Journal of Risk and Uncertainty, 15:135–150 (1997)

Mortality Risk Perceptions: A Bayesian Reassessment

JAHN K. HAKES
Senior Research Associate, Program on Empirical Legal Studies, Harvard Law School

W. KIP VISCUSI
John F. Cogan, Jr. Professor of Law and Economics, and Director of the Program on Empirical Legal Studies, Harvard Law School

Abstract

This paper uses a Bayesian learning model to assess the respective influence of different risk measurements on mortality risk perceptions. People form risk beliefs using several sources of information, including the actual population mean death risk level, the discounted lost life expectancy, and the age-specific hazard rate considered by Benjamin and Dougan (1997). The appropriate criterion for judging the validity of risk perceptions is not the perfect information case, but rather whether people form their risk beliefs in a rational manner given a world of costly and limited risk information. Although the statistical results support the overall conclusion that the learning process is rational, the character of the learning process differs depending on the risk level. Risk-related variables are much better predictors of larger risks than of small risks, which reflects the role of information costs and the benefits of learning about larger risks.

Key words: risk perception, Bayesian, mortality, learning

JEL Classification: D81, D83

Benjamin and Dougan (1997) test the hypothesis that, due to costly information, people form their perceptions of population death risks from privately held information such as hazard rates for people in their own age group. The principal novelty of their contribution is that their empirical reference point is a rational expectations model. More specifically, they examine a two-equation system. The first equation regresses the actual death rate for the population against the age cohort death rate. The second equation regresses respondents' perceived population death rate against the age cohort death rate. Their analysis indicates that these relationships are structurally similar. One cannot reject the hypothesis that subjects are properly using the age cohort death rate as a predictor of their overall death risk perceptions in much the same manner as is indicated by the true risk relationships.

It is useful to contrast this approach with that taken in the original literature on risk perception biases, such as Lichtenstein et al. (1978) and Morgan (1983). The typical starting point for such analyses is simply to link perceived population death risks with actual death risks and to note any systematic differences in this relationship. Benjamin and Dougan's (1997) approach instead hypothesizes that the set of age cohort death risks is the principal source of risk information. In their model, it is the rationality with which respondents form perceptions based on this age cohort risk information that should be the test of the accuracy of risk perceptions.

In this paper we will focus on a third reference point for analyzing risk perceptions. Our approach is grounded in a more traditional Bayesian learning framework. In particular, our analysis will address the informational roles of not only population death risks, but also age cohort death risks and other risk measures that contribute to fatality risk perceptions. It will then be possible to assess the extent to which the age cohort death risk information is a driving force in the formation of risk beliefs. This analysis also provides an explicit test of the role of age cohort hazard rates against alternative contributors to individual perception levels.

This study uses data on perceived and actual mortality risks to test several alternative Bayesian models of the factors influencing risk beliefs. The analysis in Section 1 indicates that while the hazard rate for the individual age group is an influential factor, the overall population death rate and the discounted expected number of life years lost due to the cause of death are also influential in affecting risk perceptions. Section 2 expands the analysis to consider quantile regression models of the determinants of risk beliefs. This analysis makes it possible to distinguish the nature of the influences across different segments of the risk distribution. The predictive power of a linear perception model increases with the level of the risk and is least accurate for very small risks. Section 3 concludes the paper.

1. Empirical estimates of a Bayesian learning model

The dependent variable of interest for a learning model is the perceived annual number of deaths to all Americans caused by a condition. The mortality risk survey by Lichtenstein et al. (1978) asked respondents to assess the total number of annual deaths attributed to a variety of conditions. If the estimated total number of deaths is divided by a parameter representing the population size, the resulting statistic is the estimated probability of death for a representative individual. As a result, the appropriate theoretical formulation below will be in terms of a Bayesian learning model of updating probabilities. To be consistent with past uses of these data in the literature, however, the actual specification of the learning equation will have as the dependent variable the perceived total number of deaths associated with a particular cause, not the perceived probability of death. Thus, the empirical estimates focus on a rescaled version of the probability equation that differs only by a multiplicative constant.

Unlike the Benjamin and Dougan (1997) paper that explores the parallel relationships between actual deaths and perceived population deaths through their linkage to age cohort deaths, we will focus on a single equation in which the dependent variable is individual risk perceptions. Let us denote the value of these assessed risk beliefs by p. We will consider a highly simplified Bayesian model in which there are four independent sources of information. For each source of information the individual will act as if that information source has an associated information content equal to α_i. These α_i values indicate the equivalent number of draws from an urn that is represented by each information source.

To assist our formulation of a linear model, we make the simplifying assumption in the analysis that these are independent sources of information. In addition, we assume that probabilities can be characterized by the beta distribution of prior beliefs. The analysis below consequently will give the functional form by which people use these beta probabilities, given a series of independent sources of information, in forming their posterior beliefs. It should be noted that since the beta distribution is quite flexible and can assume a wide variety of skewed and symmetric shapes that this assumption is not particularly restrictive.

The first of the four sources of information that we distinguish is that individuals may have prior risk beliefs from sources other than those indicated below, where this prior risk has a probability value of q with informational content α_1. Second, the actual risk of death a from the particular cause of death may enter these risk perceptions with an informational weight α_2. The standard analyses of the accuracy of risk beliefs, following Lichtenstein et al. (1978), typically have treated a as the only informational component that is used in assessing the accuracy of risk beliefs. Viscusi (1992) and his previous work cited therein use this variable alone to estimate a Bayesian learning model of mortality risk beliefs. Third, the hazard rate h for the age-specific group answering the survey may be a contributor to subjects' risk perceptions. This variable, which has an informational weight α_3, is the central and solitary explanatory variable in the Benjamin and Dougan (1997) analysis. Finally, the fourth contributor to risk beliefs is the discounted lost life expectancy due to the cause of death, which we denote by d with an informational content α_4.

The functional form that arises from consideration of these four different information sources consequently is

$$p = \frac{\alpha_1 q + \alpha_2 a + \alpha_3 h + \alpha_4 d}{\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4} \tag{1}$$

where $0 \leq \alpha_i$, $i \in \{1,\ldots,4\}$. If we let

$$\beta_i = \frac{\alpha_i}{(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)}, \tag{2}$$

where $0 \leq \beta_i \leq 1$, then

$$p = \beta_1 q + \beta_2 a + \beta_3 h + \beta_4 d. \tag{3}$$

As the role of individual prior risk beliefs is not observable, the first term in Equation 3 will be represented by the constant term in a regression analysis of the determinants of risk beliefs. The other coefficients will represent the fraction of the total risk information corresponding to each information source, as is indicated by the relationship in Equation 2. Since the estimation will be in terms of total deaths rather than probabilities these coefficients will represent the proportional weight placed upon the informational source multiplied by the size of the total U.S. population base.

This Bayesian analysis consequently incorporates several different concerns that have appeared in the literature. Beginning with the analysis in Lichtenstein et al. (1978), the approach used to assess the accuracy of risk perceptions has been to regress individual risk beliefs against the total number of fatalities for a condition. The primary result that has been observed is that the intercept is positive and that a slope of less than 1.0 has been associated with the actual risk variable. Viewed from a Bayesian perspective, these results are not surprising. A positive intercept simply indicates that there are sources of information other than total population death risks that will influence risk beliefs. To the extent that there is heterogeneity in risk or other legitimate reasons for risk beliefs to vary across respondent groups, then one would expect there to be a positive intercept term. Empirical results consistent with such a Bayesian approach are reported in Viscusi (1992) and in previous works cited therein.

The first novel addition included in Equation 1 is the hazard rate variable h. Benjamin and Dougan (1997) hypothesized that the Lichtenstein et al. (1978) finding is due to costly information. Consequently, respondents rely on their age-group hazard rates rather than overall population death risks in forming their risk beliefs. Since the age group of 15–24 year-olds is pertinent for the sample of college students used as respondents, our analysis includes the 1992 death rate for that group as an explanatory variable in the analysis.[1] Empirical tests of our equations using functional forms identical to those of Benjamin and Dougan (1997) yielded results that were either similar to their findings or more supportive, given the empirical structure of their rational expectations analysis.

Our analysis consequently embodies a more general learning structure than that of Benjamin and Dougan (1997). They consider the extreme learning case in which only the age-specific hazard rate matters. This implicitly assumes other information sources are infinitely costly, whereas the previous psychology literature assumed that only the general population risk is of consequence. Our analysis includes both risks as potential information sources. Moreover, since the survey question asked respondents to state the population risk not the age-specific hazard, it seems reasonable to test the learning process with reference to that relationship.

The final risk variable included in equations 1 and 3 measures the duration of life lost. Viscusi, Hakes, and Carlin (1997) hypothesize that people may not be assessing the total mortality risk for the population but in addition may be taking into account the discounted lost life expectancy of victims when estimating death risks. Causes of death that lead to more substantial life expectancy loss such as automobile accidents may be treated differently and have greater weight in the formation of risk beliefs than causes of death associated with small life expectancy loss, such as respiratory failure of the elderly due to air pollution. Following their analysis, we include an interaction term between the discounted lost life expectancy and actual deaths. Inclusion of this term in the analysis will serve two functions. First, it will serve as a benchmark for judging the extent of the influence of the hazard rate term from the Benjamin and Dougan (1997) model as compared to other explanatory variables contributing to risk beliefs. Second, including both the hazard rate and the discounted lost life expectancy-related measure will test the importance of the Benjamin and Dougan hazard rate variable in the context of the Viscusi, Hakes, and Carlin (1997) model of the formation of risk perceptions.

1.1. Data sources

Data on the actual and perceived deaths from each condition were from Lichtenstein et al. (1978). These data also defined the medical conditions used in the analysis. Age-group hazard rate information and the data used to compute discounted lost life expectancies are drawn from the National Center for Health Statistics (1994) report, *Vital Statistics of the United States*, and the National Safety Council (1993) *Accident Facts* report. Hazard rates are stated in units of deaths per 100,000 individuals in that cohort. Viscusi, Hakes and Carlin (1997) discuss the construction of the discounted lost life expectancy variable in greater detail.[2]

The Lichtenstein et al. (1978) study asked college students at the University of Oregon for their judgments of the number of annual deaths in the United States which could be expected from each of 42 different causes. As a frame of reference, the survey told half of the students the true number of auto accident fatalities each year and the other half the true number of electrocution fatalities each year. Benjamin and Dougan (1997) pool these two groups of observations, and we follow their approach in order to maintain consistency with their study. Regression analyses with separate intercept terms failed to indicate statistically significant differences in the models.

Information on age-group hazard rates is not available for all 42 conditions in the Lichtenstein et al. (1978) dataset. Benjamin and Dougan (1997) used 1972–1973 death rates and were able to find age-group hazard rates for 29 of the 42 conditions. Using 1992 death rates, we independently also found hazard rates for 29 conditions, but each group found data for three conditions the other group did not. Changes in relative death rates between 1972 and 1992 were sufficiently small as not to substantially alter the results.

Merely showing that people use the age-group hazard rate in forming their perceptions in the same manner in which the variable is correlated with population death rates is not sufficient to indicate that this is the sole source of information from the standpoint of the Bayesian learning model. First, it is necessary to control for other significant factors to ascertain that the relationship depicted above does not rely on covariates. Second, once these other factors have been taken into account, it must be determined whether the age-group hazard rate variable adds explanatory power above that obtained from use of the other variables alone. If this cannot be shown, it is quite possible that the relationship reported by Benjamin and Dougan (1997) is actually attributable to other factors, such as discounted lost life expectancy, which are highly correlated with age-group hazard rates.

1.2. Empirical estimates

Table 1 shows a series of models designed to illustrate the usefulness of age-group hazard rates as a regressor. Each equation is a variant of Equation 3, except that one or more variables may have been omitted so as to test for their influence in the model. All standard errors are heteroskedasticity-adjusted. The first three columns are simple single-regressor models. The second column represents the linear version of the regression as modeled in Benjamin and Dougan (1997), who use a logarithmic formulation. The parameter estimate

Table 1. Regression Analysis of Perceived Total Population Deaths by Cause

	Coefficient (standard error)[1]						
Independent variable	1	2	3	4	5	6	7
Intercept	5851 **	2975 **	4500 **	−90.5	−100.9	−268.2	−240.6
	(2224)	(1406)	(1716)	(941.0)	(779.0)	(889.0)	(697.0)
Actual Population Deaths	0.049 ***			0.040 ***	−0.356 ***		−0.342***
	(0.004)			(0.005)	(0.052)		(0.054)
Hazard Rate for 15–24 Age Group		1331 ***		1075 ***		950.5 **	63.8
		(335)		(324)		(310.5)	(106.7)
Actual Deaths × Discounted Lost Life Expectancy			0.008 ***		0.057 ***	0.006 **	0.055 ***
			(0.001)		(0.007)	(0.001)	(0.007)
$\bar{R}^2$	0.46	0.44	0.57	0.74	0.92	0.77	0.91

*Parameter estimate significant at 90 percent confidence level (two-tailed test)
**Parameter estimate significant at 95 percent confidence level (two-tailed test)
***Parameter estimate significant at 99 percent confidence level (two-tailed test)
[1]All standard errors are heteroskedasticity-adjusted using the procedure in White (1980).
[2]Lost life expectancy values are discounted at 3 percent and include a 10 year lag where appropriate for diseases with a gestation period.

of 1331 suggests that each additional point of age-group hazard rate would increase an individual's estimate for the number of population deaths by 1331. Stated in another way, an increase in hazard rate of 1 death per 100,000 15–24 year olds would increase the perceived population death rate by 0.649 per 100,000.

Comparing the three single-variable models, overall deaths and age-group hazard rates seem about equally useful in determining risk perceptions, as both models have strongly significant risk variables that explain between 44 and 46 percent of the variation in people's estimates of deaths from a condition. The model using an interaction term representing the cross-product of actual deaths and the discounted lost life expectancy associated with the condition performs even more strongly with an explanatory power of 57 percent of the variation in risk beliefs.

Columns 4 through 6 of Table 1 present the possible two-variable models using these regressors. Each of these models represents an improvement over the single-variable models, as both variables have significant parameter estimates in each model, and the explanatory power is much greater. However, it is the model which uses actual deaths and the interaction term between actual deaths and discounted lost life expectancy that is most effective, as these two variables explain 92 percent of the variation in individuals' risk perceptions, with each variable being statistically significant at the 99 percent confidence level, two-tailed test.

The set of age-group hazard rates does not represent the best single predictor of individuals' risk perceptions. Nor is this variable part of the best two-variable prediction model for risk perceptions.

The final test of the age-specific hazard variable's explanatory power appears in column 7 of Table 2. This equation adds the age-group hazard variable to the regression used in column 5. That addition causes the $\bar{R}^2$ of the model to drop slightly. More importantly, the coefficient for this variable is not statistically significant and does not pass the pertinent F-test for inclusion in the equation.[3]

The significance of the age-group hazard rate variable in Benjamin and Dougan (1997) arises primarily from the log-log functional form those authors used for their regressions. Even though the learning model in Equation 3 is linear, we will also take a logarithmic transformation of each variable for comparability with their study. Each regression shown in Table 2 represents the logarithmic analog of the linear regression results reported in the corresponding column of Table 1. The age-group hazard rate variable has a significant parameter estimate (at the 90 percent or better significance level) in each regression where it appears. However, this variable is still not a very powerful predictor of individuals' risk perceptions. The single-variable regressions, which appear in columns 1 through 3 of Table 2, all indicate significant parameter estimates for the risk variable regressor. The column 1 equation using the natural logarithm of actual deaths and the column 3 equation using the natural logarithm of the interaction term between actual deaths and discounted lost life expectancy explain between 78 and 82 percent of the variation in the dependent variable (the natural logarithm of perceived deaths). However, the model using the logarithm of the age-group hazard rate has an $\bar{R}^2$ value of only 0.31. Either of the other variables used alone explains twice as much variation in the logarithm of risk perceptions as does the natural logarithm of age-group hazard rate.

Table 2. Regression Analysis of Log of Perceived Total Population Deaths by Cause

	Coefficient (standard error)[a]						
Independent Variable	1	2	3	4	5	6	7
Intercept	3.059 ***	7.821 ***	1.241 **	3.615 ***	0.399	1.814 **	1.600
	(0.364)	(0.201)	(0.481)	(0.405)	(0.933)	(0.511)	(0.961)
ln(Actual Deaths)	0.532 ***			0.472 ***	−0.283		−0.061
	(0.041)			(0.041)	(0.232)		(0.240)
ln(Hazard Rate for 15–24 Age Group)		0.473 ***		0.191 ***		0.129 **	0.122 *
		(0.101)		(0.058)		(0.056)	(0.061)
ln(Actual Deaths × Discounted Lost Life Expectancy)[b]			0.574 ***		0.869 **	0.525 **	0.592 **
			(0.042)		(0.256)	(0.042)	(0.262)
$\bar{R}^2$	0.78	0.31	0.82	0.82	0.82	0.84	0.84

*Parameter Estimate significant at 90 percent confidence level (two-tailed test)
**Parameter Estimate significant at 95 percent confidence level (two-tailed test)
***Parameter Estimate significant at 99 percent confidence level (two-tailed test)
[a]All standard errors are heteroskedasticity-adjusted using the procedure in White (1980).
[b]Lost life expectancy values are discounted at 3 percent and include a 10 year lag where appropriate for diseases with a gestation period.

The three possible two-variable combinations of variables are shown in columns 4 through 6. Each equation represents an improvement over any of the single-variable models. While the model using logged age-group hazard rates together with the logged actual death totals offers the advantage that both parameter estimates are highly significant (99 percent confidence level or better), all three equations have approximately equal overall explanatory power. The increase in the explanatory power of the model brought about by the addition of the age-group hazard rate variable to the two-variable model depicted in column 5 is significant.[4]

1.3. Implications of the learning model

The implication of these results for the role of the age-group hazard rate is quite mixed. In each case, this variable is significant individually, but is of less significance when included in equations with other risk variables likely to affect respondent risk beliefs. Overall, the main implication of recognition of the role of the hazard rate is that respondents take into account a variety of partially correlated risk information sources informing their overall risk beliefs, including actual deaths, the age group hazard rate, and the discounted lost life expectancy. Assessment of individual rationality consequently must be appropriately framed within the context of taking advantage of available sources of information in forming risk beliefs. The results here indicate that while risk perceptions are not identically equal to actual risk levels, the factors influencing these risk beliefs function in a manner that is quite consistent with the predictions of a rational Bayesian learning model in which individuals avail themselves of diverse forms of information and use this information in a sensible manner.

2. Quantile regression estimates

The analysis thus far has considered the determinants of risk perceptions based on a variety of types of information that might be available. Whereas the standard empirical analyses in this area simply investigate a link between risk perceptions and the level of the actual risk for a cause of death, this analysis has broadened the group of concerns to also add hazard rates and discounted lost life expectancy to these perceptional influences.

The main result that has held true in the literature is that overall people tend to overestimate risks of low probability events and underestimate the risks associated with extremely likely events. Thus, the character of the risk perceptions and the way in which people think about risk may differ depending on the level of the risk. Although a variety of statistical analyses such as those explored above can be instructive, they do not isolate the specific factors driving risk beliefs at different levels of risk. For example, it may be the case that people think differently about very rare risks that they face as opposed to more frequently occurring risks. Risk perceptions of extremely rare events may not be affected to a great extent by the individual's own experience with those events or knowledge of individuals who have experienced the risk outcome. Instead, the types of information sources may be much more indirect and may rely on more fragmentary assessments by the media. An interesting economic and risk perception issue more generally is whether people avail themselves of these different information sources to the same degree when considering risks of differing magnitude.

The economics of information would suggest that people should be better informed of larger risks than small risks. Since information acquisition is costly, people should be willing to incur more costs to learn about truly substantial risks to their lives. Learning through experience will also better enable people to learn about large risks since observations pertaining to frequently occurring risks such as automobile accidents will enable people to estimate these probabilities more accurately than for very rare events for which experiences are less informative.

To explore this issue rigorously, the approach we will use is a series of quantile regression models. The quantile regression models examine the determinants of the magnitude of risk perceptions at different quantiles of the risk perception distribution. This analysis can be undertaken for both the linear variation of the model as well as for the logarithmic transformation of the risk of values. In terms of the characteristics of the estimation model, the estimated coefficients of the risk perceptions p at the τ satisfy

$$Quant_{\tau}(p \mid X) = \beta_{\tau}^{'}X, \tag{4}$$

where the vector of coefficients for the τ th quantile is designated by β_{τ}.[5]

To see examples of the kinds of risks that the different percentiles of the risk distribution that will be the focus of the analysis, consider the two panels in Table 3. The top panel gives examples of the risk distribution for the subsample that was given the motor-vehicle

Table 3. Examples of Types of Risks at Different Quantiles of the Perceived Death Risk Distribution

Motor Vehicle Accident-Anchored Subsample:

	Percentile of Death Risk Perception Distribution				
	0.10	0.25	0.50	0.75	0.90
Cause of death	Measles	Appendicitis	Diabetes	Homicide	All Cancer
Perceived number of deaths	331	880	2,138	8,441	47,523
Deaths in 1978 (population 205 million)	5	902	38,950	18,860	328,000
Hazard rate for 15–24 year-olds (per 100,000)	0.0	0.0	0.3	19.9	4.9
Lost Life Expectancy	73.6	13.5	12.8	43.5	14.4

Electrocution-Anchored subsample:

	Percentile of Death Risk Perception Distribution				
	0.10	0.25	0.50	0.75	0.90
Cause of death	Measles	Tuberculosis	Drowning	Homicide	Motor vehicle accident
Perceived number of deaths	85	448	1,425	3,691	33,884
Deaths in 1978 (population 205 million)	5	3,690	7,380	18,860	55,350
Hazard rate for 15–24 year-olds (per 100,000)	0.0	0.1	2.0	19.9	34.1
Lost Life Expectancy	73.6	17.2	42.9	43.5	37.2

accident risk as their anchor. The bottom panel gives the risk percentile distribution for the subsample with the electrocution risk anchor. In each case, at the 10th percentile the risk level for the respondents risk perceptions is that associated with measles. At the first quartile, appendicitis is the pertinent risk for the motor-vehicle subsample, and tuberculosis is the pertinent risk for the electrocution-anchored subsample. At the median risk level, diabetes and drowning are the respective risks assessed by respondents. At the third quartile, homicides are the assessed risk in each of the two risk anchor cases. Finally, the risk measure representing all cancers is at the 90th percentile of the motor-vehicle accident distribution, and motor-vehicle accidents are at the 90th percentile of the electrocution-anchored risk perception distribution.

The second row of each panel indicates the perceived number of deaths associated with each percentile. As is evident in each case, the number of national deaths is fairly modest through the median, but then becomes grows geometrically as one moves to the 90th percentile of the risk distribution. Although the perceived number of deaths increases dramatically as one moves to the higher percentiles, it is evident that the actual number of deaths increases even more dramatically.

At the very low risk percentiles, respondents overassess the small number of deaths by a large amount in proportional risk terms but not by a substantial amount in terms of the absolute number of deaths. At the upper percentiles of the risk distribution, there is a much more substantial under-estimation of the absolute risks involved, particularly in the case of cancer. The final two rows present the distributions of two of the other explanatory variables in the analysis. The hazard rate for 15–24 year-olds is negligible for the risks at all quantiles below the 75th percentile. Moreover, for the 90th percentile in the motor-vehicle accident anchored subsample the pertinent cause of death is cancer, which also has a very low hazard rate for the youthful age group that constitutes the respondent set. It would consequently not be surprising that this respondent group might underassess this risk since it is not greatly pertinent to the risk factors to which they are exposed. This is the basic message of the Benjamin and Dougan (1997) paper.

The final row in Table 3 summarizes the discounted lost life expectancy for the different groups. The extent of life lost does not increase steadily with the risk perception percentile. Measles, which is at the lowest risk perception percentile shown, has the highest associated discounted lost life expectancy since it is primarily a childhood disease. In contrast, cancer has a relatively low discounted lost life expectancy even though it has a relatively high overall mortality rate. What the quantile regression analysis will do is ascertain how each of these different contributors to individual risk perceptions affects the character of risk perceptions at the different percentiles of the respondent risk perception distribution.

Table 4 presents the quantile regression results for both the linear perception model which appears in the upper panel and the log perception model which appears in the lower panel. In each case, the first column of results consists of the ordinary least squares estimates, which provide a reference point for assessing the overall effects for the sample. The subsequent five columns present the coefficient and standard error estimates for the quantiles 0.10, 0.25, 0.50, 0.75, and 0.90, respectively.

Table 4. Quantile Regressions of Perceived Population Deaths

Linear Perception Models:

	coefficient (standard error)[a]					
		Quantile				
Explanatory Variable	OLS	0.10	0.25	0.50	0.75	0.90
Constant	−240.6	−763.9	82.1	171.7	323.6	834.1
	(696.6)	(949.3)	(229.5)	(290.5)	(121.5)**	(300.3)***
Actual Deaths	−0.342	−0.115	−0.117	−0.240	−0.361	−0.321
	(0.054)**	(0.071)	(0.073)	(0.120)**	(0.093)**	(0.130)***
Hazard Rate for 15-24 Year Olds	63.75	−60.02	−0.656	−9.125	170.45	469.59
	(106.7)	(172.5)	(74.5)	(244.6)	(233.7)	(428.7)
Actual Deaths × Discounted Lost Life Expectancy[b]	0.055	0.023	0.023	0.041	0.060	0.055
	(0.007)**	(0.011)**	(0.011)**	(0.017)**	(0.013)**	(0.018)***
Pseudo R^2	0.91	0.43	0.53	0.66	0.79	0.88

Log Perception Models:

	coefficient (standard error)[a]					
		Quantile				
Explantory Variable	OLS	0.10	0.25	0.50	0.75	0.90
Constant	1.600	−0.107	1.618	1.374	3.242	2.956
	(0.961)	(1.201)	(1.095)	(1.517)	(2.194)	(1.349)*
In(Actual Deaths)	−0.061	−0.356	0.085	0.074	0.047	−0.297
	(0.240)	(0.354)	(0.338)	(0.391)	(0.476)	(0.348)

Table 4. Continued

Linear Perception Models:						
	coefficient (standard error)[a]					
	Quantile					
Explanatory Variable	OLS	0.10	0.25	0.50	0.75	0.90
In(Hazard Rate for	0.122	−0.002	0.041	0.077	0.238	0.248
15-24 Year Olds	(0.061)*	(0.082)	(0.044)	(0.090)	(0.102)**	(0.055)**
In(Actual Deaths × Discounted						
Lost Life Expectancy)[b]	0.592	0.880	0.432	0.497	0.418	0.755
	(0.262)**	(0.349)**	(0.353)	(0.429)	(0.549)	(0.379)***
Pseudo R^2	0.84	0.64	0.61	0.60	0.62	0.66

*Parameter estimate significant at 90% confidence level (two-tailed test).
**Parameter estimate significant at 95% confidence level (two-tailed test).
***Parameter estimate significant at 99% confidence level (two-tailed test).
[a]OLS standard errors are heteroskedasticity-adjusted using the procedure in White (1980). Quantile standard errors are bootstrap standard errors.
[b]Lost life expectancy values are discounted at 3 percent and include a 10 year lag where appropriate for diseases with a gestation period.

The results for the linear perception model are quite striking. Two findings are most noteworthy. First, the hazard rate for the 15–24 year olds is not statistically significant at the usual levels in any of the six regressions reported in that table. For the linear regression model, the variable representing actual deaths interacted with discounted lost life expectancy is statistically significant at every quantile, as are the actual deaths at quantiles at the median or above. Second, the explanatory power of the models steadily increases as one moves from the lowest to the highest quantiles. At the 90^{th} percentile, the fraction of the variation in the perceived population death risk levels that is explained by the equation is 0.88, which is just over double the value at the 0.10 quantile. This result suggests that the various risk explanatory variables are increasingly influential in predicting perceived risk values as the level of the risk levels becomes greater. For very small risks, risk perceptions tend to be much more random and less well explained by available sources of information that could guide individual decision making.

The bottom panel in Table 4 presenting the log perception models tempers this result. In particular, there is very little difference in the percentage of the overall variation in the log of perceptions across the different quantiles. By taking the logarithmic transformation of perceptions, much of the variation that would have occurred in the extreme cases has been reduced so that there is less of an opportunity to explain large discrepancies in risk perception. The finding with respect to age-group hazard rates is similar to the earlier results using the log perception model. This variable also tends to perform relatively better in the log perception model case than in the linear perception model case. It is, however, statistically significant in only the OLS regression and for the 0.75 and 0.90 quantiles.

The results from the top panel that the variables are relatively poor predictors of risk perceptions for the lower quantiles does seem to be borne out, notwithstanding the relatively stable pseudo R^2 values. At the 0.10 quantile, only the variable measuring discounted lost life expectancy is statistically significant at the usual levels, and at the 0.25 and 0.50 quantiles none of the variables is statistically significant. At the 0.75 quantile, one variable is significant, whereas at the 0.90 quantile two of the substantive variables are statistically significant. The predictive power of the model and the ability of these variables to be substantive contributors to risk perceptions seems to be much less pronounced in the logged perception model case than in the linear perception model results in the top panel of Table 4.

The quantile regression results consequently reinforce the earlier findings. The hazard rate for the 15–24 year old group offers negligible explanatory power in the linear risk perception models and somewhat greater explanatory power in the logged perception models. Other contributors to risk beliefs, such as the overall death rate and the discounted lost life expectancy amount also appear to be much more consistent contributors to the level of risk beliefs. Finally, small risks seem to pose particular difficulties for risk perceptions in that the usual sources of risk information are less powerful predictors of the risk beliefs that respondents have.

3. Conclusion

The character of mortality risk beliefs has proven to be of continuing interest to researchers on choice under uncertainty. Much of this interest has been stimulated by the fact that there are well-defined reference points for mortality risk values so that assessing their accuracy provides a useful benchmark for assessing the biases in risk perceptions. The longstanding result in the literature has been that people overassess low probability events and underassess larger risks, leading to the well-established size-related bias in risk perceptions.

Exactly what such a pattern of risk beliefs means is more problematic. The analysis by Benjamin and Dougan (1997) suggests that risk beliefs may not be erroneous at all. Rather, the expressed risk beliefs may simply be the rational expectations of the actual values given the age-specific hazard rates facing the respondent group. When viewed from that perspective, the relation between perceived risks and the true age-specific risks is not significantly different than the statistical correlation between actual population risks and the age-specific risk level.

The approach in this paper is somewhat different in that it utilizes the Bayesian learning approach introduced in Viscusi (1992) and in his earlier papers to explain this biased risk perception phenomenon. There is a positive intercept in a regression of perceived risk values on actual risk levels whenever individuals have prior risk beliefs that utilize information from sources other than the actual risk level. Because one such source could be the age-specific hazard rate, the traditional Bayesian learning model and the rational expectations model of Benjamin and Dougan (1997) are not necessarily inconsistent. Rather, they represent alternative perspectives and alternative tests of the underlying rationality of individuals' risk beliefs. By either approach, the appropriate reference point is not perfect risk beliefs that are identical to actual risk levels, but rather risk beliefs that reflect use of rationally incomplete information sets. Our approach, however, provides the additional advantage of simultaneous assessment of how multiple information sources affect population risk perceptions under conditions where costly and limited information cause people to be rationally uninformed.

The results from the Bayesian learning model formulated in this paper suggest that individuals use three sources of information: the actual death risk, the discounted lost life expectancy associated with the cause of death, and to a lesser extent the age-specific hazard rate. The quantile regression results were particularly instructive in that they indicated that at the quantiles associated with relatively small risk values individuals' risk perceptions could not be systematically explained by the principal contributing risk factors. These risk variables were much more influential at the upper quantiles where risk levels are larger. The difficulties people have in making judgments about low probability events may stem in part from the limited guidance that the usual sources of information provide to them in their thinking about the level of rarer hazards. Moreover, in a world of costly information there will be stronger incentives to learn about large risks than small risks.

Notes

1. This variable differs in various minor ways from the Benjamin and Dougan (1997) measure. In particular, whereas we use the hazard rate for age 15–24 year olds, they use the hazard rate for 15–19 year olds.
2. The theoretical genesis of this measure is that discounted lost life expectancy is a principal economic concern affecting fatality risk decisions. See Viscusi and Moore (1989) and Viscusi (1993).
3. The F-test statistic for removing the implicit restriction constraining the age-group hazard rate parameter to zero is 0.287, which is far from significant at any commonly used level of confidence.
4. The calculated F-statistic with 1 numerator and 46 denominator degrees of freedom is 4.58. This is significant at the 95 percent confidence level, for which the critical value is about 4.06.
5. See Roger Koenker and Gilbert Basset, Jr. (1978) for further discussion as well as Gary Chamberlain (1991). The estimator is characterized by

$$\underset{\tau}{Min}\frac{1}{n}\sum_{i=1}^{n}[\tau P(p_i \geq \beta' X_i) + (1-\tau)P(p_i < \beta' X_i)]|p_i \geq \beta' X_i, \tag{5}$$

where n is the number of causes of death, i designates cause i, and P is an indicator function that takes on a value of 1 if the designated inequality in parentheses holds and zero if it does not. We use a bootstrap estimator to obtain the value of the asymptotic standard errors.

References

Benjamin, Daniel K. and William R. Dougan. (1997). “Individuals’ Estimates of the Risks of Death: Part I—A Reassessment of the Previous Evidence” *Journal of Risk and Uncertainty* 15:2.

Lichtenstein, Sarah, Paul Slovic, Baruch Fischhoff, Mark Layman, and Barbara Combs. (1978). “Judged Frequency of Events,” *Journal of Experimental Psychology: Human Learning and Memory* 4:6, 551–578.

Morgan, M. Granger (instructor), Engineering and Public Policy/Carnegie-Mellon University Graduate Research Methods Class. (1983). “On Judging the Frequency of Lethal Events: A Replication,” *Risk Analysis* 3:1, 11–16.

National Center for Health Statistics. (1994). *Vital Statistics of the United States, 1990, Vol. II, Mortality, Part A*. Washington: Public Health Service.

National Safety Council. (1993). *Accident Facts*. Washington: National Safety Council.

Viscusi, W. Kip. (1992). *Fatal Tradeoffs: Public and Private Responsibilities for Risk*. New York: Oxford University Press.

Viscusi, W. Kip. (1993). “The Value of Risks of Life and Health,” *Journal of Economic Literature* 31:4, 1912–46.

Viscusi, W. Kip, Jahn K. Hakes and Alan Carlin. (1997). “Measures of Mortality Risks,” *Journal of Risk and Uncertainty* 14:3.

Viscusi, W. Kip and Michael J. Moore. (1989). “Rates of Time Preference and Valuations of the Duration of Life,” *Journal of Public Economics* 38, 297–317.