a] | 4.34[a] | 4.44[a] | 4.24[a] | 4.29[a] | 4.44[a] |
| | (0.439) | (0.437) | (0.435) | (0.443) | (0.440) | (0.441) |
| Day | −0.410[a] | −0.411[a] | −0.402[a] | −0.406[a] | −0.434[a] | −0.408[a] |
| | (0.0907) | (0.0908) | (0.0907) | (0.0919) | (0.0920) | (0.0920) |
| Evening | −0.497[a] | −0.476[a] | −0.481[a] | −0.486[a] | −0.469[a] | −0.496[a] |
| | (0.104) | (0.105) | (0.104) | (0.106) | (0.106) | (0.106) |
| Weekend | 0.0672 | 0.0534 | 0.0638 | 0.0729 | 0.0780 | 0.0779 |
| | (0.0703) | (0.0734) | (0.0703) | (0.0715) | (0.0712) | (0.0716) |
| Season | −0.238[a] | −0.229[a] | −0.234[a] | −0.243[a] | −0.233[a] | −0.248[a] |
| | (0.0662) | (0.0663) | (0.0662) | (0.0676) | (0.0677) | (0.0676) |
| Length | 0.0030[a] | 0.0030[a] | 0.0032[a] | 0.0029[a] | 0.0030[a] | 0.0032[a] |
| | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0006) | (0.0006) |
| Length × density | 0.0101[a] | 0.0102[a] | 0.0099[b] | 0.0102[a] | 0.0099[a] | 0.0100[a] |
| | (0.0025) | (0.0025) | (0.0025) | (0.0025) | (0.0025) | (0.0026) |
| F1 | 2.54[a] | 2.54[a] | 2.54[a] | 2.56[a] | 2.56[a] | 2.56[a] |
| | (0.0802) | (0.0803) | (0.0802) | (0.0813) | (0.0815) | (0.0814) |
| F2 | 4.30[a] | 4.31[a] | 4.31[a] | 4.31[a] | 4.31[a] | 4.31[a] |
| | (0.102) | (0.102) | (0.102) | (0.102) | (0.103) | (0.103) |
| F3 | 5.34[a] | 5.37[a] | 5.31[a] | 5.40[a] | 5.30[a] | 5.28[a] |
| | (0.156) | (0.156) | (0.156) | (0.163) | (0.156) | (0.157) |
| F4 | 6.73[a] | 6.79[a] | 6.69[a] | 6.56[a] | 6.60[a] | 6.53[a] |
| | (0.262) | (0.262) | (0.262) | (0.263) | (0.262) | (0.264) |
| F5 | 7.35[a] | 7.27[a] | 7.21[a] | 7.33[a] | 7.20[a] | 7.22[a] |
| | (0.900) | (0.897) | (0.901) | (0.896) | (0.899) | (0.900) |
| Constant | −5.54[a] | −5.80[a] | −5.43[a] | −5.89[a] | −6.22[a] | −5.27[a] |
| | (0.332) | (0.310) | (0.298) | (0.346) | (0.327) | (0.279) |
| Pseudo $R^2$ | 0.183 | 0.185 | 0.184 | 0.184 | 0.185 | 0.183 |
| Log-likelihood | −9510 | −9493 | −9505 | −9201 | −9185 | −9202 |

[a] Statistical significance at the 0.01 level in a two-tailed test.
[b] Statistical significance at the 0.05 level in a two-tailed test.
[c] Statistical significance at the 0.10 level in a two-tailed test.

for all tornadoes in a year, so 1990s 1-yr state FAR will depend more heavily on recent warning performance in the regions with the most active tornado years. These annual FARs will not correspond with the FARs reported by the NWS for the year since our FARs depend on warnings issued prior to 1990 because of the 1-yr lag.

Next we regress the annual FAR (for each definition) on a constant and a linear time trend over the period 1987–2004 using OLS. We use the estimated coefficient on the time-trend variable to impute the reduction in FAR over our sample period (Table 9). The point estimates indicate a reduction in the FAR over the period of

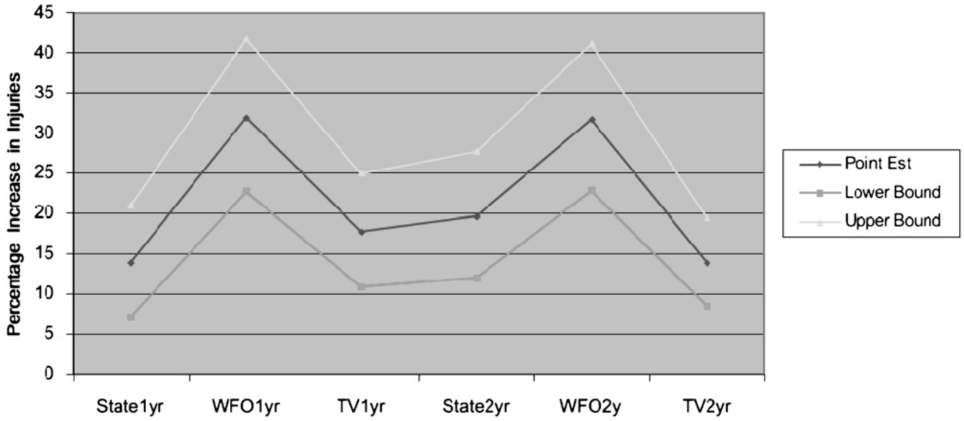


FIG. 6. Effect of a one-std-dev increase in FAR on injuries.

our study of between 0.036 and 0.087, depending on the definition applied, which is only about one-half of the difference between the first- and last-year FARs. Based on the regression coefficients for fatalities and injuries (Tables 6, 7), we can then estimate the effect of the observed decline in FAR. The point estimates indicate reductions in fatalities of between 4% and 11%, while the reduction in injuries is between 5% and 12%.

### b. The trade-off between warnings and false alarms

The warning process for tornadoes trades off the POD and FAR. Forecasters could increase the POD by warning for every thunderstorm but would produce innumerable false alarms. The FAR could be reduced by issuing warnings only when a tornado is reported to be on the ground, but POD and lead time would fall. Our regression models include both warnings and false alarms and thus we can quantify the casualty trade-off and examine if NWS forecasters could reduce casualties by trading off a higher FAR for a higher POD, or vice versa.

The observed *increase* in fatalities for warnings with lead times over 15 min relative to no warning complicates this evaluation. The point estimates for lead times over 15 min suggest that issuing fewer warnings would produce fewer fatalities. But these results are due to a handful of particularly deadly tornadoes and do not appear to be robust. Consequently we will make calculations assuming that an increase in POD leads to warnings with lead times in the 0–15-min range, based on the relative frequency of lead times in these intervals, and apply the casualty reductions observed in these intervals. The specification with 1-yr state in Table 6 implies reductions in fatalities of 13%, 46%, and 29%, respectively, for lead 1–5, lead 6–10, and lead 11–15, and the relative frequency of warnings in these three intervals was 0.332, 0.352, and 0.316. Combining these frequencies and fatality reductions yields a 29.8% reduction in expected fatalities when a tornado is warned for. The reduction in expected injuries for these lead time intervals is 37%, 29%, and 41%, so a warning for a tornado reduces expected injuries by 35.2%.

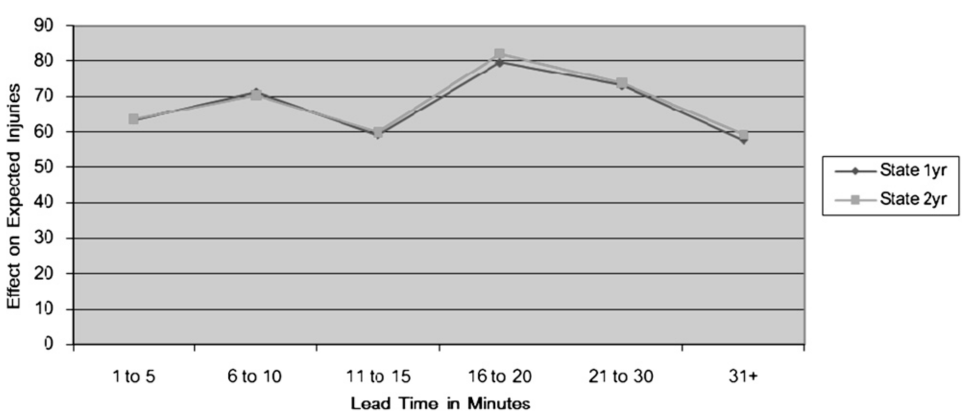


FIG. 7. Warning lead time and tornado injuries.

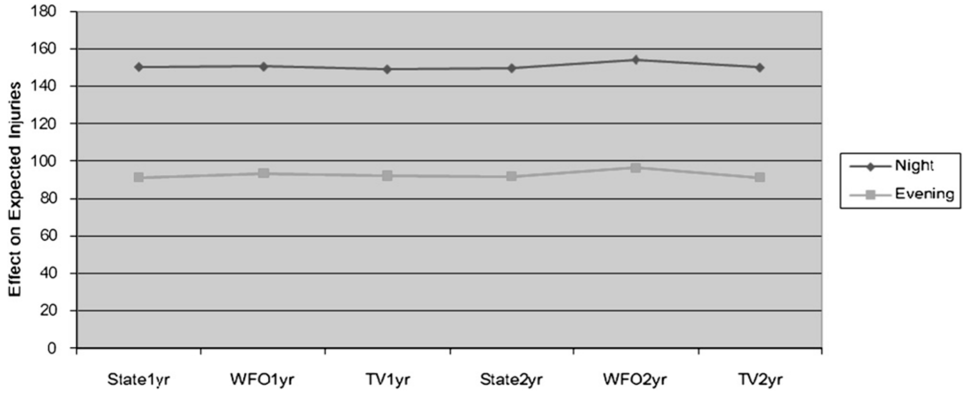


FIG. 8. Time of day and tornado injuries.

Increasing the POD results in more tornadoes being warned for, which results in these tornadoes being less dangerous, but the corresponding increase in FAR applies to all tornadoes. The trade-off between a hypothetical increase in the POD and increase in the FAR can be illustrated as follows: suppose that the POD can be increased by 0.10 but as a result the FAR increased by 0.10. With the 1-yr state FAR, warning for an additional 10% of tornadoes reduces expected fatalities for these tornadoes by 29.8% and for all tornadoes by 2.98%; the increase in FAR applies to all tornadoes and increases expected fatalities by 10.2%. In this case, the hypothetical change in warning strategy would increase expected fatalities by 6.97%. The same change in warnings would increase expected injuries by 8.22%.

We now turn to a more accurate depiction of the trade-off between the POD and FAR. Brooks (2004) uses national warning performance, signal detection theory, and an assumption about the climatological frequency of tornadoes to trace out the apparent trade-off between POD and FAR. The curve fitting the warning performance between 1990 and 2000 is approximately linear in the region between (POD = 0.4, FAR = 0.7) and (POD = 1.0, FAR = 0.9). Assuming that this is an exploitable trade-off for the tornado warning process, what would be the consequences for casualties of shifting from (POD = 0.4, FAR = 0.7) to (POD = 1.0, FAR = 0.9)? Table 10 presents the analysis using each of the six different definitions of FAR. To illustrate, consider the calculations for 1-yr state. An extra 60% of tornadoes would occur with a warning, and, given the 29.8% and 35.2% reductions in expected fatalities and injuries due to a warning, increasing the POD from 0.4 to 1.0 would reduce fatalities by 17.9% and injuries by 21.1%. Based on the point estimates in Tables 6 and 7, the 0.20 increase in 1-yr state FAR would increase expected fatalities and injuries by 21.4% and 24.9%. Overall, the two effects nearly cancel out, producing a 0.3% decrease in fatalities and a 1.5% decrease in injuries. The FAR and warning effects essentially cancel out for other FAR definitions,

TABLE 8. Further analysis of false alarms and tornado casualties (Poisson regression model with standard errors in parentheses). The specifications each contain all of the control variables in Tables 6 and 7, but these coefficients are not reported here to conserve space. The full results are available from the authors.

| | Fatalities | | Injuries | |
|---|---|---|---|---|
| FAR | 0.707[a] | 0.811[b] | 1.15[b] | 1.14[b] |
| | (0.343) | (0.330) | (0.279) | (0.264) |
| No warnings | −1.11[a] | −1.16[a] | 0.0861 | −0.235 |
| | (0.423) | (0.423) | (0.366) | (0.253) |
| Doppler | −0.419[a] | — | −0.438[c] | — |
| | (0.203) | — | (0.172) | — |
| Warning | — | −0.0270 | — | −0.360[b] |
| | — | (0.0895) | — | (0.0691) |
| Pseudo $R^2$ | 0.593 | 0.592 | 0.182 | 0.183 |
| Log-likelihood | −1976 | −1978 | −9528 | −9517 |

[a] Statistical significance at the 0.05 level in a two-tailed test.
[b] Statistical significance at the 0.01 level in a two-tailed test.
[c] Statistical significance at the 0.10 level in a two-tailed test.

TABLE 9. Effect of observed change in FAR 1987–2004 on tornado casualties.

| | 1-yr state | 1-yr WFO | 2-yr state | 2-yr WFO |
|---|---|---|---|---|
| FAR 1987/88 | 0.819 | 0.820 | 0.820 | 0.834 |
| FAR 2004 | 0.738 | 0.699 | 0.741 | 0.726 |
| Change in FAR (from linear regression) | −0.041 | −0.067 | −0.036 | −0.062 |
| Reduction in fatalities | −3.9% | −6.7% | −8.6% | −10.7% |
| Reduction in injuries | −4.6% | −9.1% | −8.2% | −11.7% |

except with 2-yr state and 2-yr WFO, where the net impact is increases of 37% and 20%. For injuries, 1-yr WFO, 2-yr state, and 2-yr WFO yield net increases of 10%–18% while the other measures yield small decreases. Table 9 also reports the net impact on casualties with the upper and lower bounds of the 95% confidence interval for the FAR variables. In only three cases do the confidence intervals for fatalities or injuries not include zero. Consequently, our analysis finds no evidence of potential casualty reductions from either increasing the FAR or increasing the POD, given the trade-off faced in the warning process.[9]

## 7. Conclusions

This paper has investigated the impact of false alarms on tornado casualties and, by implication, response to warnings. Intuition and theory suggest that if residents place little confidence in warnings, if they think that warnings are "always false alarms," they might ignore warnings, resulting in higher casualties. But evidence of a false-alarm effect in the field for hazard warnings generally has been elusive. We have found strong evidence that a higher local, recent FAR significantly increases tornado fatalities and injuries, justifying its use as a performance metric for the NWS. For robustness we constructed FAR variables using three different geographies for local and 1- and 2-yr windows for recent; the finding is robust for all of these plausible definitions of a local, recent FAR. The positive relationship between FAR and casualties is apparent even in tabulations of fatalities and injuries per tornado. Each different FAR variable attains statistical significance in a regression analysis, and the magnitudes are consistent across definitions, with a 1-std-dev increase in the FAR increasing expected fatalities by 12%–29% and expected injuries by 13%–32%.

Our approach has several limitations that suggest avenues for future research. First, as mentioned, we have examined casualties and not warning response directly. This allows us to use a dataset of over 20 000 tornadoes to exploit differences in warning performance across the nation to identify a false-alarm effect. But we have no direct evidence that false alarms affect the perceived credibility of warnings. Future survey or qualitative research might focus on the link between local recent false alarms and the perception of warning quality or warning response. Second, our demographic control variables are based on the counties struck by tornadoes, which are large when compared with the damage paths of individual tornadoes. While the county-based variables attain significance, and in the expected direction (with the exception of income), future research could attempt to refine the damage paths of the tornadoes. Last, as in all econometric work, omitted variables may explain some portion of the effect on casualties we attribute here to false alarms or other variables.

We stress once more in concluding that evidence that false alarms affect casualties does not imply the existence of a false-alarm "problem." The term problem suggests that the FAR is higher than optimal given the trade-off between the FAR and POD. Given the trade-off between FAR and POD described in Brooks (2004), a reduction in FAR, given the offsetting decrease in POD, would not appear to reduce fatalities and injuries. The point estimates indicate that the false-alarm effect is almost equal to the warning effect at the observed trade-off. We find that a substantial false-alarm effect does exist, and based on this we cannot recommend maximizing the POD while ignoring the impact on false alarms.

*Acknowledgments.* We thank Joseph Schaefer for supplying us with the Storm Prediction Center tornado archive and William Lerner for sharing NOAA's tornado warning verification archive with us. This research was supported in part by the Department of Commerce/NOAA Grant 74-6002942. We thank Walker Ashley, Jamie Kruse, seminar attendees at Northern Illinois University, Lisa Dilling, and three referees for helpful comments on an earlier draft. Xiaoyi Mu, Ying Xiao,

[9] This statement applies at the current trade-off on the margin; shifting the frontier for feasible warnings to allow an increase in the POD *and* reduction in FAR would reduce casualties. If the optimal warning lead time is assumed (the interval with the largest reduction in casualties relative to no warning) instead of the average effect across lead 1–5, lead 6–10, and lead 11–15, the decrease in expected fatalities and injuries of increasing the POD from 0.4 to 1.0 is −27.1% and −24.0%.

TABLE 10. Evaluating the trade-off between FAR and POD. The numbers are percentage changes for a trade-off of (POD = 0.4, FAR = 0.7) for (POD = 1.0, FAR = 0.9). The main numbers are calculated for the point estimates of the FAR variable; the last two columns use the lower and upper bounds (LB and UB, respectively) of the 95% CI for the FAR impact on casualties.

| | | | Combined effects | | |
|---|---|---|---|---|---|
| | Impact of POD | Impact of FAR | Point estimates | LB 95% CI | UB 95% CI |
| Fatalities | | | | | |
| 1-yr state | −17.9 | +21.4 | −0.3 | −12.3 | +13.5 |
| 1-yr WFO | −17.3 | +23.2 | +1.9 | −7.3 | +11.9 |
| 1-yr TV | −20.6 | +25.4 | −0.4 | −8.4 | +8.3 |
| 2-yr state | −17.3 | +65.6 | +37.1 | +15.9 | +62.3 |
| 2-yr WFO | −16.6 | +43.9 | +20.0 | +5.4 | +36.7 |
| 2-yr TV | −20.1 | +24.9 | −0.3 | −8.5 | +8.7 |
| Injuries | | | | | |
| 1-yr state | −21.1 | +24.9 | −1.5 | −11.2 | +9.4 |
| 1-yr WFO | −20.0 | +37.7 | +10.2 | +1.3 | +19.8 |
| 1-yr TV | −20.5 | +19.1 | −5.4 | −11.3 | +0.9 |
| 2-yr state | −21.0 | +43.0 | +13.0 | −0.9 | +28.8 |
| 2-yr WFO | −21.1 | +49.5 | +18.0 | +6.6 | +30.6 |
| 2-yr TV | −20.4 | +19.2 | −5.2 | −11.1 | +1.2 |

Xiujian Chen, Abby Schroeder, Adil Suliman, and Radhames Lizardo provided research assistance.

## REFERENCES

Anbarci, N., M. Escaleras, and C. A. Register, 2005: Earthquake fatalities: The interaction of nature and political economy. *J. Public Econ.,* **89,** 1907–1933.

Ashley, W. S., 2007: Spatial and temporal analysis of tornado fatalities in the United States, 1880–2005. *Wea. Forecasting,* **22,** 1214–1228.

——, A. J. Krmenec, and R. Schwantes, 2008: Nocturnal tornadoes. *Wea. Forecasting,* **23,** 795–807.

Atwood, L. E., and A. M. Major, 1998: Exploring the ''cry wolf'' hypothesis. *Int. J. Mass Emerg. Disasters,* **16,** 279–302.

Barnes, L. R., E. C. Gruntfest, M. H. Hayden, D. M. Schultz, and C. Benight, 2007: False alarms and close calls: A conceptual model of warning accuracy. *Wea. Forecasting,* **22,** 1140–1147.

Breznitz, S., 1984: *Cry Wolf: The Psychology of False Alarms*. Lawrence Earlbaum Associates, 265 pp.

Brooks, H. E., 2004: Tornado-warning performance in the past and future: A perspective from signal detection theory. *Bull. Amer. Meteor. Soc.,* **85,** 837–843.

——, and C. A. Doswell III, 2002: Deaths in the 3 May 1999 Oklahoma City tornado from a historical perspective. *Wea. Forecasting,* **17,** 354–361.

Doswell, C. A., III, and D. W. Burgess, 1988: On some issues of United States climatology. *Mon. Wea. Rev.,* **116,** 495–501.

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

Golden, J. H., and C. R. Adams, 2000: The tornado problem: Forecast, warning, and response. *Nat. Hazards Rev.,* **1,** 107–118.

Greene, W. H., 2000: *Econometric Analysis*. 4th ed. Prentice Hall, 1004 pp.

Kahn, M. E., 2005: The death toll from natural disasters: The role of income, geography, and institutions. *Rev. Econ. Stat.,* **87,** 271–284.

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

Schmidlin, T. W., B. Hammer, and J. Knabe, 2001: Tornado shelters in mobile home parks in the United States. *J. Amer. Soc. Prof. Emerg. Planners,* **8,** 1–15.

Simmons, K. M., and D. Sutter, 2005: WSR-88D radar, tornado warnings, and tornado casualties. *Wea. Forecasting,* **20,** 301–310.

——, and ——, 2008: Tornado warnings, lead times and tornado casualties: An empirical investigation. *Wea. Forecasting,* **23,** 246–258.

Trapp, R. J., S. A. Tessendorf, E. S. Godfrey, and H. E. Brooks, 2005: Tornadoes from squall lines and bow echoes. Part I: Climatological distribution. *Wea. Forecasting,* **20,** 23–34.

Viscusi, W. K., J. M. Vernon, and J. E. Harrington Jr., 2000: *Economics of Regulation and Antitrust*. 3rd ed. MIT Press, 864 pp.