1 **A model for U.S. tornado casualties involving interaction between damage**
2 **path estimates of population density and energy dissipation**

3 James B. Elsner*, Tyler Fricker, and William D. Berry

4 *Florida State University, Tallahassee, Florida*

5 **Corresponding author address:* James B. Elsner, Department of Geography, Florida State Univer-
6 sity, 113 Collegiate Loop, Tallahassee, FL 32306.

7 E-mail: jelsner@fsu.edu

ABSTRACT

8 A recent study showed the importance of tornado energy as a factor in a
9 model for tornado deaths and injuries (casualties). The model was additive
10 under the assumption of uniform threat. Here we test two explicit hypothe-
11 ses designed to examine this additive assumption. The first hypothesis con-
12 cerns energy dissipation's effect conditional on population density and the
13 second concerns population's effect conditional on energy. Both hypotheses
14 are tested using a regression model that contains the product of population
15 density and energy dissipation. Results show that the elasticity of casualties
16 with respect to energy dissipation increases with population density. That is,
17 the percentage increase in casualties with increasing energy dissipation *in-*
18 *creases* with population density. Similarly, the elasticity of casualties with
19 respect to population density increases with energy dissipation. That is, the
20 percentage increase in casualties with increasing population density *increases*
21 with energy dissipation. Allowing energy and population elasticities to be
22 conditional rather than constant provides a more complete description of how
23 tornado casualties are influenced by these two important factors.

24 **1. Introduction**

25 Nearly a fifth of all natural-hazards fatalities in the United States are the direct result of tornadoes
26 (National Oceanic and Atmospheric Administration, 2015). A tornado that hits a city is capable
27 of inflicting hundreds or thousands of casualties (death or injury). Data from the Storm Prediction
28 Center (SPC) of the National Oceanic and Atmospheric Administration (NOAA) show that the
29 27 April 2011 Sawyerville-Eoline, Alabama tornado produced 1564 casualties with 64 of them
30 resulting in death. Less than a month later the 22 May 2011 Joplin, MO tornado produced 1308
31 casualties with 158 of them resulting in death. More recently, the 26 December 2015 Garland-
32 Rowlett, Texas tornado resulted in ten deaths and 468 injuries.

33 Casualties occur when a tornado strikes people at home, work, school, etc. Extreme wind speeds,
34 relatively short warning lead times, and low quality built environments increase the potential for
35 casualties (Greenough et al., 2001). Risk factors for a casualty include taking shelter in a manufac-
36 tured home, being an older adult, and inaccessibility to a safe room in a basement or in a reinforced
37 structure (CDC, 2012). Early warnings and strong shelters are the most effective way to reduce
38 deaths and injuries (Bohonos and Hogan, 1999). But the potential for casualties is not uniform
39 across the country (Ashley and Strader, 2016). Southern states are particularly at risk because of
40 the abundance of manufactured homes and the frequency of reduced visibility conditions (trees,
41 hills, and heavy rain).

42 Our goal is to better understand how factors related to tornado exposure and ferocity combine
43 to determine the number of casualties. The objective is a statistical model that describes the rate
44 of casualties given known tornado path characteristics and the estimated number of people within
45 the path. Our exclusive focus is on casualty-producing tornadoes. Given that a tornado produces

46 a casualty, how do population density and energy dissipation combine to explain the number of
47 casualties?

48 **2. Hypotheses**

49 Fundamentally, it is clear that a tornado's *casualty rate* (how many people get injured or killed)
50 depends on the *density of population* in the tornado's path, and the tornado's *strength* (i.e., the
51 power of the winds it generates, or using more technical language, the amount of atmospheric
52 energy it dissipates).¹ Fricker et al. (2017) estimate the elasticity of tornado casualties with respect
53 to each of population density and energy dissipation, and use these estimates to calculate that
54 when the population density within the path of a tornado is held constant, a doubling of its energy
55 increases casualties, on average, by 33% ($\pm 3\%$); and when the energy dissipation of a tornado is
56 held constant, a doubling of population density along its path increases casualties by an average
57 of 21% ($\pm 3\%$).

58 However, the statistical analysis yielding these findings is based on a model that assumes the
59 effect of a tornado's strength on its casualty rate is the same regardless of the population density
60 along its path. Yet, it seems more reasonable to expect that the effect of tornado strength on the
61 casualty rate is conditional on population density. When the population density along a tornado's
62 path is zero, by definition, the casualty rate must be zero—since regardless of the strength of the
63 tornado, there are no people present to be injured or killed. This guarantees that tornado energy has
64 no effect on the casualty rate when population density is zero. In contrast, when population density
65 along a tornado's path is very high, tornado energy should have a strong effect on the casualty
66 rate—with low-energy storms uniformly producing a small number of casualties and high-energy
67 storms often resulting in a large number of casualties. Indeed, as the population density in the path

¹ We treat as synonymous the terms tornado *strength*, tornado *energy*, and energy *dissipation*.

68 of a tornado rises from zero, the effect of tornado energy on casualties should gradually grow in
69 magnitude. We can formalize this reasoning about the way in which the effect of tornado strength
70 on the casualty rate is conditional on population density in the following testable hypothesis:

71 **Energy's Effect Conditional on Population Hypothesis:** (i) As long as the population den-
72 sity along a tornado's path is greater than zero, the effect of the tornado's energy on its casu-
73 alty rate is positive regardless of the density of population in its path. However, (ii) the effect
74 of tornado energy on the casualty rate increases in magnitude as population density rises.

75 If the effect of tornado strength on the casualty rate is conditional on population density, the
76 symmetry of conditionality guarantees that the effect of population density on the casualty rate is
77 conditional on tornado strength (Berry et al., 2012). However, the guaranteed conditionality does
78 not establish the sign of the effect of population density at any level of tornado energy. Additional
79 reasoning is required. Since even the weakest of tornadoes is strong enough to kill and injure,
80 at any tornado energy level, an increase in the population density along the tornado's path puts
81 more people in harm's way, and should result in an increase in the casualty rate. As a tornado's
82 energy rises, the potential of the storm to inflict casualties on the population in its path should
83 increase, and so the magnitude of the positive effect of population density on the casualty rate
84 should increase.

85 The expected positive effect of population density on the casualty rate is not due solely to the
86 increase in the number of people in harm's way. Low population density is most likely to be found
87 in rural areas, and high density is most likely to be present in urban areas; and for several reasons,
88 an urban area tends to be more prone to tornado casualties than a rural area that is impacted
89 by a tornado of identical strength. Donner (2007) argues that the types of relationships rural
90 communities create, as well as the awareness people in such regions hold of their habitats, might

91 help to protect otherwise vulnerable populations. He maintains that strong social bonds expand
 92 the scope and availability of potential sources of weather information, access to shelters, and
 93 other resources that help with protective action. He also speculates that urban areas are more
 94 vulnerable to casualties because the rare occurrence of tornadoes in urban communities leads
 95 to lack of preparedness among residents, along with skepticism about the risks. Moreover, the
 96 presence of tall buildings and other objects may make it difficult to visually confirm the presence
 97 of tornadoes in urban communities. Without an environmental cue, warning responses tend to be
 98 lower (Mileti and O'Brien, 1992).

99 We can formalize our reasoning about the way in which the effect of population density on the
 100 casualty rate is conditional on a tornado's strength in the following testable hypothesis:

101 **Population's Effect Conditional on Energy Hypothesis:** (i) The effect of population density
 102 along the path of a tornado on the tornado's casualty rate is positive regardless of the tornado's
 103 energy. However, (ii) the effect of population density increases in magnitude as tornado
 104 energy rises.

105 3. Model

106 We test our two hypotheses by modifying a model introduced by Fricker et al. (2017). They
 107 estimate a negative binomial regression with the tornado as the unit of analysis and the event
 108 count, the number of casualties resulting from the tornado (denoted C), as dependent variable:

$$C \sim \text{NegBin}(\hat{\mu}, n) \tag{1a}$$

$$\ln(\mu) = \ln(\beta_0) + \beta_P \ln(P) + \beta_E \ln(E) \tag{1b}$$

109 The casualty rate, C , is assumed to be adequately described by a negative binomial distribution
 110 with a rate parameter μ and a size parameter n . P denotes population density, and E denotes
 111 energy dissipation. The coefficient for $\ln(P)$, β_P , represents the elasticity of the tornado casualty
 112 rate with respect to population density (to be abbreviated *population elasticity*), and represents the
 113 percentage change in the casualty rate associated with a one percent increase in population density.
 114 Similarly, β_E represents the elasticity of the tornado casualty rate with respect to energy dissipa-
 115 tion (to be abbreviated *energy elasticity*), and reflects the percentage change in the casualty rate
 116 associated with a one percent increase in tornado energy. The functional form of Eq. 1b imposes
 117 the assumption that energy elasticity is the same regardless of the value of population density, and
 118 population elasticity is the same regardless of the value of energy dissipation. These assumptions
 119 are directly contrary to our two conditional hypotheses.

120 However, we can adapt Fricker et al. (2017)'s model to be compatible with our conditional
 121 hypotheses by adding a multiplicative term, $\ln(P) \cdot \ln(E)$, to Eq. 1b to yield the model:

$$C \sim \text{NegBin}(\hat{\mu}, n) \quad (2a)$$

$$\ln(\mu) = \ln(\beta_0) + \beta_P \ln(P) + \beta_E \ln(E) + \beta_{P \cdot E} [\ln(P) \cdot \ln(E)] \quad (2b)$$

122 To reveal how this revised model makes energy elasticity conditional on population density, we
 123 fix P in Eq. 2b at some constant, P^* :

$$\ln(\mu) = \ln(\beta_0) + \beta_P \ln(P^*) + \beta_E \ln(E) + \beta_{P \cdot E} [\ln(P^*) \cdot \ln(E)] \quad (3)$$

124 Rearranging terms yields

$$\ln(\mu) = [\ln(\beta_0) + \beta_P \ln(P^*)] + [\beta_E + \beta_{P \cdot E} \ln(P^*)] \cdot \ln(E) \quad (4)$$

125 In this equation, each of $[\ln(\beta_0) + \beta_P \ln(P^*)]$ and $[\beta_E + \beta_{P \cdot E} \ln(P^*)]$ is a constant, so that $\ln(\mu)$ is
 126 expressed as a linear function of $\ln(E)$. Accordingly, the coefficient for $\ln(E)$, $[\beta_E + \beta_{P \cdot E} \ln(P^*)]$,

127 is the elasticity of the tornado casualty rate with respect to energy dissipation when population
128 density, P , equals P^* . Thus, it is clear that Eq. 2b specifies energy elasticity as conditional on the
129 value of population density; more specifically, energy elasticity is a linear function of the logarithm
130 of population density, $\ln(P)$. Similarly, we can fix E in Eq. 2b at the constant, E^* , and manipulate
131 terms to show that $[\beta_P + \beta_{P \cdot E} \ln(E^*)]$ is the elasticity of the tornado casualty rate with respect to
132 population density when energy dissipation, E , equals E^* —implying that population elasticity is
133 conditional on the value of energy dissipation.

134 Because Eq. 2b models the elasticity of the tornado casualty rate with respect to each of energy
135 dissipation and population density as conditional on the value of the other, we can describe the
136 equation as assuming that energy dissipation and population density *interact* in influencing tornado
137 casualties. This is in contrast to Eq. 1b, which assumes that energy dissipation and population
138 density are *additive* in their influence on tornado casualties. The coefficients β_P , β_E and $\beta_{P \cdot E}$
139 from Eq. 2b can be used to calculate the elasticity of the tornado casualty rate with respect to
140 each of energy dissipation and population density at any specified value of the other variable,
141 thereby permitting an empirical test of each of our two hypotheses. Moreover, the coefficient
142 $\beta_{P \cdot E}$ in Eq. 2b measures the magnitude of the interaction between population density and energy
143 dissipation, i.e., the magnitude of the relationship between (i) population elasticity and the log of
144 energy dissipation, and (ii) energy elasticity and the log of population density.

145 Our overall modeling approach is similar to that of recent work that examines factors related
146 to tornado casualties (Lim et al., 2017; Zahran et al., 2013; Simmons and Sutter, 2008; Donner,
147 2007). However it differs in a couple of key ways. First, we use a continuous measure of a
148 tornado's energy dissipation—rather than path area, EF rating, or total damage—as the indicator
149 of tornado strength. Correlation among these variables is high but energy dissipation has a better
150 physical justification as it is wind power that is proximal to the cause of casualties. Second, we

151 exclude tornadoes without casualties, focusing on factors influencing the casualty rate among those
152 tornadoes producing at least one casualty. This makes our models simpler, and the interpretation
153 of the results more straightforward, since for a variety of reasons most tornadoes do not result in
154 casualties.² Third, our study focuses on the interaction between the environment (tornado energy)
155 and a demographic factor (population density) in explaining tornado casualties. A focus on the
156 statistical *interaction* between environmental and demographic factors is unique in research about
157 tornado casualties, and provides a new lens through which to view factors that influence the rate
158 of injuries and death from tornadoes.

159 In brief, the results from this study offer support for both our conditional hypotheses. We find
160 that the elasticity of the tornado casualty rate with respect to energy dissipation is positive and
161 statistically significant across most of the range of population density values, and that this elasticity
162 increases with population density. We also find that the elasticity of the tornado casualty rate with
163 respect to population density is positive and statistically significant across most of the range of
164 energy dissipation values, and that this elasticity increases with tornado energy. Supplementary
165 analysis shows that if our results are biased due to error in measuring population density, this
166 bias may result in an *underestimation* of the extent of interaction between energy dissipation and
167 population density. The paper continues in Section 4 with a description of the indicators we use
168 to measure tornado casualties, tornado energy and population density. Results are presented in
169 Section 5. A summary of what was done and prospects for future work are given in Section 6.

²The question of *whether* a tornado produces casualties is interesting but it has been addressed elsewhere (e.g., Ashley et al. (2008)).

4. Data and Variables

a. Tornado casualty rates

Interest centers on the count of injuries and fatalities listed in the Storm Prediction Center's tornado database. The database is compiled from the National Weather Service's (NWS) *Storm Data*, and includes all known tornadoes dating back to 1950. A tornado casualty is defined as a fatality or injury directly attributable to a tornado, or to impact by airborne, falling, or moving debris. An example of a direct fatality would be a driver killed when a motor vehicle is tossed over. To be considered a direct injury, the injury must require treatment by a first-responder or at a medical facility. An example would be a person treated for a laceration caused by flying debris. Indirect casualties are not considered. An example of an indirect casualty would be electrocution during debris removal.

NWS *Storm Data* also include for each tornado, the straight-line track, start location, damage path dimensions of length and maximum width, and the maximum Enhanced Fujita (EF) rating (wind speed damage rating on an ordered scale of categories from zero to five). Reports in the database are compiled by the NWS offices and reviewed by the National Climate Data Center (Verbout et al., 2006). We consider only tornadoes occurring within the conterminous United States.

During the 22-year period 1995–2016 there are 26,863 tornadoes recorded in the conterminous United States. Of these, 2,208 are linked to 25,968 casualties. Only 6.7 percent of all casualties lead to death. Most casualty-producing tornadoes result in just a few casualties. Indeed, even among casualty-producing tornadoes, the median number of casualties is 3.

A relatively small number of casualty-producing tornadoes result in many casualties. The Tuscaloosa-Birmingham, Alabama tornado of 27 April 2011 and the Joplin, Missouri tornado

193 of 22 May 2011 top the list of the most casualties in the period. The Tuscaloosa-Birmingham
194 tornado was one of 362 tornadoes occurring between 25 and 28 April in the largest outbreak in
195 recorded United States history. The Joplin tornado is the deadliest in the modern recorded-keeping
196 era (since 1950).

197 During the period 1995–2016, tornado casualties were most common across the southeastern
198 quarter of the country (Fig. 1). Alabama had the most (3,937) followed by Oklahoma (2,500),
199 Missouri (2,025), Arkansas (1,759), Texas (1,704), and Tennessee (1,678). Rounding out the top
200 ten are Georgia (1,610), Mississippi (1,400), Kentucky (1,051), and North Carolina (985). Outside
201 the South, Massachusetts (ranked 22nd) stands out in New England with 232 casualties from only
202 three tornadoes.

203 The number of casualties is positively related with a tornado’s EF (wind speed) rating (Table 2).
204 Among casualty-producing tornadoes, on average an EF0 tornado results in two casualties, an
205 EF1 tornado results in 3.1 casualties, and an EF2 tornado results in 6.5 casualties. Beyond EF2,
206 expected casualties increase dramatically. On average, an EF3 tornado results in 20 casualties, an
207 EF4 tornado results in 62 casualties, and an EF5 tornado results in 236 casualties.

208 *b. Population density*

209 We create a damage path from the tornado track using the width variable and then estimate the
210 total number of people within the path boundary. Population data are obtained from the Gridded
211 Population of the World, version four (GPW, v4) from the Socioeconomic Data and Applications
212 Center at Columbia University. The database contain decennial census density estimates for 1990,
213 2000, and 2010 represented as residential population/km². The native cell resolution is .0083°
214 latitude/longitude, which at 36° N latitude means a cell having the dimension of .9 km in the north-
215 south direction and .7 km in the east-west direction. For each tornado we assign the population

216 density corresponding to the closest decennial estimate so, for example, a tornado that occurred
217 in 1996 uses the 2000 decennial estimate while a tornado that occurred in 2012 uses the 2010
218 estimate.³

219 For the set of 2,192 tornadoes with at least one casualty⁴ the median population density per
220 tornado is 31.9 people/km² with an inter-quartile range between 9.88 and 137 people/km². There
221 are seven orders of magnitude separating the lowest from highest per-tornado population density
222 (Table 1). We estimate the EF2 Brooklyn, NY tornado of 8 August 2007 that injured nine had
223 the highest population density in its path at 13,949 people/km². Of the ten casualty-producing
224 tornadoes with the highest population densities, three occurred in New York, two occurred in
225 Florida, and two occurred in California. Pennsylvania, Massachusetts, and Texas each had one.

226 The distribution of population and population density in the path of tornadoes by EF rating
227 is shown in Table 3. We estimate that about 2.5 million people have been within the path of a
228 tornado and 1.5 million people within the path of a casualty-producing tornado. Because of their
229 frequency and path area, EF1 tornadoes account for the largest accumulated number of people at
230 risk. But given a casualty-producing tornado, it is the EF3 tornadoes that, as a group, affected
231 the most people. The number of people affected is a strong function of EF rating since more
232 damaging tornadoes tend to have larger damage paths (Brooks, 2004; Elsner et al., 2014; Ashley
233 and Strader, 2016). There is no strong relationship between median population density and EF
234 rating. There are thirty non-casualty-producing tornadoes without an EF rating with an estimated
235 total population within the paths of thirteen people.

³Linear interpolation of the population to the tornado year had no significant influence on the model results.

⁴Sixteen (16) of the 2,208 casualty-producing tornadoes occurred over areas without population and are removed from further analysis.

236 *c. Energy dissipation as a measure of tornado strength*

237 Tornadoes dissipate a large amount of atmospheric energy (Fricker et al., 2014, 2017). Following
238 what was done with hurricanes (Emanuel, 2005), the energy dissipation of a tornado is computed
239 as

$$E = A_p \rho \sum_{j=0}^5 w_j v_j^3, \quad (5)$$

240 where the summation is over the six possible EF ratings (0, 1, 2, 3, 4 and 5), A_p is the area of
241 the tornado's path [units of square meters], ρ is air density [1 kg m^{-3}], v_j is the midpoint wind
242 speed [m s^{-1}] for each damage rating (EF scale) j , w_j is the corresponding fraction of path area
243 by damage rating, and J is the maximum damage rating. Multiplying the units from the individual
244 terms results in E being measured in a unit of power [$\text{kg m}^2 \text{ s}^{-3} = \text{Joule/s} = \text{Watt (W)}$].

245 The fraction of path area is that recommended by the U.S. Nuclear Regulatory Commission [see
246 Fricker and Elsner (2015)], which combines a Rankine vortex with empirical estimates (Ramsdell
247 and Rishel, 2007). Threshold wind speeds for the EF ratings are a three second gust. With no
248 upper bound on the EF5 wind speeds, the midpoint wind speed is set at 97 m s^{-1} (7.5 m s^{-1}
249 above the threshold wind speed consistent with the EF4 midpoint speed relative to its threshold).
250 Additional details and justification for energy dissipation as a valid measure of tornado strength
251 are given in Fricker et al. (2017).

252 For the set of 2,192 tornadoes with at least one casualty, the median energy dissipation is 97.2
253 gigawatts (GW) with an inter-quartile range between 14 and 514 GW. The Tallulah-Yazoo City-
254 Durant tornado (Louisiana and Mississippi) of 24 April 2010 that killed ten and injured 146 has
255 the highest energy dissipation at 66,200 GW. There are seven orders of magnitude separating the
256 lowest from the highest per-tornado energy dissipation.⁵

⁵The distributions of population density and energy dissipation are approximately normal on a logarithmic scale.

5. Results

a. Visual inspection of the data

Fig. 3 plots the tornado casualty rate for each tornado path in the sample against both the storm's energy dissipation and the population density in its path. The correlation (Pearson) between energy dissipation and population density is $-.07$. Given a tornado with at least one casualty, the number of casualties increases with the estimated number of people in the path and with energy dissipation. It is rare for a tornado with less than .1 GW of energy dissipation to produce more than a dozen casualties when the population density is less than 100 people/km². Casualty-producing tornadoes with an order of magnitude more energy dissipation tend to produce many more casualties especially when affecting areas with population densities exceeding 10 people/km².

Energy dissipation depends strongly on the EF damage rating. Tornadoes that hit more targets in developed areas (built environments) have the potential to be rated higher all else being equal because the rating comes from the worst damage. In built environments with more targets there are more people so energy dissipation might contain some population effect. However, the correlation between energy dissipation and population density is small, indicating this is not a problem.

b. Model results

We estimate the interactive negative binomial regression model of Eqs. 2a and 2b using our data set of 2,192 tornadoes.⁶ The resulting coefficient estimates are presented in the top portion of Table 4. To understand the empirical results, we use these coefficients to estimate the (i) energy elasticity at each possible value of population density, and (ii) population elasticity at each possi-

⁶All empirical analysis is performed using the open-source R language for statistical computing (R Core Team, 2016) with freely-available data.

All figures were made using functions from the `ggplot2` package (Wickham, 2009). The R code and links to the data for this study are available at <https://rpubs.com/jelsner/interaction>.

277 ble value of energy dissipation. However, the model’s estimated elasticities are more trustworthy
278 for values of energy dissipation and population density near which there are abundant data, since
279 estimates “far from the data” are less influenced by data, and more influenced by the assumed
280 functional form of the model (King and Zeng, 2006). Thus, in all figures displaying model es-
281 timates below, we restrict attention to values for population density and energy dissipation over
282 which the density of observations in the sample of tornadoes is relatively large.

283 In particular, we focus our attention on estimates over the region internal to the gray polygon su-
284 perimposed over the joint distribution for population density and energy dissipation in Fig. 3. This
285 region was identified by visual inspection of the plot as excluding pairs of values for population
286 density and energy dissipation near which there are very few observations. The polygon—which
287 contains 83.4% of the tornadoes in our sample—excludes all values of population density below
288 1.4 people/km², and above 1500 (a range containing of 92.3% the tornadoes in the sample), and
289 all values of energy dissipation below 1 GW and above 10,000 (a range containing 92.7% of torna-
290 does). Effectively our choice is the minimum area encompassing at least 83% of the data values.
291 Estimates are the same if we do not omit any data but uncertainty is higher for estimates made
292 outside this area.

293 Estimated elasticities are used to construct Fig. 4.⁷ The black lines in Panels A and B show,
294 respectively, estimated energy elasticity (along with a 95% confidence band) conditional on pop-
295 ulation density, and estimated population elasticity conditional on energy dissipation. Panel A
296 shows, for example, that when population density equals 100 people/km², energy elasticity is
297 estimated to equal .357 (with a confidence interval extending from .315 to .400)—a value that

⁷Although Eq. 2b is framed using natural logarithms, to facilitate clarity the plots in Fig. 4 and elsewhere have axes scaled using (more widely familiar) common logarithms.

298 indicates that a one percent increase in energy dissipation is associated, on average, with a .357
299 percent increase in the number of casualties.

300 The Energy's Effect Conditional on Population Hypothesis, predicts that (i) the effect of energy
301 dissipation on the casualty rate is positive at any population density exceeding zero, but (ii) the
302 magnitude of this effect grows as population density rises. Regarding prediction (i), Panel A shows
303 that energy elasticity is positive and statistically significant at each population density value over
304 which elasticities are displayed. If we extend the focus to the entire range of values for population
305 density in the sample (i.e., .003 to 13,949), estimated energy elasticity is positive and statistically
306 significant at all population densities greater than .21 people/km²; this is a range of values in which
307 99% of the tornadoes in the sample lie. In contrast, there are no values for population density at
308 which estimated energy elasticity is negative and statistically significant.

309 Given the functional form of Eq. 2b, prediction (ii) of the Energy's Effect Conditional on Pop-
310 ulation Hypothesis requires that the coefficient for the multiplicative term, $\beta_{P.E}$, be positive, and
311 indeed the estimate of this coefficient is positive and statistically significant (with a p -value less
312 than .0001), implying that energy elasticity increases as population density rises. More specif-
313 ically, the point estimate for $\beta_{P.E}$ of .0415 indicates that the elasticity of the casualty rate with
314 respect to energy dissipation increases by .0288 with each doubling of population density.⁸ Panel
315 A helps clarify the extent to which energy elasticity varies with the value of population density.
316 The plot shows that a shift in population density from 1.4 people/km² (roughly the density in
317 Plainview, AR) to 1500 people/km² (roughly the density in Eastlake, OH) results in an increase
318 in estimated energy elasticity from .179 (.110, .247) to a value more than two and half times in
319 magnitude: .470 (.393, .549).

⁸To get the value .0288, one multiplies the coefficient $\beta_{P.E}$ by $\ln(2)$: (.0415)(.6931) = .0288.

Panel B allows an evaluation of the Population's Effect Conditional on Energy Hypothesis—which predicts that (i) the impact of population density on the casualty rate is positive at any value of tornado energy, but (ii) the strength of the impact rises as energy dissipation increases. Regarding prediction (i), Panel B reveals that population elasticity is positive and statistically significant at each energy dissipation value greater than 2.1 GW⁹; this is a range of tornado energy values in which 90.6% of the tornadoes in the sample lie. In contrast, there are no energy dissipation values at which estimated population elasticity is negative and statistically significant.

The statistically significant positive coefficient estimate for the multiplicative term, $\beta_{P.E}$, is evidence that population elasticity increases as energy dissipation rises—consistent with prediction (ii) of the Population's Effect Conditional on Energy Hypothesis. Furthermore, Panel B clarifies the implications of the coefficient for the multiplicative term about the degree to which population elasticity varies with the level of tornado energy. Increasing energy dissipation in the sample from 1 GW to 10,000 GW prompts an increase in estimated energy elasticity from .047 (−.041, .133) to a value nearly ten times its size: .432 (.330, .537).

To better understand the extent of interaction between energy dissipation and population density in influencing the tornado casualty rate, we can use the coefficients for the model of Eqs. 2a and 2b to compute the estimated tornado casualty rate conditional on a tornado's values for energy dissipation and population density. Figure 5 displays the estimated relationship between energy dissipation and the casualty rate at three different values for population density: the minimum value contained within the polygon in Fig. 3 (1.4 people/km²), the maximum value within the polygon (1500 people/km²), and the median value across tornadoes in the sample (31.9 people/km²). As

⁹For comparison a 40 m/s hurricane dissipates about 12.5 GW of kinetic energy over the average lifetime of a tornado (12 min) [see (Emanuel, 1999)]

energy dissipation increases from 1 to 10,000 GW, the casualty rate is expected to increase by a difference of

- 6.55 (from 1.54 to 8.08) when population density is at its *minimum*,
- 29.2 (from 1.79 to 31.0) when population density is at its *median*, and
- 160.5 (from 2.15 to 162.6) when population density is at its *maximum*.

Thus, Figs. 4 and 5 provide strong empirical evidence for both our hypotheses: the Energy's Effect Conditional on Population Hypothesis, and the Population's Effect Conditional on Energy Hypothesis. Specifically, the effect on a tornado's casualty rate of each of energy dissipation and population density is positive at nearly any value of the other variable, but with a magnitude that grows with the value of the other variable.

c. The consequences of shifting from an additive model to an interactive model

The conditionality in the effects of energy dissipation and population density would be undetectable if we were to analyze the data using the additive model, since the additive model assumes that the elasticity of tornado casualties with respect to each of energy dissipation and population density is the same regardless of the value of the other variable. Yet, it is interesting to assess more systematically the implications for our understanding of the consequences of tornadoes of abandoning the additive model (Eqs. 1a and 1b) in favor of an interactive model (Eqs. 2a and 2b).

Using the same data we employ to estimate the interactive model, we estimate the additive model (i.e., the model created by deleting the multiplicative term in Eq. 2b); the resulting coefficient estimates are presented in the bottom half of Table 4. The statistical significance of the coefficient for the multiplicative term in the interactive model constitutes initial evidence that the interactive model provides a better fit to the data than the additive model. However, it may be that

the improvement in fit by specifying that population density and energy dissipation interact is too small to justify abandoning the simpler additive model in favor of the more complex interactive model. Thus, we conduct two tests suitable for comparing additive and interactive models in a way that balances model fit and model complexity by imposing a penalty for greater complexity: Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Both criteria indicate the superiority of the interactive model—with AIC decreasing from 13323 to 13223, and BIC decreasing from 13340 to 13246, when the multiplicative term is added to the model.

We can analyze the impact of shifting from an additive model to an interactive model on the magnitude of estimated elasticities. The horizontal red lines in Fig. 4 plot the estimated value of energy elasticity (in Panel A) and population elasticity (in Panel B) obtained from the additive model. The fact that each red line is horizontal is a consequence of the additive model's assumption that the elasticity of the casualty rate with respect to each of energy dissipation and population density is constant. We saw above (from Panel A) that the interactive model yields the estimate that as population density increases from 1.4 to 1500 people/km², energy elasticity increases from .179 to .470. This contrasts with the additive model's estimate of a constant energy elasticity of .321. The interactive model also estimates that as energy dissipation increases from 1 to 10,000 GW, population elasticity increases from .047 to .432, in contrast with the additive model's estimate of a constant population elasticity of .213.

For a final analysis of the implications of shifting from an additive to an interactive model, we use the coefficients for each model to compute the predicted casualty rate for each tornado path in the estimation data set, given the tornado's values for energy dissipation and population density. Then, for each tornado path, we compute the absolute difference between the two models' predicted casualty rates. The median absolute difference is .480, and the mean is 1.37. These values may, at first glance, seem small. However, the typical tornado produces relatively few casualties;

387 even if we exclude the majority of tornadoes that cause no casualties, the median casualty rate
388 across all (casualty-producing) tornadoes is 3. Given the median of 3, we view an absolute differ-
389 ence of .480 as consequential. For 11% of tornadoes, the absolute error exceeds 3, and for 5.5%,
390 the absolute difference is greater than 5. Moreover, a nontrivial share of tornadoes are character-
391 ized by an absolute difference that is very high; for 1.6% of tornadoes (36), the absolute difference
392 is greater than 10. Clearly, modifying the statistical analysis of tornadoes to allow estimated en-
393 ergy and population elasticities to be conditional rather than constant yields a much more nuanced,
394 and richer, characterization of the effects of a tornado’s strength and of the population in its path
395 on the resulting number of casualties.

396 *d. A potential bias from the population data*

397 Ideally, our measure of population density should reflect the number of people in the path of
398 the tornado *when the tornado strikes*. However, because we do not have access to this informa-
399 tion, we rely on the number of people that *reside* within a tornado’s path—ignoring the fact that
400 when the tornado strikes, some people residing in the path will be away from home, and others
401 who live outside the path will be within the path. It seems reasonable to assume that residential
402 population more closely approximates the number of persons in a tornado’s path when the tornado
403 strikes at night (and the vast majority of people are at home) than when it strikes during the day.
404 If so, estimating the interactive model on the subset of nighttime tornadoes could be conceived
405 as estimating the model using those tornadoes in our sample for which population density in the
406 path is measured with the least error. The top half of Table 4 shows that when the model is esti-
407 mated using all tornadoes, the coefficient for the multiplicative term—which reflects the extent of
408 interaction between energy dissipation and population density in influencing the tornado casualty
409 rate—is .042. In contrast, when the model is estimated using *just* nighttime tornadoes (i.e., those

410 between sunset and sunrise), the coefficient for the multiplicative term increases in magnitude to
411 .049. This supplementary analysis suggests that if our main results are biased due to measurement
412 error, this bias results in an *underestimation* of the extent of interaction between population density
413 and energy dissipation on casualties.

414 6. Summary

415 A recent study showed the importance of energy dissipation on tornado casualties using an
416 additive regression model (Fricker et al., 2017). In this paper we test two explicit hypotheses
417 designed to examine the additive assumption implicit in this earlier study. The first hypothesis
418 concerns energy's effect conditional on population and the second concerns population's effect
419 conditional on energy. The hypotheses are tested with a regression model that contains a term as
420 the product of population density and energy dissipation.

421 New results here show that energy elasticity increases with population density. That is, the per-
422 centage increase in casualties with increasing energy dissipation *increases* with population density.
423 Similarly, population elasticity increases with energy dissipation. That is, the percentage increase
424 in casualties with increasing population density *increases* with energy dissipation. The conclu-
425 sion is unambiguous in that allowing energy and population elasticities to be conditional rather
426 than constant provides a more complete description of how casualties are influenced by these two
427 factors as indicated by more accurate predictions of casualty rates.

428 This study focused on the two dominant factors that determine the number of casualties resulting
429 from a tornado: tornado energy and population density. Undoubtedly, casualties are influenced by
430 other variables as well. Future work will use the modeling framework described here to address
431 the role specific socioeconomic and demographic variables have in statistically explaining tornado
432 casualties. Toward that end it will be necessary to obtain consistent estimates of these variables at

433 the tornado level and the model will need to include random effects related to when the tornado
434 occurs (time of day, month).

435 *Acknowledgments.* The work was partially supported by crowd funding to the second author (TF)
436 through `experiment.com`.

References

- Ashley, W. S., A. J. Kremenec, and R. Schwantes, 2008: Vulnerability due to nocturnal tornadoes. *Weather and Forecasting*, **23**, 795–807.
- Ashley, W. S., and S. M. Strader, 2016: Recipe for disaster: How the dynamic ingredients of risk and exposure are changing the tornado disaster landscape. *Bulletin of the American Meteorological Society*, **97**, 767–786.
- Berry, W. D., M. Golder, and D. Milton, 2012: Improving tests of theories positing interaction. *The Journal of Politics*, **74** (3), 653–671, doi:10.1017/S0022381612000199.
- Bohonos, J. J., and D. E. Hogan, 1999: The medical impact of tornadoes in North America. *The Journal of Emergency Medicine*, **17**, 67–73.
- Brooks, H. E., 2004: On the relationship of tornado path length and width to intensity. *Weather and Forecasting*, **19**, 310–319.
- CDC, 2012: Tornado-Related Fatalities Five States, Southeastern United States, April 25–28, 2011. Tech. rep., Centers for Disease Control and Prevention. URL www.cdc.gov/mmwr/preview/mmwrhtml/mm6128a3.htm.
- Donner, W. R., 2007: The political ecology of disaster: An analysis of factors influencing U.S. tornado fatalities and injuries, 1998–2000. *Demography*, **44** (3), 669–685, doi:10.1353/dem.2007.0024, URL <https://doi.org/10.1353/dem.2007.0024>.
- Elsner, J. B., T. H. Jagger, H. M. Widen, and D. R. Chavas, 2014: Daily tornado frequency distributions in the United States. *Environmental Research Letters*, **9** (2), 024018.
- Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436**, 686–688.

459 Emanuel, K. A., 1999: The power of a hurricane: An example of reckless driving on the informa-
 460 tion superhighway. *Weather*, **54**, 107–108.

461 Fricker, T., and J. B. Elsner, 2015: Kinetic energy of tornadoes in the United States. *PLoS ONE*,
 462 **10**, e0131090, doi:10.1371/journal.pone.0131090.

463 Fricker, T., J. B. Elsner, P. Camp, and T. H. Jagger, 2014: Empirical estimates of kinetic energy
 464 from some recent U.S. tornadoes. *Geophysical Research Letters*, **41**, 4340–4346.

465 Fricker, T., J. B. Elsner, and T. H. Jagger, 2017: Population and energy elasticity of tornado
 466 casualties. *Geophysical Research Letters*, **44**, 3941–3949, doi:10.1002/2017GL073093.

467 Greenough, G., M. McGeehin, S. M. Bernard, J. Trtanj, J. Riad, and D. Engelbert, 2001: The
 468 potential impacts of climate variability and change on health impacts of extreme weather events
 469 in the United States. *Environmental Health Perspectives*, **109**, 191–198.

470 King, G., and L. Zeng, 2006: The dangers of extreme counterfactuals. *Political Analysis*, **14**,
 471 131–159.

472 Lim, J., S. Loveridge, R. Shupp, and M. Skidmore, 2017: Double danger in the double wide:
 473 Dimensions of poverty, housing quality and tornado impacts. *Regional Science and Urban*
 474 *Economics*, **65**, 1–15, doi:10.1016/j.regsciurbeco.2017.04.003, URL [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.regsciurbeco.2017.04.003)
 475 [regsciurbeco.2017.04.003](https://doi.org/10.1016/j.regsciurbeco.2017.04.003).

476 Mileti, D. S., and P. W. O’Brien, 1992: Warnings during disasters: Normalizing communicated
 477 risk. *Social Problems*, **39**, 40–57.

478 National Oceanic and Atmospheric Administration, 2015: National Weather Service Weather Fa-
 479 tality, Injury, and Damage Statistics. URL {[http://www.nws.noaa.gov/om/hazstats/resources/](http://www.nws.noaa.gov/om/hazstats/resources/weather_fatalities.pdf)
 480 [weather_fatalities.pdf](http://www.nws.noaa.gov/om/hazstats/resources/weather_fatalities.pdf)}.

481 R Core Team, 2016: *R: A Language and Environment for Statistical Computing*. Vienna, Austria,
 482 R Foundation for Statistical Computing, URL <https://www.R-project.org/>.

483 Ramsdell, J. V., Jr, and J. P. Rishel, 2007: Tornado Climatology of the Contiguous United States.
 484 Tech. Rep. NUREG/CR-4461, PNNL-15112, Pacific Northwest National Laboratory, P.O. Box
 485 999, Richland, WA 99352.

486 Simmons, K. M., and D. Sutter, 2008: Tornado warnings, lead times, and tornado casualties: An
 487 empirical investigation. *Weather and Forecasting*, **23**, 246–258.

488 Venables, W. N., and B. D. Ripley, 2002: *Modern Applied Statistics with S*. 4th ed., Springer, New
 489 York, URL <http://www.stats.ox.ac.uk/pub/MASS4>, iISBN 0-387-95457-0.

490 Verbout, S. M., H. E. Brooks, L. M. Leslie, and D. M. Schultz, 2006: Evolution of the U.S. tornado
 491 database: 1954-2003. *Weather and Forecasting*, **21**, 86–93.

492 Wickham, H., 2009: *ggplot2: elegant graphics for data analysis*. Springer New York, URL <http://had.co.nz/ggplot2/book>.
 493

494 Zahran, S., D. Tavani, and S. Weiler, 2013: Daily variation in natural disaster casualties: Informa-
 495 tion flows, safety, and opportunity costs in tornado versus hurricane strikes. *Risk Analysis*, **33**,
 496 1265–1280.

497	LIST OF TABLES	
498	Table 1.	Descriptive statistics for the variables used to estimate the model of Eqs. 2a and
499		2b observed over the 2,192 tornadoes in our sample. Data period: 1995–2016.
500		GW: gigawatts (1 GW = 10^9 Watts) 27
501	Table 2.	Casualties by EF rating. Data period: 1995–2016. 28
502	Table 3.	Population and population density in the path of tornadoes (Casualty-Producing
503		and All) by EF damage rating. Data period: 1995–2016. 29
504	Table 4.	Coefficient estimates for the interactive model of Eqs. 2a and 2b and the ad-
505		ditive model of Eqs. 1a and 1b using data for $n = 2,192$ observed tornadoes.
506		Std. Error is the standard error of the estimate and $\Pr(> t)$ is the probability
507		of the t -value more extreme under the null hypothesis that the coefficient is
508		zero. The AIC (BIC) for the interactive model is 13223 (13246) and for the
509		additive model is 13323 (13340). The coefficients were determined via an iter-
510		ative maximum likelihood approach with the <code>glm.nb</code> function from the MASS
511		package for R (Venables and Ripley, 2002). 30

512 TABLE 1. Descriptive statistics for the variables used to estimate the model of Eqs. 2a and 2b observed over
513 the 2,192 tornadoes in our sample. Data period: 1995–2016. GW: gigawatts (1 GW = 10^9 Watts)

	Variable	Value
	Number of Tornadoes	2192
	Average number of casualties per tornado	12
	Median number of casualties per tornado	3
	Highest number of casualties	1564
	Lowest number of casualties	1
	Standard deviation of casualties	53
	Average population density per tornado (people/km ²)	223
	Greatest population density (people/km ²)	13949
	Least population density (people/km ²)	.003
	Standard deviation of population density (people/km ²)	708
	Average energy dissipation per tornado (GW)	855
	Maximum energy dissipation (GW)	66,200
	Minimum energy dissipation (GW)	.0057
	Standard deviation of energy dissipation (GW)	2900

TABLE 2. Casualties by EF rating. Data period: 1995–2016.

EF	Casualties	Tornadoes	Fatalities	Injuries	Avg. Casualties/Tor.
0	362	183	8	354	2.0
1	2300	750	79	2221	3.1
2	4888	753	218	4670	6.5
3	7933	392	536	7397	20.2
4	7187	116	457	6730	62.0
5	3298	14	432	2866	235.6

514 TABLE 3. Population and population density in the path of tornadoes (Casualty-Producing and All) by EF
515 damage rating. Data period: 1995–2016.

	Number of	Median Path	Total	Avg. No.	Median No.	Avg. Pop.	Median Pop.
EF	Tor.	Length (km)	People	People/Tor.	People/Tor.	Density (/km ²)	Density (/km ²)
Casualty-Producing Tornadoes							
0	183	2	9386	51	4	440	55
1	750	5	217053	289	17	293	37
2	753	11	397781	528	72	186	29
3	392	24	492471	1256	261	88	24
4	116	34	289722	2498	824	103	27
5	14	54	89981	6427	3318	82	25
All Tornadoes							
0	16453	1	196183	12	0	128	5
1	7463	4	698008	94	5	139	14
2	2143	9	633813	296	29	119	15
3	620	21	555967	897	125	77	17
4	140	28	294953	2107	550	91	18
5	14	54	89981	6427	3318	82	25

516 TABLE 4. Coefficient estimates for the interactive model of Eqs. 2a and 2b and the additive model of Eqs. 1a
517 and 1b using data for $n = 2,192$ observed tornadoes. Std. Error is the standard error of the estimate and $\Pr(> |t|)$ is
518 the probability of the t -value more extreme under the null hypothesis that the coefficient is zero. The AIC (BIC)
519 for the interactive model is 13223 (13246) and for the additive model is 13323 (13340). The coefficients were
520 determined via an iterative maximum likelihood approach with the `glm.nb` function from the MASS package for
521 R (Venables and Ripley, 2002).

	Estimate	Std. Error	t -value	$\Pr(> t)$
Interactive Model				
$\hat{\beta}_0$	-3.0268	0.9606	-3.15	0.0016
$\hat{\beta}_P$	-.8121	0.2294	-3.54	0.0004
$\hat{\beta}_E$	0.1661	0.0377	4.41	< 0.0001
$\hat{\beta}_{P:E}$	0.0415	0.0092	4.51	< 0.0001
Additive Model				
$\hat{\beta}_0$	-6.8825	0.5302	-13.0	< 0.0001
$\hat{\beta}_P$	0.2133	0.0248	8.62	< 0.0001
$\hat{\beta}_E$	0.3212	0.0196	16.4	< 0.0001

522

LIST OF FIGURES

523

524

Fig. 1. Location of casualty-producing tornadoes (1995–2016) by EF rating. The size of the circle is proportional to the number of casualties. 32

525

526

527

528

Fig. 2. People exposed and energy dissipation for all casualty-producing tornadoes (1995–2016). (A) Annual average number of people exposed per tornado, (B) distribution of per tornado population density, (C) annual average energy dissipation per tornado, and (D) distribution of per tornado energy dissipation. 33

529

530

531

Fig. 3. Scatter plot showing casualty-producing tornadoes by population density and energy dissipation on logarithmic scales. The number of casualties per casualty-producing tornado is given in color with the color ramp on a logarithmic scale. 34

532

533

534

535

536

537

538

539

540

Fig. 4. Estimated relationship between (A) population density and the elasticity of tornado casualties with respect to energy dissipation, and (B) energy dissipation and the elasticity of tornado casualties with respect to population density (based on coefficient estimates for the interactive model of Eqs. 2a and 2b in the top portion of Table 4). In each panel the plotted black line denotes point estimates of elasticities, and the gray ribbon indicates the 95% uncertainty bound around the point estimates. The red lines denote the point estimates of elasticities from the additive model of Eqs. 1a and 1b presented in the bottom half of Table 4. In each panel a histogram showing the distribution of the horizontal-axis variable is displayed at the bottom. 35

541

542

543

Fig. 5. Estimated effect of tornado energy dissipation on the casualty rate at various levels of population density (colors) based on coefficient estimates for the interactive model of Eqs 2a and 2b. 36

544 FIG. 1. Location of casualty-producing tornadoes (1995–2016) by EF rating. The size of the circle is propor-
545 tional to the number of casualties.

FIG. 2. People exposed and energy dissipation for all casualty-producing tornadoes (1995–2016). (A) Annual average number of people exposed per tornado, (B) distribution of per tornado population density, (C) annual average energy dissipation per tornado, and (D) distribution of per tornado energy dissipation.

FIG. 3. Scatter plot showing casualty-producing tornadoes by population density and energy dissipation on logarithmic scales. The number of casualties per casualty-producing tornado is given in color with the color ramp on a logarithmic scale.

FIG. 4. Estimated relationship between (A) population density and the elasticity of tornado casualties with respect to energy dissipation, and (B) energy dissipation and the elasticity of tornado casualties with respect to population density (based on coefficient estimates for the interactive model of Eqs. 2a and 2b in the top portion of Table 4). In each panel the plotted black line denotes point estimates of elasticities, and the gray ribbon indicates the 95% uncertainty bound around the point estimates. The red lines denote the point estimates of elasticities from the additive model of Eqs. 1a and 1b presented in the bottom half of Table 4. In each panel a histogram showing the distribution of the horizontal-axis variable is displayed at the bottom.

559 FIG. 5. Estimated effect of tornado energy dissipation on the casualty rate at various levels of population
 560 density (colors) based on coefficient estimates for the interactive model of Eqs 2a and 2b.