Risk Analysis, Vol. 24, No. 3, 2004

Dead Reckoning: Demographic Determinants of the Accuracy of Mortality Risk Perceptions

Jahn Karl Hakes[1]* and W. Kip Viscusi[2]

General patterns of bias in risk beliefs are well established in the literature, but much less is known about how these biases vary across the population. Using a sample of almost 500 people, the regression analysis in this article yields results consistent with the well-established pattern that small risks are overassessed and large risks are underassessed. The accuracy of these risk beliefs varies across demographic factors, as does the switch point at which people go from underassessment to overassessment, which we found to be 1,500 deaths annually for the full sample. Better educated people have more accurate risk beliefs, and there are important differences in the risk perception by race and gender that also may be of policy interest.

KEY WORDS: Mortality risks; risk perception

1. INTRODUCTION

Early work on biases in the public's risk beliefs focused on patterns of mortality risk perceptions. These analyses used regression analysis of the geometric mean of risk beliefs for different causes of death against actual deaths to demonstrate a systematic overestimation of low probability risks and underestimation of higher probability risks.[3] More recent papers, such as Viscusi, Hakes, and Carlin[(1)] and Benjamin and Dougan,[(2)] expanded the set of regressors to include lost life expectancy for different causes of death and age group death rates, respectively. They found that mortality risk beliefs were strongly dependent on the length of life lost with different causes of death as well as the pertinence of the risk to the respondent's age group. However, in a subsequent analysis Benjamin *et al.* were unable to reject their null hypothesis of unbiasedness in perceptions for large magnitude mortality risks.[(3)] Due to the unavailability of demographic data and small sample sizes, these studies did not investigate the variations in mortality risk perception patterns across demographic groups or the extent of variation in risk perceptions among individuals.

This article makes several advances on the literature. First, the empirical analysis uses risk perception data on an individual basis rather than average values for accident groups. This difference not only creates a much greater sample size but, more importantly, it makes possible explicit econometric adjustment to account for systematic person-specific differences in risk beliefs. Second, because of the focus on individual data, it will be possible to assess how patterns of risk overestimation and underestimation vary with demographic factors.

While there has been no multivariate analysis of the effect of demographic characteristics on perceptions of mortality risks, past research has addressed demographic variations in other types of risk perceptions. The most thoroughly examined demographic influence is the effect of gender. Savage's analysis

1 Clemson University, Clemson, SC, USA.
2 John F. Cogan, Jr. Professor of Law and Economics and Director of the Program on Empirical Legal Studies at Harvard Law School.
* Address correspondence to Jahn Karl Hakes, John E. Walker Department of Economics, Clemson University, 222 Sirrine Hall, Box 1309, Clemson, SC 29634-1309; jhakes@clemson.edu.
3 See Lichtenstein *et al.* (1978) and Morgan (1983).[(17,21)]

of perceived risks from aviation accidents, house fires, auto accidents, and stomach cancer indicated that there was higher dread regarding these hazards among women, as well as among blacks, the lesser educated, and the young.[(4)] Similarly, Gustafson's summary of the literature on fear of crime found greater risk beliefs among women.[(5)]

Research on women's environmental risk beliefs has yielded similar results. Davidson and Freudenburg's review of the environmental risk perception literature found that women had higher risk beliefs for risks of nuclear power and radioactive wastes but did not differ significantly for other risks.[(6)] Flynn *et al.* similarly found higher risk beliefs for women, with white men having the lowest risk beliefs.[(7)] Women also indicated more concern about risk of hazardous waste and global warming in the study by Bord and O'Conner.[(8)] Even female scientists report higher risk assessments for nuclear technologies in the study by Barke *et al.*[(9)]

Researchers have found higher risk beliefs among women in other contexts as well. Dosman *et al.* found that women perceive greater risks from food, while better educated and younger respondents reported lower risk beliefs.[(10)] Hersch found that women were more likely to engage in protective health-related behaviors pertaining to smoking, seatbelt use, exercising, and preventative dental care, while black men were least likely to take risk-reducing efforts.[(11)] Such gender differences are apparent at an early age, as girls have higher perceived risks of injury from play than boys,[(12)] are more able to identify risks,[(13)] and are less overly optimistic about avoiding traffic accident risks when they become young adults.[(14)]

Many of the studies above found race and education effects. The role of education was also found in Viscusi's analysis of judges' risk beliefs, who exhibited smaller biases in perceptions than those reported for the general public in the literature.[(15,16)]

Section 2 will describe the sample used for our analysis. Section 3 will compare perceived mortality risks to actual death counts to identify significant over- and underassessment of risks, and will illustrate how those perceptions vary for subsamples selected by education, gender, and race. Section 4 furthers the analysis by using multiple regression analysis to estimate the effects upon risk perceptions of the various demographic variables holding other factors constant, and thus removing biases caused by any correlations of demographic variables among individuals in our sample. Section 5 uses quantile regression analysis to illustrate how people form perceptions differently for risks of different magnitudes. Section 6 concludes the article with a summary of results and a brief discussion of the policy implications of the findings.

Table I. Sample Characteristics

Variable	Mean	Standard Error
Female = 1	0.684	0.0217
Nonwhite = 1	0.095	0.0137
College = 1	0.374	0.0225
Age	44.452	0.7136
Usable responses	22.225	0.1208

Notes: $N = 462$. Observations deleted where (1) age, race, or educational attainment not reported; (2) respondent refused to estimate actual deaths; or (3) respondent gave an extreme outlier response.

2. SAMPLE DESCRIPTION

Our sample is from a safety risk survey administered for this article during the summer of 1998 to 493 adults. The sample was recruited by a marketing firm in Phoenix, Arizona. Subjects came to a central location to participate in the survey and were reimbursed for their time. Each respondent reported standard demographic information on age, race, gender, and educational attainment, answered a series of questions eliciting their opinions on hypothetical risk situations, and completed a survey question listing 23 different potential causes of death.[4]

The completion rate for the risk assessment rate of the survey was over 90% for those who participated in the study. As Table I indicates, the sample included a wide range of demographic groups. However, the sample is not representative of the entire U.S. population, as women, for example, compromise two-thirds of the respondents.[5] The education level responses

4 The causes of death include both diseases and types of accidental deaths, but for ease of exposition will be referred to hereafter as "ailments."

5 Nonresponses were deleted, as were 18 extreme outliers that appear to have reflected respondent confusion. The overall response rate was over 90%, with nonrespondents typically refusing the entire page. These results lessen the concerns of selectivity bias. The 18 estimates that were deleted indicated over 500,000,000 annual deaths from the ailment in question. These outliers were removed to reduce the skewing effects upon the arithmetic mean, although the effects of these outliers on the geometric means were negligible. The following analysis was tested for sensitivity to the removed observations, and the results were qualitatively unaffected.

were recoded into a binary variable taking the value of one if the respondent had completed a college degree or graduate degree, and a value of zero if the respondent indicated a lower level of attainment.[6] Just under 10% of the sample were members of a racial or ethnic minority group. The modest number of racial and ethnic minorities among the respondents, which limits the ability to distinguish the statistical significance of different racial categories, led us to pool the minority groups so that the analysis will focus on average parameter effects across all minorities. We will consequently examine results from different subsamples and multivariate regression results rather than focusing on sample means alone. The risk assessment question was as follows: "In 1990 47,000 people in the United States died in automobile accidents. How many people died from the other causes of death listed below? You are not expected to know any of these answers exactly. Your best estimate will do."[7] The survey instructed respondents to put their estimates in the blanks to the right of the ailment name.

3. PATTERNS OF RISK PERCEPTION

The first three columns of Table II compare actual fatalities from each of 23 causes of death to the arithmetic means and geometric means of the respondents' mortality risk estimates.[8] Previous studies have focused on geometric means because that measure decreases the distorting effect of outliers, and for comparability we too focus on geometric means. The differences between the arithmetic and geometric means reflect the extreme right-skewness of the perception distributions, which is enhanced by the lower bound on estimates at zero.[9] For each of the 23 listed ailments, the median is quite similar to the geometric mean and as a consequence will not be reported here.[10]

The results in Table II follow the well-known empirical patterns that people overestimate the small risks of death, such as botulism and fireworks, and underestimate the larger risks, such as heart disease and diabetes. The influence of outliers on the arithmetic means makes the standard errors so large that while the overestimation of low probability risks is apparent, the arithmetic means never fall below the actual value of fatalities for large risks.

The remaining pairs of columns in Table II consequently focus upon geometric mean responses for subsamples selected by demographic groups. Columns 4 and 5 of Table II show the (geometric) means by gender, Columns 6 and 7 report results by education, and Columns 8 and 9 report the results by race.

Table II designates the mean values for which the 95% confidence interval lies above the actual deaths with an "*," means for which the 95% confidence interval is below the number of deaths by "**," and geometric means outside the subsample 95% confidence interval by "***." While there are relatively few statistically significant differences in the subsample geometric means for low-frequency causes of death, for more frequent fatalities some patterns emerge. For the most common causes of death, males underestimate the risk by less than do females, and college graduates and whites report perceived death counts closer to actual deaths.

4. MULTIPLE REGRESSION RESULTS

To determine the separate effects of gender, race, and education upon mortality risk perceptions, we use

6 The analytical results reported below are qualitatively similar to those obtained using an alternative coding for estimated number of years of schooling, wherein "some high school" = 9; "high school diploma" = 12; "some college" = 14; "college degree" = 16; and "graduate degree" = 20, while respondents reporting "other" were not coded and were dropped from the analysis.

7 Although the use of the 1990 statistic from the National Safety Council is not ideal, at the time of the survey it was the most recent year of data available for the entire set of risks in the survey and we sought to match the traffic fatality risk more closely to available risk data. It requires the respondent to consider the possibility of changes in auto fatality counts over an eight-year interval. Auto fatality numbers have been relatively steady throughout the 1990s, falling to between 41,000 and 43,000 fatalities per year by the end of the decade. Estimates by NHTSA (2001) indicate 45,000 deaths in 1990. The low coefficient of variation in auto fatalities across the years suggests that any bias would be minor.[22]

8 The fatality totals are drawn from *Vital Statistics of the United States* for the year 1993, the most recent year for which data were available at the time of the survey.[23]

9 These distortions are particularly apparent for the eight most infrequent causes of death, which have similar geometric means, but widely varying arithmetic means, particularly for birthing difficulties and measles. While most of the arithmetic means on the list are between 8 and 20 times greater than the geometric mean, these ratios for appendicitis, birthing difficulties, and measles are 42, 57, and 114, respectively.

10 The average absolute deviation between the median and the geometric mean for the ailments is 15.8%, with a median absolute deviation of 15.3% and a maximum absolute deviation of 38.1%. Twelve of the 23 geometric means were below the median, with the other 11 exceeding the median. Furthermore, for each ailment, the confidence interval for the geometric mean includes the median and vice versa.

Table II. Basic Beliefs for Full Sample

		Full Sample Perceptions		Geometric Mean of Perceived Deaths, by Gender		Geometric Mean of Perceived Deaths, by Educational Attainment		Geometric Mean of Perceived Deaths, by Race	
Cause of Death	Actual deaths in 1993 (1)	Arithmetic mean (2)	Geometric mean (3)	Men (4)	Women (5)	No college degree (6)	College degree (7)	White (8)	Nonwhite (9)
Botulism	2	13,754*	623*	668	604	680	542	636***	506
Measles	5	26,670	233*	254	225	216	261	218***	461
Fireworks	5	5,360*	229*	208	239	223	239	224***	282
Lightning	89	3,494*	241*	173***	280***	247	230	248***	181
Birthing difficulties	320	101,440	1766*	1,937	1,692	1,572	2134	1,760	1,831
Appendicitis	500	18,908	445	558	402***	422	483	440	488
Accidental electrocution	670	12,003*	1,043*	1,441***	899***	1,027	1,069	1,094***	648
Hepatitis	677	19,578*	1,365*	1,791	1,207***	1,230	1,618	1,379	1,237
Accidental firearm discharges	1,416	61,307	3,235*	2,419***	3,703***	3,271	3,178	3,203	3,570
Accidental drowning	3,979	23,635	2,935**	2,685	3,057	2,876	3,034	2,966	2,650
Fire and flames	4,175	45,243	3,076**	2,734	3,246	2,970	3,259	3,207***	2,039
Asthma	4,750	50,249*	3,543**	3,531	3,548	3,006***	4,644***	3,819***	1,725
Accidental poisoning	5,200	21,117*	2,504**	3,110	2,266***	2,468	2,563	2,537	2,206
Accidental falls	12,313	15,356	1,279**	1,180	1,327	1,044***	1,779***	1,380***	599***
Stomach cancer	13,640	61,643*	6,441**	6,186	6,561	4,715***	10,607***	6,539	5,569
Homicide	24,614	378,780*	23,000	23,787	22,647	22,823	23,296	23,261	20,647
Breast cancer	45,000	142,506*	15,276**	14,601	15,595	10,458***	28,491***	15,875***	10,561
Diabetes	47,664	48,004	4,846**	5,821	4,456***	3,655***	7,670***	4,935	4,082
Stroke	144,088	186,727	17,466**	23,393***	15,320***	12,525***	30,320***	18,219***	11,691
Lung cancer	145,000	268,667	17,017**	23,506***	14,688***	11,739***	31,279***	18,203***	8,937
All cancers	505,322	782,048	73,514**	91,476	66,401***	54,439***	119,937***	75,451***	56,824
Heart disease	720,000	748,231	41,962**	54,950	37,115***	27,768***	82,571***	45,225***	20,881
All causes	2,148,463	7,132,862*	536,915**	885,538***	423,049***	355,051***	1,034,598***	544,152	471,616

*95% confidence interval for mean lies entirely above actual number of deaths.
**95% confidence interval for mean lies entirely below actual number of deaths.
***Geometric mean for corresponding subsample is outside 95% confidence interval for this subsample.
Notes: $N = 10{,}268$. Observations deleted where (1) age, race, or educational attainment not reported; (2) respondent refused to estimate actual deaths; or (3) respondent gave an extreme outlier response.

multiple regression analysis. The dependent variable for the regression in Table III is the natural logarithm of the respondent's perceived deaths. The natural logarithm transforms the number of perceived deaths into a distribution that more closely approximates the normal and allows for interpretation of unit changes in explanatory variables in terms of a percentage change in perceived deaths. This functional form in Model 1 of Table III permits comparison to earlier results such as those reported in Lichtenstein *et al.*[17]

There are, however, two important differences. First, we use individual data for each respondent's assessment for each risk category, leading to a sample size of 10,268 rather than 23 if we had relied on geometric means. Second, the reported regressions in Model 1 and Model 2 in Table III each include recognition of fixed person-specific effects. Thus, the constant term in the regression is in effect permitted to have a different value for every individual in the sample so that differences in the average level of perceived deaths across respondents are taken into account.

The estimated relationship for Model 1 in Table III has a positive intercept, a positive slope in the relevant range of actual deaths, and is concave upward. By comparison, Lichtenstein *et al.* used the Base 10 logarithm rather than the natural logarithm, and estimated curves through the logged geometric means (log GM) of estimated deaths as a function of the logged true frequency (log TF) and logged true frequency squared ($(\log \mathrm{TF})^2$) of each cause.[17] For respondents who had their perceptions "anchored"

Table III. Risk Perception Regressions for Full Sample

Dependent Variable	(1) Fixed Effects ln(perceived deaths in 1993)	(2) OLS ln(perceived deaths in 1993)	(3) OLS ln(perceived deaths in 1993)	(4) OLS ln(perceived deaths in 1993)
ln(actual deaths in ’93)	−0.024	−0.032**	−0.032**	−0.176**
	(0.017)	(0.016)	(0.016)	(0.070)
ln(actual deaths in ’93)^2	0.033***	0.033***	0.033***	0.033***
	(0.001)	(0.001)	(0.001)	(0.001)
Age	–	–	0.041	−0.018
			(0.029)	(0.040)
Age^2	–	–	-3.78×10^{-4}	3.00×10^{-4}
			(3.08×10^{-4})	(4.26×10^{-4})
Female = 1	–	–	−0.116	0.116
			(0.162)	(0.223)
College = 1	–	–	0.332**	−0.315
			(0.151)	(0.219)
Nonwhite = 1	–	–	−0.122	0.270
			(0.304)	(0.362)
College × ln(deaths93)	–	–	–	0.078***
				(0.018)
Female × ln(deaths93)	–	–	–	−0.028
				(0.019)
Nonwhite × ln(deaths93)	–	–	–	−0.047*
				(0.028)
Age × ln(deaths93)	–	–	–	0.007**
				(0.003)
Age^2 × ln(deaths93)	–	–	–	-8.21×10^{-5}**
				(2.76×10^{-5})
Constant	5.722***	5.766***	4.714***	5.900***
	(0.096)	(0.100)	(0.611)	(0.837)
Observations	10268	10268	10268	10268
R^2	0.38	0.38	0.38	0.39

Notes: Standard errors in parentheses. Model 1 reports clustered std. errors. Models 2, 3, and 4 report clustered std. errors that are also robust with regard to heteroskedasticity; $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Crossover points (in “actual deaths in 1993”)

Model 1: crossover = 1495; Model 2: crossover = 1457;

Model 3: crossover = 1448.

Males:	1676;	Females:	1357
College:	2122;	No college:	1156
White:	1479;	Nonwhite:	1188

All errors approximately equal.

Model 4: crossover = 1268.

Males:	1405;	Females:	1208
College:	1703;	No college:	1089
White:	1282;	Nonwhite:	1151

Males had lower error than females at 95% level. Other pairs approximately equal.

by motor vehicle accident fatalities, their model estimated that log GM = 2.27 + 0.03 log TF + 0.07 (log TF)^2. They did not report estimated standard errors.[11]

If people had accurate risk beliefs, the intercept in Model 1 of Table III would be zero and the slope coefficients for ln(actual deaths) and ln(actual deaths)^2 would be one and zero, respectively, as a 1% difference in actual deaths would result in a 1% difference in perceived deaths. In Model 1, however, we see a positive and significant intercept of 5.722, so that at risk levels of one death annually people perceive 305 deaths for the illness. The

[11] A regression with each case representing the geometric mean number of perceived deaths for each ailment, which replicates the Lichtenstein *et al.*[(17)] analysis, resulted in the estimated equation ln(perceived) = 5.785*** −0.040 ln(actual deaths) +0.034***(ln(actual deaths))^2, with $R^2 = 0.897$.

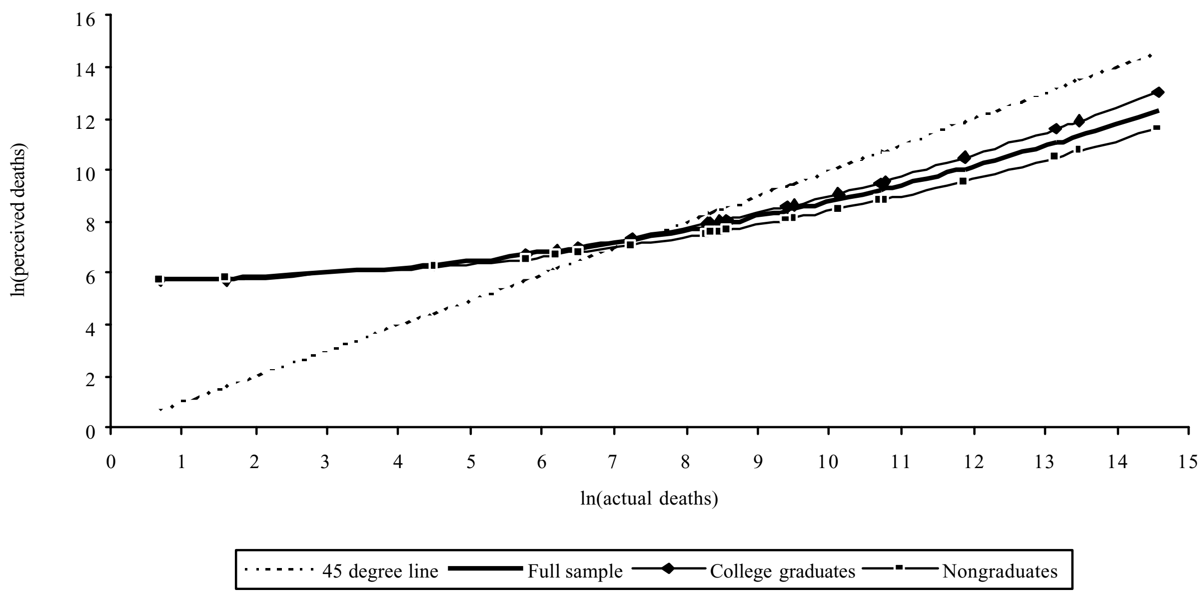

Fig. 1. Perceived mortality risks, by educational attainment (logged scale).

parameter estimate for ln(actual deaths) is insignificant, but the parameter estimate for the ln(actual deaths)^2 quadratic term is positive and significant. As actual deaths increase, perceived deaths increase at an increasing rate.[12] The slope coefficient below remains the 1.0 value in the perfect information case since $\frac{\partial(\ln(perceived))}{\partial(\ln(actual))} = 2\beta_2(\ln(actual)) + \beta_1$, yielding a point estimate for the partial derivative ranges from 0.022 to 0.938 as the ln(actual deaths) increases. The slope is 0.521 at the mean, so that mortality risk perceptions at this level incorporate only about half the information from differences in actual deaths.

Model 2 in Table III repeats the estimation without fixed effects using ordinary least squares (OLS) regression with standard errors that are both robust and clustered by individual. Robust estimation of standard errors is based on work by Huber (1967) and White (1980), and here it is used to correct for the effects of heteroskedasticity between perceived and actual deaths, while our use of clustering techniques allows us to relax the assumption of independence of residuals for observations from the same individual.[(18,19)13] Model 2 yields very similar parameter estimates, except now both of the ln(actual death) variables are significant. Systematic person-specific differences in risk belief consequently account for the significant influence of the linear risk terms.

Fig. 1 illustrates the overestimation of low probability risks and underestimation of high probability risks as seen in the classic mortality risk perception literature. The concavity of the estimated regression curve is also apparent, which indicates that the responsiveness of risk perceptions to changes in actual risk increases as the level of actual risk increases. The curvature means that larger risks are perceived more accurately than are small risks. This pattern is consistent with previous empirical findings, and is consistent with models of rational behavior when information gathering is costly.

Another similarity to the previous literature in which automobile accident fatalities are used as an anchor point for mortality risks is that the level of risk at which patterns of overestimation switch to underestimation is at approximately 1,500 annual deaths for both Models 1 and 2.[14] For death risks from which

12 It should be noted that the upward concavity is not due to either catastrophic risk or dread risk. The causes of death with the highest fatality counts in our list are typically chronic diseases, and are similar to the diseases with lower fatality counts in that mortalities from these causes do not tend to "cluster," as a disease epidemic might cause, nor are the deaths from these causes particularly sudden or dramatically violent.

13 The clustering technique is described in Stata Corp.[(24)]

14 Due to the imperfect fit of the regression, however ($R^2 = 0.38$), the confidence interval for this crossover point is quite large. For Model 1 in Table III, for example, the 95% confidence interval for the crossover point is (710, 4320). The crossover point would be expected to vary if a different anchor than automobile accidents was used.

fewer than 1,500 people die annually in the United States there is a tendency to overestimate the mortality risk. The subsample means reported in Table II do not yield a clear pattern of significant differences in the magnitude of overestimation of small risks, but differences in underestimation of causes of death resulting in more than 1,500 annual fatalities are more frequent. Using heart disease fatalities as an example, the actual number of deaths in 1993 was 720,000. The geometric mean of perceived heart disease deaths is 41,962, and the confidence interval lies entirely below the true level of risk. But the geometric mean for college educated individuals is 82,571, an estimate three times higher than for nongraduates (27,768). Similarly, males perceive higher levels of heart disease fatalities than do women (54,950–37,115), and whites' perceptions of heart disease fatalities are higher than those for nonwhites (45,225–20,881). For each paired subsample, there is at least one instance where the pairs have significantly different mortality risk perceptions.

The last two regression models shown in Table III disentangle these demographic influences. The variables used are the age of the respondent in years, the square of the respondent's age, and dummy variables for race, sex, and college completion. The square of the respondent's age was included because acquisition of additional information about mortality risks might not be linear in age, as one may receive more information as the mortality rate of one's own birth cohort increases in middle age. Empirically, the qualitative results vary somewhat with various combinations of these two variables and their cross-products with actual deaths.

Model 3 includes three demographic group indicator variables so as to identify differential intercepts. For instance, the coefficient of 0.332 for the college graduate indicator variable in Model 3 means that for a given number of actual deaths, college educated respondents, on average, perceive 33.2% more fatalities from a given cause of death than do noncollege graduate respondents of the same race and gender. The coefficients of the other indicator variables reveal that there are no statistically significant differences in mortality risk perceptions by age, race, or gender.

But as we know from above, a higher perceived number of deaths may represent either more accurate risk perceptions (as with larger risks), or less accurate risk perceptions (as with small risks). To allow a better understanding of whether demographic factors correlate with more or less accurate perceptions, Model 4 also includes not only the differential intercept terms but also interactions of each demographic factor with actual deaths. These latter terms identify differential slopes in the relationship between perceived deaths and actual deaths. With this further refinement, we see, for instance, that while the –0.315 difference in the intercept coefficient for college graduates and noncollege graduates is statistically insignificant, for each 1% increase in actual deaths (while moving from one cause of death to another), the number of additional deaths perceived by college graduates increases by 7.8% more than does the number of additional deaths perceived by noncollege graduates. From the relatively steeper slope coefficient, we learn that the mortality risk perceptions of college graduates are more elastic than the risk perceptions of noncollege graduates. That is, college graduates are more responsive to changes in actual deaths and closer to the ideal of 1.0 that indicates complete understanding of the differential threats of small and large risks.

To put this in perspective, consider the perceived number of deaths from breast cancer (where ln(actual deaths) is about 10.7). There is no differential intercept shift for college graduates as opposed to noncollege graduates. But due to the 0.078 differential slope coefficient on the college $*$ ln(actual deaths) interaction term, college graduates, other things equal, will on average perceive 83.5% more fatalities than will noncollege graduates from actual risks that are the magnitude of breast cancer. Fig. 1 illustrates the differences by education group, and shows that they are most apparent for very large risks.

The largely insignificant differential intercept estimates in Model 3 of Table III might suggest that there are not significant differences in the overall level of mortality risk perceptions by demographic group. But adding flexibility to the functional form through inclusion of the differential slope terms in Model 4 of Table III yields several interesting results. While the differential intercept terms are not statistically significant, four of the five differential slope terms are significant. While the slope for males does not differ significantly from that for females, the negative coefficient on the (nonwhite $*$ ln(actual deaths)) interaction variable indicates that at the 90% confidence level, the mortality risk perceptions of nonwhites are significantly less responsive to actual risks than are those of white graduates.

The effects of age upon perceptions are more difficult to interpret, as there are four parameter estimates to combine. Although mortality risk perceptions for infrequent causes of death do not vary significantly by age, as indicated by the insignificant coefficients on age and age^2 in Model 4 of Table III, other things equal, the responsiveness of individuals to actual deaths increases with age, peaking between the ages of 50 and 60. That is, middle-aged individuals have a better understanding of differences in levels of actual risk. The overall relationship between age, actual deaths, and perceived deaths is similar to the effects of education, and is consistent with the concept of learning from experience as one ages, with diminishing returns to additional experience. Eventually, the returns to additional experience become negative, either because of a loss in cognitive ability or because the differential character of risks faced by the elderly is influential. It is also possible that senior citizens are more likely to underestimate the current population of the United States, and consequently give overly low estimated population death counts.

To compare the accuracy of perceptions for each demographic group, the estimated equations in Table III were fitted at each of the 23 levels of ln(actual deaths) for each demographic group using the conditional mean values of the other right-hand side variables. This allows us to determine the crossover point and residual squared error for each group. The crossover points for each group appear below Table III. The only statistically significant difference in errors was the partition by gender, for which the crossover point for males is at a higher level of deaths than for females. The fitted regression line for the males, taking all 23 causes of death into consideration, lies significantly closer to the 45° line ideal than does the fitted regression line for females.

While the interaction terms between the demographic variables and actual deaths are informative, the role of education and other variables may vary across different subsamples. For example, education may be a more important component of mortality risk perception formation for females than it is for males. To fully explore these interactions with cross-product terms for every combination of demographic variables would result in a regression model with coefficients that would be very difficult to interpret. To observe the most important of these interactions while preserving simplicity of modeling, the subsequent subsections explore the results from separate regressions for different demographic groups rather than relying on more limited interaction terms.

4.1. Risk Perceptions by Education Group

Table IV presents regression results for respondents without college education and those with college education. Among those without college educations, females and nonwhites have significantly lower responsiveness to actual deaths than do similarly educated males and whites, respectively. Age does not significantly affect the perceptions of college nongraduates. Looking jointly at the parameters on ln(actual deaths) and ln(actual deaths)^2—the combined net effect of which is dominated by the quadratic term—the responsiveness to actual deaths is lower than for their college educated counterparts.

The subsample of college graduates yields no significant race or gender differences in perceptions. Instead, apart from actual deaths, the only significant determinant of perceived deaths is age, suggesting that education alleviates the differences in risk perceptions that otherwise exist between men and women, and between whites and nonwhites. The four parameters related to age combine to indicate positive, but diminishing, returns to experience, with those returns concentrated on more frequent causes of death. These latter results match those that were found in Table III and suggest that it is the subsample of college educated people that is driving those results.

Our F-test for determining whether it is appropriate to pool the subsamples of college graduates and college nongraduates resulted in a test statistic of 4.56, suggesting that it is not appropriate to pool the sample. College graduates exhibit greater responsiveness to actual mortality risks and are characterized by a regression equation that is more precise, as is reflected in the much higher R^2 statistic for the college degree holders in Table IV.

4.2. Risk Perceptions by Gender

As was seen in Columns (4) and (5) of Table II, males tended to estimate more annual fatalities than females for the more frequent causes of death, while results are mixed for less frequent causes of death. Contrary to the environmental risk perception literature, as discussed in Davidson and Freudenburg, there do not seem to be systematic differences in mortality risk perception based upon the type of risk (accident vs. disease).[(6)] What evidence exists to support their hypothesis in our data is limited to frequent causes of death, so that the differences might be based on risk magnitude. The deviations between male and

Table IV. Risk Perception Regressions, by Educational Attainment

	No College Degree ln(perceived deaths in 1993)	College Degree ln(perceived deaths in 1993)
ln(actual deaths in ’93)	−0.069	−0.356
	(0.082)	(0.148)**
ln(actual deaths in ’93)^2	0.030	0.038
	(0.001)***	(0.002)***
Age	0.018	−0.136
	(0.048)	(0.071)*
Age^2	-0.92×10^{-4}	15.52×10^{-4}
	(5.06×10^{-4})	(7.65×10^{-4})**
Age × ln(deaths93)	0.005	0.014
	(0.004)	(0.007)**
Age^2 × ln(deaths93)	-6.24×10^{-5}	-15.13×10^{-5}
	(3.82×10^{-5})	(7.57×10^{-5})**
Female = 1	0.225	−0.022
	(0.319)	(0.311)
Female × ln(deaths93)	−0.058	0.011
	(0.024)**	(0.029)
Nonwhite = 1	0.320	0.241
	(0.478)	(0.511)
Nonwhite × ln(deaths93)	−0.060	−0.029
	(0.036)*	(0.043)
Constant	4.965	8.434
	(0.995)***	(1.600)***
Observations	6368	3900
R^2	0.327	0.492

Notes: Standard errors in parentheses are clustered, and are also robust with regard to heteroskedasticity.
$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.
F-statistic for test of $\beta_{\text{no college}} = \beta_{\text{college}} = 4.56$.
Crossover points: no college = 1053; college = 1710.
Error ratio: 1.57 (no college higher), significant at 99% level.
Crossovers do not necessarily correspond with errors, as slopes vary as well as intercepts.

female fatality estimates are not strongly correlated with gender-specific death rates.[15]

Table V presents separate regression results for males and females. The significantly larger constant term and the lower magnitudes of the coefficients for ln(actual deaths) and ln(actual deaths)^2 for females are consistent with the positive differential intercept coefficient and negative differential slope coefficient for the female indicator variable in Model 4 of Table III.[16] The beneficial influence of education upon the accuracy of mortality risk perceptions is much greater for females than for males. This finding corresponds to the result in Table IV that male-female differences in risk perceptions by nongraduates disappear for college graduates.

Females also increase their fatality estimates for frequent causes of death as they age, with the additional information on these frequent causes of death making their estimates more accurate, albeit with diminishing returns. The indicator variables for race show that the differences in perceptions between white and nonwhite males is predominantly a slope effect, with nonwhites less responsive to differences in the actual number of deaths. As above, the implication of this finding is that white males have a better appreciation of the differences in levels of risk than nonwhite males. By contrast, nonwhite females offer estimates of fatalities that are not significantly different from those of white females. As with educational attainment, the F-test of pooling the male and female subsamples rejects the null that the vectors of slope parameters are the same for the two subsamples.

15 A separate exploration of the Benjamin and Dougan hypothesis regarding the importance of gender-specific death rates in affecting risk beliefs was not borne out in our data.(2)

16 Note, however, that the coefficients in Model 4 of Table III were not statistically significant.

Table V. Risk Perception Regressions, by Sex

	(1) Males ln(perceived deaths in 1993)	(2) Females ln(perceived deaths in 1993)
ln(actual deaths in '93)	−0.222	−0.197
	(0.131)*	(0.081)**
ln(actual deaths in '93)^2	0.038	0.031
	(0.002)***	(0.001)***
Age	0.050	−0.051
	(0.071)	(0.049)
Age^2	-3.55×10^{-4}	6.21×10^{-4}
	(7.50×10^{-4})	(5.21×10^{-4})
Age × ln(deaths93)	0.007	0.008
	(0.006)	(0.004)**
Age^2 × ln(deaths93)	-8.72×10^{-5}	-8.23×10^{-5}
	(6.60×10^{-5})	(3.99×10^{-5})***
College = 1	−0.288	−0.399
	(0.389)	(0.269)
College × ln(deaths93)	0.030	0.103
	(0.031)	(0.021)***
Nonwhite = 1	−0.027	0.459
	(0.619)	(0.442)
Nonwhite × ln(deaths93)	−0.081	−0.028
	(0.047)*	(0.033)
Constant	4.595	6.741
	(1.430)***	(1.045)***
Observations	3215	7053
R^2	0.412	0.384

Notes: Standard errors in parentheses are clustered, and are also robust with regard to heteroskedasticity.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.
F-statistic for test of $\beta_{\text{male}} = \beta_{\text{female}} = 9.23$.
Crossover points (in "actual deaths in 1993"): males = 1113; females = 2030.
Error ratio: 1.07 (males higher), not significant.

4.3. Risk Perceptions by Race

Table VI repeats the comparison of subsamples, this time differentiating between whites and nonwhites, which include several nonwhite racial groups whose effect is averaged across that subsample. The results above for the geometric means for individual ailments in Table II indicated that wherever a significant difference in means exists, the whites have the higher estimate, with exceptions for measles and fireworks.

Comparison of the parameter estimates in Table VI for the ln(actual deaths) and ln(actual deaths)^2 variables for the two subsamples is consistent with the results found in Table III that the estimates of whites are more responsive to actual risk. Among whites, as above in Table III, college graduates are more responsive to actual deaths than are nongraduates, males are more responsive to actual deaths than are females, and the responsiveness to actual deaths increases with age until about age 45, and then diminishes until the point where the mortality risk perceptions of 70 year olds are similar in responsiveness to actual risk as those of otherwise similar 20 year olds. By contrast, the subsample of nonwhites shows no significant differences in mortality risk perceptions by gender, and the regression results suggest that perceived risks increase for all causes of death as nonwhites age, with the effect not related to the magnitude of the actual risk. The regression upon the subsample of nonwhites does echo the findings for whites in that college graduates are more responsive to actual risk. The goodness of fit for the two models indicates that the risk perceptions of whites are more completely explained by the model than are the perceptions of nonwhites, but the pooling of ethnicities under the nonwhite banner may have contributed to this difference. The F-test to determine whether it is valid to separately estimate parameters for whites and nonwhites yields a very significant F-statistic of 11.75.

Table VI. Risk Perception Regressions, by Race

	Whites ln(perceived deaths in 1993)	Nonwhites ln(perceived deaths in 1993)
ln(actual deaths in 1993)	−0.203	−0.253
	(0.075)***	(0.256)
ln(actual deaths in 1993)^2	0.033	0.036
	(0.001)***	(0.004)***
Age	−0.002	−0.321
	(0.041)	(0.194)
Age^2	1.37×10^{-4}	0.0042
	(4.43×10^{-4})	(0.0025)*
Age × ln(deaths93)	0.0087	0.0087
	(0.0034)**	(0.0132)
Age^2 × ln(deaths93)	-0.96×10^{-4}	-1.52×10^{-4}
	(0.36×10^{-4})***	(1.56×10^{-4})
College = 1	−0.340	−0.395
	(0.231)	(0.678)
College × ln(deaths93)	0.074	0.146
	(0.019)***	(0.056)**
Female = 1	0.030	0.939
	(0.236)	(0.655)
Female × ln(deaths93)	−0.034	−0.005
	(0.020)*	(0.052)
Constant	5.632	10.950
	(0.882)***	(3.424)***
Observations	9312	956
R^2	0.402	0.332

Notes: Standard errors in parentheses are clustered, and are also robust with regard to heteroskedasticity.
$^*p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.
F-statistic for test of $(\beta_{\text{white}} = \beta_{\text{nonwhite}}) = 11.75$.
Crossover points (in "actual deaths in 1993"): whites = 1575; nonwhites = 1014.
Error ratio: 1.27 (nonwhites higher), not significant.

5. DIFFERENCES BY RISK LEVEL: QUANTILE REGRESSION RESULTS

How are risk perceptions at different magnitudes of actual risk determined? The factors that drive perception formation for low-magnitude mortality risks may not be the same as for more common causes of death. The manner in which demographic variables affect high-frequency and low-frequency risk beliefs differently can be explored using quantile regression analysis. Whereas ordinary least squares regression minimizes the sum of squared residuals, quantile regression minimizes the sum of absolute residuals, subject to the regression line resulting in q% of the residuals being negative. For the special case where $q = 50\%$, the regression line will intersect the data at the median. The econometric implication of this technique is that a low quantile regression, such as at the 10th percentile ($q(10)$), which is forced to produce 90% positive residuals, explains what is happening among the responses indicating lowest perceived risk for a given level of actual risk.

Table VII uses the functional form from Model 3 of Table III and applies it to the median (50th quantile) of the distribution of perceived deaths, as well as the 10th, 25th, 75th, and 90th quantiles. Consider the coefficients for college graduates at the $q(10)$ and $q(90)$ levels as examples of parameter estimate interpretation. The $q(10)$ coefficient for college indicates that the 10th percentile of estimates made by college graduates are 89.4% higher than those at the 10th percentile for nongraduates. Given that the intercept (1.860) is much lower than the intercepts for the other quantiles, and the strong correlation between levels of actual and perceived deaths, the overestimation by college graduates does not constitute a major error in absolute number of perceived fatalities. Taking lightning strikes as an example, the $q(10)$ fitted value for male white noncollege graduates is about 18 perceived annual deaths. While the fitted value for male

Table VII. Quantile Regression Analysis

	$q(10)$	$q(25)$	$q(50)$	$q(75)$	$q(90)$
ln(actual deaths)	0.035	−0.028	−0.037	−0.003	−0.096
	(0.039)	(0.021)	(0.024)	(0.024)	(0.036)***
ln(actual deaths)^2	0.031	0.036	0.033	0.031	0.036
	(0.003)***	(0.001)***	(0.002)***	(0.002)***	(0.002)***
Female = 1	0.001	−0.066	−0.101	−0.152	−0.251
	(0.092)	(0.048)	(0.054)*	(0.053)***	(0.079)***
College = 1	0.894	0.612	0.263	0.083	−0.011
	(0.091)***	(0.046)***	(0.052)***	(0.051)	(0.077)
Nonwhite = 1	−0.669	−0.329	0.033	0.122	0.127
	(0.152)***	(0.078)***	(0.086)	(0.083)	(0.122)
Age	−0.0005	0.022	0.037	0.055	0.070
	(0.0171)	(0.009)**	(0.010)***	(0.010)***	(0.015)***
Age^2	1.62×10^{-4}	-1.62×10^{-4}	-3.53×10^{-4}	-5.61×10^{-4}	-7.16×10^{-4}
	(1.84×10^{-4})	(0.97×10^{-4})*	(1.09×10^{-4})***	(1.06×10^{-4})***	(1.58×10^{-4})***
Constant	1.860	3.326	4.971	6.015	7.488
	(0.395)***	(0.205)***	(0.228)***	(0.222)***	(0.337)***
PseudoR^2	0.180	0.210	0.235	0.239	0.246

Note: Standard errors in parentheses are robust with regard to heteroskedasticity.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

white college graduates is relatively much higher, at 44 perceived annual deaths, the absolute difference of 26 perceived annual deaths is of little practical importance.

In the $q(50)$ model, however, where the most relevant part of the regression line—for the causes of death near the middle rows of Table II—lies underneath the 45° line depicted in Fig. 1, the higher risk assessments by college graduates represent risk perceptions that are on average more accurate by several hundred annual deaths. In the $q(50)$ regression, for male white noncollege graduates, the expected number of perceived fatalities from accidental poisoning (5,200 actual deaths) is 1,394. For male white college graduates, the estimate is 26.3% higher, resulting in an average perception of 1,813 annual deaths. While the coefficient is smaller, the practical importance of the coefficient upon the absolute number of perceived deaths is larger than in the $q(10)$ regression. The insignificant coefficient for college graduates in the $q(75)$ and $q(90)$ regressions indicates that among the highest stated perceived numbers of deaths, the responses of college graduates are statistically similar to those of other groups.

Other demographic factors also play different roles depending on the level of the risk. Nonwhites are more prone to responding with a lower number of perceived deaths for a given cause at $q(10)$ and $q(25)$ than whites. Females have lower risk beliefs at $q(50)$ and above, indicating that they are more prone to underestimating very large risks. The combined effects of the age terms indicates that numbers of perceived deaths increase at a decreasing rate as individuals age, such that older people are more likely to give the highest estimates of fatalities at a given level of actual deaths. The overall slope of the regression line (as described by ln(actual deaths) and ln(actual deaths)^2) is remarkably constant over the range of ln(perceived deaths).

The quantile regressions explain more of the variation in perceived deaths, as measured by pseudo-R^2, at the upper quantiles than they do for the lower quantiles, indicating that there is more randomness involved in the responses indicating the lowest numbers of perceived deaths.[17] People are less well informed about small risks since the information available about such risks is more limited. At the median level of ln(perceived deaths) (at 8.51, or about 5,000 perceived deaths), we see that college graduates and older respondents are making higher estimates than nongraduates and the young, and that females are more likely to give a lower response than are males.

[17] The pseudo-R^2 statistic indicates the fraction of variation in the dependent variable explained by the model. The formula is 1 – (sum of weighted deviations about estimated quantile)/(sum of weighted deviations about raw quantile), based upon the likelihood for a double exponential distribution.[(24)]

6. CONCLUSION

The general result that people overestimate small risks and underestimate large risks holds true for all major demographic groups. To the extent that risk perception biases establish a rationale for policy intervention, the case for such policies is broadly based.

Different groups in society do differ in their performance, which suggests that differential targeting of interventions such as risk information policies may be warranted. The better educated have more accurate risk beliefs and also are much more adept at incorporating age-related experiences in their risk assessments. Minorities tend to be less well informed about risk and are less responsive to the age-related experiences than whites. Females are more influenced by education and age-related information about risks than are males, and have lower risk beliefs than males for the large risks.

People exhibit more systematic thinking about large risks than small risks, which is what one might expect given the greater information that we have about frequently occurring risks. This more systematic performance should not, however, be a rationale for policy inaction because the potential health losses from underestimating large risks may be severe.

One theory for the high degree of observed risk aversion in public policy decisions is based upon public overestimation of small risks and underestimation of large risks, as argued in Viscusi.[20] According to this theory, the public's difficulty in distinguishing between differing magnitudes of risks leads to similar amounts of spending for reducing each risk. As a result, the resulting regulatory costs per statistical life saved are much higher for low probability risks, whereas the greatest gains in lifesaving will be from reducing very large risks.

Improved policy treatment of risks, assisted particularly by improved communication of risks, holds the potential to increase the cost effectiveness of public policy.[18] Our findings that improved education and age (life experience) both correlate with a greater understanding of risk differentials, and have interaction effects indicating that they are complementary, suggests that a more highly educated population may be better able to receive quantitative risk communication efforts. Given the lower accuracy of quantitative risk perceptions here for some demographic segments of the population, such as nonwhites and the lesser educated, it seems possible that these groups may need additional attention in the risk communication effort to ensure that they are adequately informed.

[18] See Fischhoff *et al.* and Kunreuther *et al.* for information on how policies are constrained by public perceptions, and on the communication of low-level risks, respectively.[25,26]

ACKNOWLEDGMENTS

Helpful comments were provided by two anonymous referees and seminar participants at the University of Florida. Viscusi's research is supported by the Harvard Olin Center for Law, Economics, and Business.

REFERENCES

1. Viscusi, W. K., Hakes, J. K., & Carlin, A. (1997). Measures of mortality risks. *Journal of Risk and Uncertainty*, *14*(3), 213–233.
2. Benjamin, D. K., & Dougan, W. R. (1997). Individuals' estimates of the risks of death: Part I—A reassessment of the previous evidence. *Journal of Risk and Uncertainty*, *15*(2), 115–133.
3. Benjamin, D. K., Dougan, W. R., & Buschena, D. (2001). Individuals' estimates of the risks of death: Part II—New evidence. *Journal of Risk and Uncertainty*, *22*(1), 35–57.
4. Savage, I. (1993). Demographic influences on risk perceptions. *Risk Analysis*, *13*(4), 413–420.
5. Gustafson, P. E. (1998). Gender differences in risk perception: Theoretical and methodological perspectives. *Risk Analysis*, *18*(6), 805–811.
6. Davidson, D. J., & Freudenburg, W. R. (1996). Gender and environmental concerns: A review and analysis of available research. *Environmental Behavior*, *28*, 302–339.
7. Flynn, J., Slovic, P., & Mertz, C. K. (1994). Gender, race, and perception of environmental-health risks. *Risk Analysis*, *14*(6), 1101–1108.
8. Bord, R. J., & O'Connor, R. E. (1997). The gender gap in environmental attitudes: The case of perceived vulnerability to risk. *Social Science Quarterly*, *78*(4), 830–840.
9. Barke, R. P., Jenkins-Smith, H., & Slovic, P. (1997). Risk perceptions of men and women scientists. *Social Science Quarterly*, *78*(1), 167–176.
10. Dosman, D. M., Adamowicz, W. L., & Hrudey, S. E. (2001). Socioeconomic determinants of health- and food safety-related risk perceptions. *Risk Analysis*, *21*(2), 307–317.
11. Hersch, J. (1996). Smoking, seatbelts, and other risky consumer decisions: Differences by gender and race. *Managerial and Decision Economics*, *17*(5), 471–481.
12. Hillier, L. M., & Morrongiello, B. A. (1998). Age and gender differences in school-age children's appraisals of injury risk. *Journal of Pediatric Psychology*, *23*(4), 229–238.
13. Morrongiello, B. A., & Rennie, H. (1998). Why do boys engage in more risk taking than girls? The role of attributions, beliefs, and risk appraisals. *Journal of Pediatric Psychology*, *23*(1), 33–43.
14. Dejoy, D. M. (1992). An examination of gender differences in traffic accident risk perception. *Accident Analysis and Prevention*, *24*(3), 237–246.
15. Viscusi, W. K. (1999). How Do Judges Think About Risk? *American Law and Economics Review*, *1*(1/2), 26–62.
16. Viscusi, W. K. (2001). Jurors, judges, and the mistreatment of risk by the courts. *Journal of Legal Studies*, *30*(1), 107–142.
17. Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal*

of Experimental Psychology: Human Learning and Memory, *4*(6), 551–578.
18. Huber, P. J. (1967). The behavior of maximum-likelihood estimates under non-standard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Sciences and Probability*. Berkeley, CA: University of California Press.
19. White, H. (1980). A heteroskedasticity-consistent covariance estimator and a direct test for heteroskedasticity. *Econometrica*, *48*, 817–830.
20. Viscusi, W. K. (1998). *Rational Risk Policy*. Oxford: Oxford University Press.
21. Morgan, G., *et al.* (1983). On judging the frequency of lethal events: A replication. *Risk Analysis*, *3*, 11–16.
22. U.S. Department of Transportation, National Highway Traffic Safety Administration. (2000). *Traffic Safety Facts: 2000: Overview*. Washington, DC: NHTSA.
23. National Center for Health Statistics. (1994). *Vital Statistics of the United States*, 1990, Vol. II, Part A. Washington, DC: Public Health Service.
24. Stata Corp. (2001). *Stata Statistical Software: Release 7.0 (User's Guide*). College Station, TX: Stata Corporation.
25. Fischhoff, B., Lichtenstein, S., Slovic, P., Derby, S. L., & Keeney, R. (1981). *Acceptable Risk*. Cambridge: Cambridge University Press.
26. Kunreuther, H., Novemsky, N., & Kahneman, D. (2001). Making low probabilities useful. *Journal of Risk and Uncertainty*, *23*(2), 103–120.