Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Differences in Subjective Risk Thresholds: Worker Groups as an Example

Anil Gaba, W. Kip Viscusi,

Differences in Subjective Risk Thresholds: Worker Groups as an Example

Anil Gaba • W. Kip Viscusi
INSEAD, Boulevard de Constance, Fontainebleau 77305, France
John F. Cogan Jr. Professor of Law and Economics, Harvard Law School, Hauser 302, Cambridge, Massachusetts 02138

Subjective risk perceptions are often encoded as responses to 0-1 questions in surveys or other qualitative risk scales. However, reference points for assessing an activity as risky are confounded by various characteristics of the respondents. This paper uses a sample of workers for whom quantitative risk assessments as well as dichotomous risk perception responses are available. It is shown that, given a quantitative risk measure, the thresholds for assessing an activity as "risky" vary systematically, particularly by education. The differences in such thresholds across worker groups are estimated. The resulting implications of using qualitative risk variables for assessing wage-risk tradeoffs are estimated, yielding results that are also relevant for many other areas involving similar qualitative variables.
(*Risk Perceptions*; *Risk Thresholds*; *Risk Premiums*; *Dichotomous Variables*)

1. Introduction

Surveys and field research in risk analysis and decision analysis often use questions that obtain qualitative characterizations of variables. Surveys of worker risk beliefs, for example, inquire whether the worker is exposed to dangerous or unhealthy conditions.[1] Similarly, the U.S. government frequently runs surveys that ask respondents to rate hazards in terms of whether or not they are truly very dangerous threats to individual health. Much of the research on cigarette smoking risk perceptions is of that character, as is research dealing with assessment of the risks of alcohol and other personal activities.[2] For example, one major government survey inquired whether the respondent believed that the "product is somewhat/very harmful." Asking whether a respondent perceives an activity as being risky or dangerous in some manner is possibly more the norm than is eliciting quantitative risk perception information.[3] Such qualitative characterization of variables, however, is not limited to studies of risk perceptions. Weather forecasters, for example, often give deterministic forecasts (e.g., in the form of "adverse weather" or "no adverse weather").[4]

From an empirical standpoint, these qualitative responses are coded in 0-1 terms. For example, if an activity is perceived as being risky, then it receives a value of 1; if not, it receives a value of 0. This is a legitimate quantitative metric for scoring risk perceptions. However, comparisons across different groups of people will be pertinent only if they have a comparable reference point for assessing the presence of a risk. If, for example, college-educated respondents designate an activity as risky when the probability of the hazard is modest, whereas those without a college education designate an activity as risky only once a much higher probability of

[1] This wording is, for example, included in the University of Michigan Survey of Working Conditions and the Quality of Employment Survey.

[2] See, for example, the U.S. Bureau of Alcohol, Tobacco, and Firearms (1988) and the U.S. Department of Health and Human Services (1989).

[3] There are, however, exceptions. See, for example, Kunreuther et al. (1978), in which respondents are presented with a quantitative risk scale based on individual longevity. Relative risk ratings are also frequently employed. (See Fischhoff et al. 1981).

[4] See, for example, Murphy (1977).

0025-1909/98/4406/0801$05.00

the adverse outcome is reached, then comparisons across these two groups based on subjectively coded risk variables will tend to overstate the risk levels of populations who have lower risk thresholds. A difficulty arises as to what the risk rating means if the threshold for assessing some activity as being risky differs across groups of respondents.[5]

This paper focuses on workers who have potentially hazardous jobs, but the issue it raises is quite general. Do smokers have a different threshold for what they consider risky, compared to nonsmokers? Do people who have chosen to live near toxic waste dumps or nuclear power plants similarly have different ways in which they would characterize the riskiness of their exposures? Do weather forecasters have different probability thresholds for giving a deterministic forecast of ''adverse weather'' versus ''no adverse weather?'' The issue here is not the familiar one of valuation or quantitative assessment. It may be, for example, that workers in hazardous jobs place a lower value on their health and also underestimate the quantitative magnitude of the risks. Neither of these concerns is the issue here. Rather, it is whether for any particular value of a quantitative risk assessment they are more likely to assess their job as being hazardous or risky when given a qualitative question of that type. In particular, are there different scales that people use in triggering the response that some danger or risk is being encountered?

This paper seeks to ascertain whether there are in fact differences in such risk thresholds, whether these differences vary systematically with respondent characteristics, and whether such variations are of empirical consequence. Our intent is broader than simply identifying a bias in risk categorizations. We also develop a procedure for debiasing qualitative risk judgments so that they reflect a common quantitative risk metric across individuals. The procedure introduced in this paper for estimating risk thresholds, and the resulting implications, have broader validity beyond risk assessments. For example, a weather forecaster gives a deterministic forecast of 1 if his/her probabilistic forecast is above a certain threshold and gives a deterministic forecast of 0 otherwise. The threshold or the cut-off may differ across different groups of weather forecasters. These deterministic forecasts can be standardized using the procedure introduced here.

Section 2 demonstrates how, for a sample of worker risk perceptions, the threshold levels for assessing the presence of a hazard vary systematically in expected ways. Data from the same sample for both quantitative risk assessments and the discrete risk perception variables provide an insight into the differences in risk thresholds. Section 3 explicitly estimates the critical cutoff values of the quantitative risk level used by different groups before designating a job as dangerous. The standardization procedure we introduce in §4 adjusts the subjective risk estimates to obtain debiased values of the discrete danger perception that would occur if all respondents had the same critical quantitative risk level before designating a job as being dangerous. Section 5 explores the empirical implications of this phenomenon within the context of assessing wage-risk tradeoffs. The differences are not simply of random measurement error, and the debiased danger perception variable performs much more similarly to the manner of the quantitatively scored variable. Section 6 concludes the paper.

2. Systematic Differences in Thresholds

The sample in this study consists of 335 workers exposed to hazardous chemicals. This sample of workers, at four different chemical plants, consists of both blue-collar workers and white-collar workers with substantial chemical exposures, such as research chemists.[6] All workers participating in the survey had some contact with chemicals as part of their job. The occupations included engineers, technicians, chemists, mechanics, researchers, and supervisors. Unfortunately, published injury rate data are available only for the industry as a whole, not by occupational group. The survey also did

[5] Concern with the definition, perception, and assessment of risk is of consequence for assessing the rationality of private decisions as well as the structuring of government interventions. See Kunreuther (1976), Lichtenstein et al. (1978), Machina (1987), Lichtenstein et al. (1982), and Fischhoff et al. (1984). For more general reviews, see Kahneman et al. (1982), and Hogarth (1990).

[6] This data set is drawn from the survey by Viscusi and O'Connor (1984). The complete text of the survey appears as in Viscusi and Magat (1987, Appendix G).

not include questions pertaining to the injury experience of the respondent or the respondent's co-workers.

The distinctive feature of this data set is that it included two sets of risk questions pertaining to the worker's current job. The first question led to a quantitative coding of the risk, and the second one was qualitative. The first risk perception question presented a risk scale, ranging from very safe to very dangerous, in which the individuals were asked to indicate the level of their risk. The survey then asked the respondents: "Please indicate where you feel your own job belongs by placing an X on the line between VERY SAFE JOB and VERY DANGEROUS JOB." An objective reference point was provided by an arrow marking the U.S. private sector injury and illness rate for the average worker. The anchor appeared near the left side of the scale and did not indicate the numerical value of the probability. There was no apparent anchoring effect in terms of massing of respondents. No other risk information appeared on the scale or in the survey. The response was then converted into probabilistic terms, that is, scaled between 0 and 1. This quantitative risk assessment variable is designated by *RISK*. The metric used for this scale was the U.S. Bureau of Labor Statistics injury and illness rate for industry. This risk metric served as the reference point for quantifying the risk level that respondents thought corresponded to the risk posed by their job. The average industry risk was 0.1.

There are a variety of indicators of the reliability of the *RISK* variable. The implicit value of worker injuries reflected in workers' wage-risk tradeoffs parallel results in other studies, as do the effects of *RISK* on workers' intention to quit the job.[7] These data also have been used to estimate worker utility functions for good health and post-injury, where these results are quite robust across different functional forms and yield a wide variety of results that are consistent with economic theory and other studies in the literature.[8]

Figure 1 illustrates a beta distribution fitted to the different values of *RISK* and the actual distribution for *RISK*, in the full sample. Figure 1(a) shows the fitted beta density function and Figure 1(b) demonstrates the cumulative distribution function of the fitted beta along with cumulative frequencies at select values of *RISK* in the actual data.[9] Beta distributions are capable of approximating a wide range of types of information regarding a variable with possible values between 0 and 1. For example, beta distributions are used often to reflect a variety of prior beliefs regarding proportions in Bayesian models (e.g., Lindley and Phillips 1976). Moreover, beta distributions have been used to model response to rating scales (Givon and Shapira 1984), to model discrimination abilities in product testing (Morrison and Brockway 1979, Buchanan et al. 1987), and in models of purchase intentions (Morrison 1979) and purchase behavior (Gaba and Winkler 1992), as well as in many other applications involving attitudes and survey data.

A beta density is of the form

$$f_\beta(r|a, b) = r^{a-1}(1-r)^{b-1}/B(a, b), \tag{1}$$

where $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$, with a and $b > 0$. In practice, the values for a and b can be selected by considering various fractiles of the distribution of interest (e.g., the 0.75 fractile of the distribution—the value of the variable such that 75 percent of the observations are below or equal to it, and so on) and then fitting a beta distribution to these fractiles.[10] The beta density in Figure 1(a) with parameters $a = 1.89$ and $b = 17.95$ was found to be a good fit to the distribution of *RISK*. A mean of 0.095 and a standard deviation of 0.064 for this beta density are identical to those for the variable *RISK* in the sample. Moreover, as it can be seen in Figure 1(b), the cumulative probabilities of this beta distribution are close to cumulative frequencies for *RISK* in the sample.

The greatest density of the distribution is at a risk value below 0.10. This skewed distribution has a relatively long upper tail, which reflects the fact that most jobs do not pose a certainty of risk but rather

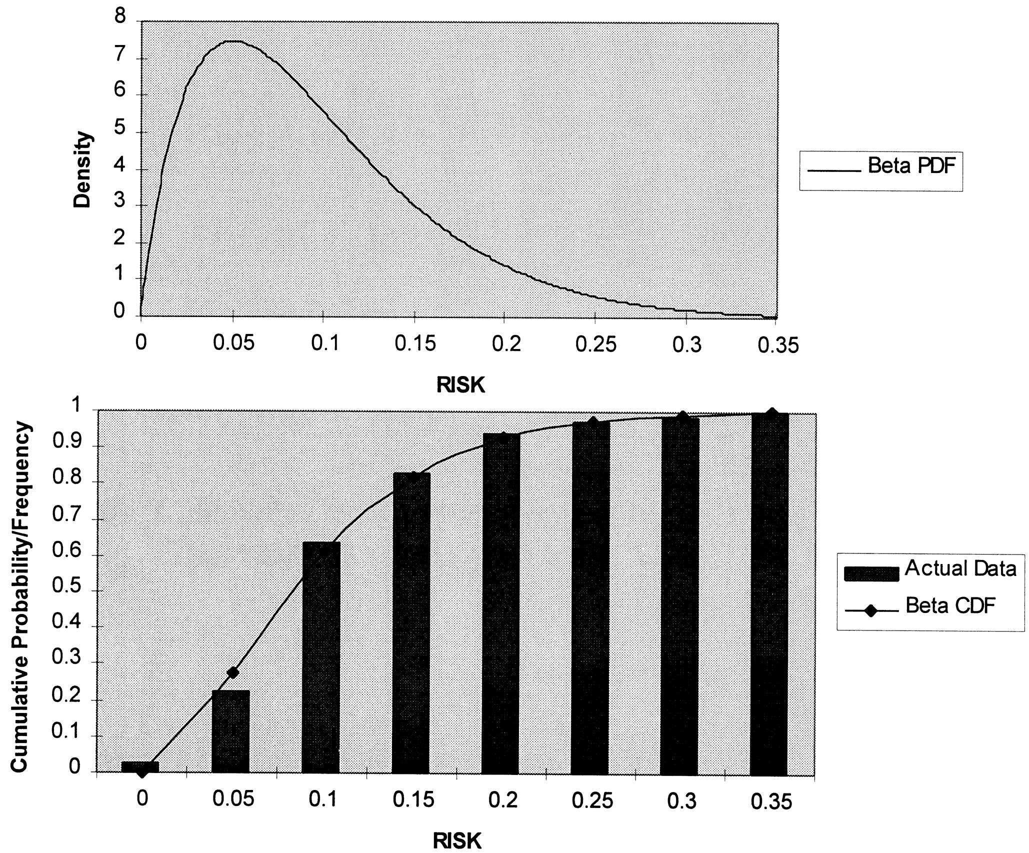
[7] See Viscusi and O'Connor (1984) for estimates of wage equations, quit intention equations, and other job attitudes variables.

[8] The estimates using these data in Viscusi and Evans (1990) explore three different functional forms, including two Taylor's series expansions. Results on the implicit value of worker injuries, the willingness to pay for reductions of risk of different magnitude, the effect of the initial risk starting point, and other matters are all consistent with theoretic predictions.

[9] By "select values," we mean that the values for *RISK* in the actual data are not limited to 0, 0.05, 0.1, and so on.

[10] See, for example, Spetzler and Staël von Holstein (1975).

Figure 1 **Distribution for the Continuous Risk Measure (a) Fitted Beta Density Function for RISK (b) Cumulative Distribution Function for Fitted Beta and Cumulative Frequencies at Select Values in Actual Data for RISK**



involve risks that tend to be events at the lower end of the probability scale. Even very high risk industries such as coal mining and construction have an injury and illness rate that is less than 1.5 times the mean assessed risk in this sample. Since the risk involved is that of nonfatal job injuries, the probability is much higher than it would be if, for example, the risk pertained to fatalities, which would be on the order of 1/10,000 annually.

The respondents also answered a qualitative risk question in which they were asked whether their job exposed them to dangerous or unhealthy working conditions. The specific wording of the question was: "Does your job at any time expose you to what you feel are physical dangers or unhealthy conditions?" This 0-1 question provided an indication of whether, on an overall basis, they considered their job as risky, thus making it possible to compare the results of this study with other worker surveys.[11]

If individuals have the same cut off levels for job risks that they consider to be risky, then the 0-1 danger perception variable suppresses some of the continuous aspects of the quantitative risk perception variable, but

[11] In particular, a very similar wording of the risk belief question was used in the University of Michigan Survey of Working Conditions. The performance of that variable and the injury rate for the worker's industry are explored in Viscusi (1979), with results paralleling those here. In principle, one could use the debiasing technique we introduce in this paper in conjunction with the University of Michigan data to adjust those results as well. The difference is that the average industry injury rate for the worker's industry would serve as the reference point, rather than the objective risk of the worker's job.

Table 1 **Proportion for whom DANGER = 1 (and *Sample Size*) within Risk Category**

Risk Measure Range	Full Sample	College-Educated	Not College-Educated	White-Collar	Blue-Collar or Technical
0–0.05	0.33	0.50	0.19	0.47	0.23
	(88)	(40)	(48)	(36)	(52)
0.06–0.10	0.52	0.68	0.40	0.58	0.46
	(129)	(56)	(73)	(60)	(69)
0.11–0.15	0.79	0.83	0.76	0.80	0.78
	(67)	(30)	(37)	(30)	(37)
0.16–0.20	0.75	0.76	0.73	0.88	0.65
	(36)	(21)	(15)	(16)	(20)
0.21–0.25	1	1	1	1	1
	(6)	(5)	(1)	(5)	(1)
0.26–0.30	1	1	1	1	1
	(4)	(1)	(3)	(1)	(3)
0.31–0.35	1	1	1	1	1
	(5)	(2)	(3)	(2)	(3)
0–0.35	0.57	0.69	0.47	0.65	0.50
	(335)	(155)	(180)	(150)	(185)

should strongly parallel it. Let c denote the cut off quantitative *RISK* value at which the respondent considers the job to be dangerous. In particular, the null hypothesis is that if individuals have the same cut off level c for what they would designate as being dangerous, then once their assessment on the BLS risk scale hits that critical level, they will score the risk as being present, leading to a coding of the danger perception variable equal to 1. The alternative hypothesis is that different worker groups have quite different values of c for which they consider the jobs to be dangerous, thus contaminating the implication of the discrete danger perception variable.

Table 1 presents the proportion in the sample for whom the subjective danger assessment value designated by *DANGER* equals 1 and the sample size within risk category. Column 1 indicates the value of the quantitative risk measure range. For example, the first category consists of all workers who scored the job risk on the BLS probability index scale as being between 0 and 0.05. The subsequent columns give the proportion in each sample group who consider their jobs as dangerous (*DANGER* = 1) and the sample size for each of the risk range rows.

The first column of Table 1 consequently gives the annual equivalent accident risk probability that the worker believes is comparable to the risk of his or her own job. The risk perception levels of workers appeared to be reasonable.[12] The second column of Table 1 indicates the fraction of workers who view their job as exposing them to dangerous or unhealthy conditions and the total number of workers in each perceived objective risk index. The focal point of this paper is how this 0-1 subjective perception variable, *DANGER*, correlates with the continuous risk measure. As can be seen from column 2 of Table 1, the fraction of workers who view their job as dangerous increases reasonably steadily with the risk level, and all workers who assess the annual injury frequency rate as being 0.21 or higher designate their job as dangerous.

There appear, however, to be some important differences across groups in the objective risk measures that trigger the *DANGER* designation. Comparison of columns 3 and 4 in Table 1 indicates that for every risk level category smaller than 0.21, college-educated workers are more likely to view their jobs as risky than those who are not college-educated. These discrepancies are

[12] The perceptions here pertain to workers' risk for their current jobs. This aspect of the results as well as the estimated compensating differentials associated with the sample are consistent with the quantitative risk perceptions being in a reasonable range.

particularly great at low risk levels, where college-educated workers are almost 3 times as likely as those who are not college-educated to view a job with a risk of 0–0.05 as hazardous. The comparisons of white-collar and blue-collar workers in the last two columns of Table 1 have similar implications. The relative differences in the risk beliefs are, however, less stark than the differences by education group, as white-collar workers in the lowest risk range are just over twice as likely as blue-collar workers to view their jobs as risky. Furthermore, note that for all worker groups a large number of respondents are in the lower risk measure ranges and relatively few respondents are in the higher risk measure ranges. This only exacerbates the reference point problem since the differences in danger assessments are greater in the lower risk ranges for all groups. Factors other than occupation and education may also affect risk beliefs, such as exposure to risk information in the media and personal risk experiences. We focus on the demographic and job determinants since these variables were included in the data set and appear in most surveys.[13]

An additional perspective on these beliefs is to use as a summary quantitative risk index whether the continuous *RISK* variable is above the average injury and illness frequency rate in the industry, where the dichotomous variable *HRISK* (0-1 dummy variable) designates whether the worker's quantitative risk belief in the sample data is a high risk job above the average industry risk of 0.1. Table 2 summarizes the cross tabulations of *HRISK* and *DANGER* for the five sample breakdowns. For the full sample, 36 percent believe that their jobs pose an above average risk. Of these workers for whom *HRISK* = 1, approximately 80 percent view their jobs as dangerous. If the objective risk score is below or equal to the industry average, the majority of workers (0.36/0.64, or 56.25 percent, from Table 2) do not view their job as risky.

The comparable patterns in Table 2 by educational status reflect the differing apparent thresholds for considering a job risky. For college-educated workers, even if *HRISK* has a value of zero, more than half of all respondents call their jobs dangerous. Four-fifths of college-educated workers in above-average risk jobs consider their positions dangerous. Workers who are not college-educated display a fairly similar pattern of responses if their jobs are below average in risk, and are much less likely to view their job as dangerous if *HRISK* equals 1. The white-collar/blue-collar split is similar but more muted, with white-collar workers being more likely to designate their jobs as dangerous if the objective risk measure is below the industry average.

Table 2 **Relation of Danger Perceptions to Continuous Risk Assessments Above The Industry Average: Proportions of Sample by Category**

Full Sample, $n = 335$

		DANGER 0	DANGER 1	Total
HRISK	0	0.36	0.28	0.64
	1	0.07	0.29	0.36
	Total	0.43	0.57	

College-Educated, $n = 155$

		DANGER 0	DANGER 1	Total
HRISK	0	0.23	0.37	0.61*
	1	0.08	0.32	0.39*
	Total	0.31	0.69	

Not College-Educated, $n = 180$

		DANGER 0	DANGER 1	Total
HRISK	0	0.46	0.21	0.67
	1	0.07	0.26	0.33
	Total	0.53	0.47	

White-Collar, $n = 150$

		DANGER 0	DANGER 1	Total
HRISK	0	0.28	0.35	0.63
	1	0.07	0.31	0.37*
	Total	0.35	0.65*	

Blue-Collar or Technical, $n = 185$

		DANGER 0	DANGER 1	Total
HRISK	0	0.42	0.23	0.65
	1	0.08	0.27	0.35
	Total	0.50	0.50	

* Row/column does not add up due to rounding.

3. Estimation of Risk Thresholds

The two sets of breakdowns of *DANGER* perceptions versus categorizations of the objective risk measure sug-

[13] Explorations using other variables, such as worker age, did not yield results that were as strong.

gest that the quantitative risk threshold that must be reached before designating a job as dangerous varies systematically with the worker population group. In this section, we explicitly estimate the implicit risk threshold levels that underlie the worker responses.

More specifically, we estimate the different values of c, the different values of the cutoff for *RISK*, for each of the worker groups. Let a beta density function $f_\beta(r|a, b)$ represent the distribution of *RISK* values in each of the worker groups. For workers in Group i ($i = 1, 2, \ldots, k$), let c_i denote the critical value of the continuous risk measure that must be attained before designating a job as dangerous. So, if a worker in Group i has a *RISK* value above c_i then that worker designates his/her job as dangerous (*DANGER* = 1), and designates his/her job as not dangerous (*DANGER* = 0) otherwise. Then the proportion of the workers in Group i with *DANGER* = 1 is

$$p_i = 1 - I_{c_i}(a, b), \tag{2}$$

and

$$1 - p_i = I_{c_i}(a, b), \tag{3}$$

is the proportion of the workers in Group i with *DANGER* = 0, where

$$I_{c_i}(a, b) = \int_0^{c_i} f_\beta(r|a, b)dr, \tag{4}$$

is an incomplete beta function. Hence, for Group i, the marginal probability that a worker is recorded with *DANGER* = 1 is p_i, and the marginal probability that a worker is recorded with *DANGER* = 0 is $1 - p_i$. Now, suppose that in a random sample of n_i workers from Group i, r_i are recorded with *DANGER* = 1 and the remaining $n_i - r_i$ with *DANGER* = 0. The likelihood function is therefore

$$l(r_1, r_2, \ldots, r_k | n_1, n_2, \ldots, n_k, p_1, p_2, \ldots, p_k) = \prod_{i=1}^{k} \binom{n_i}{r_i} p_i^{r_i}(1 - p_i)^{n_i - r_i}. \tag{5}$$

Then the maximum likelihood estimate of p_i is $\hat{p}_i = r_i/n_i$. Then it follows from (2) that the maximum likelihood estimate of $I_{c_i}(a, b)$ is

$$\hat{I}_{c_i}(a, b) = 1 - \hat{p}_i = 1 - r_i/n_i. \tag{6}$$

Further, since $I_{c_i}(a, b)$ is increasing in c_i, the maximum likelihood estimate for c_i is $\hat{c}_i$, for which

$$I_{\hat{c}_i}(a, b) = \hat{I}_{c_i}(a, b) = 1 - r_i/n_i. \tag{7}$$

Given values of a, b, n_i, and r_i, the values of $\hat{c}_i$ can be determined easily from tables on the incomplete beta function (for example, extensive tables can be found in Pearson 1934, 1968) or can be computed using the beta inverse function in statistical software such as SAS.

Using the beta density function estimated in §2 and shown in Figure 1, for the variable *RISK* and given the sample results for the *DANGER* variable, the resulting values of $\hat{c}_i$ for different groupings of the workers are shown in Table 3. These estimates clearly indicate that the white-collar/blue-collar difference is due entirely to differences in education. Within educational groups, the blue-collar values of $\hat{c}_i$ are almost identical to the estimates for all workers and only marginally different from the white-collar values.

College, however, plays a pivotal role. Workers who have completed college view a job with an accident frequency rate of 0.055 or greater as being a dangerous job, whereas workers who did not complete college have a $\hat{c}_i$ value of 0.087, which is 0.032 greater, indicating that those who are not college-educated have a greater tolerance for risk. For all workers, the $\hat{c}_i$ value is 0.071.

That college-educated workers should view their jobs as risky is quite reasonable since the research chemists in the sample are, in fact, exposed to chemical hazards.[14] What is most striking is that their characterization of risk is quite different. A much lower risk level will trigger a positive response on their part to a qualitative question of whether their job was dangerous.

The most reasonable explanation is that a difference in their valuation of risk has created a difference in subjective risk judgments.[15] College-educated workers have higher current income levels and higher lifetime wealth levels, which will raise their valuation of health status. Indeed, for this sample the income elasticity of the

[14] The composition of the risk faced by different groups might affect their risk ratings.

[15] For further analysis of the role of risk attitudes, valuations, and heterogeneity in influencing risk taking behavior, see MacCrimmon and Wehrung (1990) and Slovic and Lichtenstein (1968).

Table 3 **Estimated Cutoff Values of Continuous Risk Measure for DANGER = 1 (and *Sample Size*) by Worker Group**

	College-Educated	Not College-Educated	All Workers
White Collar	0.055 (125)	0.085 (25)	0.060 (150)
Blue Collar or Technical	0.053 (30)	0.087 (155)	0.081 (185)
All Workers	0.055 (155)	0.087 (180)	0.071 (335)

implicit value of job injuries is 1.0.[16] Because of the linkage between job risks and income, this study has focused on the influence of exogenous education characteristics on the c_i values.

An alternative hypothesis might be that the different c_i values reflect better risk information for college-educated workers. However, if this were the case, then the *RISK* values would be influenced as well. The test here is not whether college-educated workers are more likely to be aware of job risks but whether they are more likely to designate a job as dangerous if their continuous *RISK* score reaches a particular level. If college workers are more aware of risks, there is no reason to expect differential awareness that would disproportionately affect the dichotomous risk measure. In contrast, differences in valuations of risk by educational group will create a greater expected welfare loss for the college-educated from any given value of *RISK*, thus accounting for the observed discrepancy in c_i levels.

4. Standardization of Subjective Danger Thresholds

Consider, for example, a bold worker who has a cut off value of *RISK* at 0.8 before (s)he designates *DANGER* = 1 and a timid worker who similarly has a cut off value at 0.2. Based on the qualitative *DANGER* variable, comparing risk faced by the two workers in this case is meaningless since they have different reference points for assessing the presence of risk. For the dichotomous *DANGER* variable to be a valid risk measure for comparisons across workers, the cutoff value of *RISK* should be standardized across worker groups.

Let the critical value of *RISK* for designating a job as risky be the estimated c_i for all workers, 0.071. Table 4 recomputes the value of *DANGER* with this standardized cut off value, where we designate these debiased values by *STANDARDIZED DANGER*. The proportion of the sample group for whom *DANGER* = 1 appears in each cell of Table 4, and the counterpart values of *STANDARDIZED DANGER* are in parentheses below them. The changes in the danger perceptions are greatest for workers who are not college-educated. With the standardized risk cut off value, the fraction of noncollege-educated workers who view their jobs as dangerous jumps from 0.48 to 0.84 for white-collar workers and from 0.46 to 0.65 for blue-collar workers. For the college-educated, the standardization reduces the fraction with danger perceptions from 0.69 to 0.59 for white-collar workers and from 0.70 to 0.67 for blue-collar workers.

Perhaps the shifts of greatest significance are those where standardized risk cut-offs restore the two broad risk relationships that, unexpectedly, did not hold without standardization. Danger assessments for blue-collar workers now exceed those of white-collar workers (0.65 versus 0.63), whereas the reverse was true before. College-educated workers now have a lower proportion for whom *STANDARDIZED DANGER* = 1 than workers who are not college-educated (0.61 versus 0.67), whereas their unadjusted *DANGER* value was almost

Table 4 **Summary of Subjective Danger Assessments (and Danger Assessments if Identical Risk Cutoff Applies for All Workers) by Sample Group**

	Proportion with DANGER = 1 (and with STANDARDIZED DANGER = 1)		
	College-Educated	Not College-Educated	Full Sample
White Collar	0.69 (0.59)	0.48 (0.84)	0.65 (0.63)
Blue Collar or Technical	0.70 (0.67)	0.46 (0.65)	0.50 (0.65)
Full Sample	0.69 (0.61)	0.47 (0.67)	0.57 (0.64)

[16] See Viscusi and Evans (1990) for these results.

one and a half times the size of that of workers who are not college-educated. The aberrational values of *DANGER* become reversed after accounting for the different danger cutoff values.

5. Effect of Danger Standardization on Estimated Wage-Risk Tradeoffs

It is interesting to explore the effect of standardizing the *DANGER* variable in a statistical context that commonly uses such data. This exploration will indicate the extent to which the biases in respondents' codings simply generate random measurement error or in fact have statistical consequences that are more difficult to predict. We will also be able to compare the statistical performance of the debiased *STANDARDIZED DANGER* variable with that of the objective *RISK* variable to assess the extent to which our procedure generates a variable with properties similar to an objective risk measure.

The effect of differing risk thresholds on the statistical properties of the *DANGER* variables is not innocuous. Consider the influence of the differing c_i values on the value of the estimated wage-risk tradeoff. In particular, consider a standard wage equation of the form

$$EARNINGS = \alpha + \beta_1 DANGER + \sum_{i=2}^{n} \beta_i X_i + \epsilon, \quad (8)$$

where α is the constant term, the β_is are coefficients, the X_is are a series of explanatory variables, and ϵ is a random error term. If *DANGER* is subject to random measurement error, then the coefficient of *DANGER* will be biased downward. This is the standard errors-in-variables result in econometrics.[17]

As the starting point for analysis consider the estimated *EARNINGS* equation in column 1 of Table 5, where the other variables pertain to worker age (0-1 dummy variable (d.v.) *AGE*30–49, which equals 1 if age is between 30 and 49 years), race (0-1 d.v. *BLACK*), sex (0-1 d.v. *MALE*), marital status (0-1 d.v. *MARRIED*), college graduate (0-1 d.v. *COLLEGE*), and worker experience (*EXPERIENCE*, in years).[18] The results are as expected except that *BLACK* workers are paid more.[19] The estimated coefficient of *DANGER* is $2,034, which can be viewed as a risk premium—an increment to salary for a dangerous job. Given a mean value of *DANGER* of 0.50 (since 50 percent of blue collar workers stated their job as being dangerous), this implies an

Table 5 **Estimates of Earnings Equations with Subjective Danger Variable, Standardized Danger Variable, and Continuous Risk Variable (for Blue-Collar/Technical Workers)**

Dependent Variable: Earnings

Independent Variable	Coefficients (and Standard Errors)		
	DANGER MODEL	*STANDARDIZED DANGER* MODEL	*RISK* MODEL
INTERCEPT	10707*	10910*	11068*
	(526.24)	(612.95)	(611.24)
Risk Variable in the Model	2034.03*	1139.62*	7158.46*
	(407.55)	(452.25)	(3388.89)
MIDAGE	2339.09*	2141.12*	2102.81*
	(447.81)	(467.36)	(469.12)
BLACK	2218.84*	2499.63*	2264.73*
	(714.33)	(746.76)	(761.52)
MALE	1947.53*	1946.81*	1966.63*
	(411.64)	(433.97)	(436.20)
MARRIED	498.78	827.89	788.08
	(445.46)	(477.46)	(479.87)
COLLEGE	1681.39*	2105.21*	1939.92*
	(555.16)	(576.45)	(583.40)
EXPER	86.39*	78.33*	78.35*
	(16.65)	(17.37)	(17.47)
R^2	0.46	0.40	0.40

* p-value < 0.05.

[17] See, for example, Greene (1990), especially p. 294–295.

[18] The wage equations we estimate are for the blue-collar workers only. This restriction is in keeping with the common approach in the compensating differential literature. Most studies do not include white-collar workers because the different character of their jobs makes it difficult to disentangle premiums for risk from compensation for other job and personal attributes correlated with riskiness.

[19] One possible explanation for the surprising sign on the black worker variable is coding error. The race variable was one of the two variables (gender being the other) coded by the interviewer, and there may have been miscoding of mixed races and other minority groups.

average annual value of compensation for risk equal to $1,017 per blue-collar worker.

Consider the results if we replace *DANGER* by *STANDARDIZED DANGER* in the equation, thus eliminating the role of different risk cutoff values. The results in column 2 of Table 5 indicate that the risk premium is cut almost in half—to $1,140—if the job is viewed as dangerous. This reduction in the value of the estimated coefficient is the *opposite* of what one would predict if the measurement error were random. Coupled with the mean value of *STANDARDIZED DANGER* equal to 0.65 for blue-collar workers, this result implies an average annual value of compensation for risk of $740.

As a check on the appropriate level of compensation, the third column of Table 5 presents estimates for which the job risk variable is the continuous *RISK* measure. The coefficient of 7158 and the mean value of *RISK* of 0.092 for blue-collar workers imply average annual wage premium for risk of $659, which is extremely close to the $740 value obtained with *STANDARDIZED DANGER*. Thus, the danger perceptions corrected for differences in risk cut offs yield estimated risk premiums much closer to those obtained with a quantitative continuous risk index.

These assessments have two principal implications. First, the errors caused by differences in risk thresholds are not random. In this instance the result was to create an upward bias rather than the expected downward bias. Second, the standardized values generate empirical results much more similar in character to the estimates obtained using a continuous risk measure.

6. Conclusion

Qualitative variables commonly occur in research contexts. The variables often pertain to qualitative probability judgments such as those considered here. However, systematic differences across individuals in thresholds of the underlying quantitative judgments that trigger the qualitative responses might often bias the assessment of qualitative judgments.

This paper focused on perception of job hazards. There were important differences among sample members in the quantitative risk level that triggered a stated awareness of the presence of risks. Differences across educational groups and worker types both appear to be consequential, but it was the educational group bias that was by far the greatest.

Because this study analyzed a survey in which information about the underlying quantitative risk assessment as well as the dichotomous qualitative risk awareness variable was available, it was possible to estimate the differences in the risk thresholds across worker groups. These cut off values varied substantially, with college-educated workers having the lowest cut off values. Annual injury frequency rates had to be 0.032 greater for respondents who were not college-educated to indicate that a job was dangerous.

The problems raised by these differences in risk thresholds are not innocuous. In particular, they do not fit the predicted pattern for random measurement error. Rather than creating a downward bias in the estimated wage-risk tradeoffs, this difference in thresholds led to a considerable upward bias. The effects of the heterogeneity of risk thresholds consequently may be difficult to predict, in terms of both the magnitude and the direction of the influence. Researchers cannot simply dismiss these difficulties as being analogous to the random measurement error that intrinsically affects almost all data.

These results reflect a more general phenomenon in which differences in risk valuation could contaminate responses to questions that purportedly deal only with risk perception. Although, ideally, respondents should think only of the probability when asked if a job or activity is risky, their considering it to be dangerous depends also on whether or not they strongly disvalue or dislike the adverse health effects highly. Those with college educations should be less willing to incur health risks because of their greater affluence, and this in turn appears to affect their expressed risk beliefs.

The model developed here for estimating thresholds that trigger qualitative responses and then standardizing qualitative judgments to neutralize the effect of differences in thresholds could, of course, also be used in other situations such as the one considered here. To this effect, one obvious suggestion is that surveys which elicit qualitative judgments should also obtain information regarding the underlying quantitative judgments. This suggestion is consistent, for example, with studies in weather forecasting by Murphy (1977), who demonstrates that deterministic forecasts are biased by threshold effects and that from a decision-making point of view the value of probabilistic forecasts is often greater than the value of deterministic forecasts.

More generally, future surveys should include questions which elicit quantitative probability judgments (and other quantitative judgments on other variables) for which people may differ in their qualitative characterization of the quantitative values. If such quantitative information is not available, it may be possible to establish benchmark qualitative judgments using other survey data including the qualitative variable and an appropriate quantitative counterpart. Our estimation procedure could provide a methodology for calibrating the differences across respondents in the qualitative variable reference points.

References

Buchanan, B., M. Givon, and A. Goldman, ''Measurement of Discrimination Ability in Taste Tests: An Empirical Investigation,'' *J. Marketing Res.*, 24 (1987), 154–163.

Fischhoff, B., S. Lichtenstein, P. Slovic, S. L. Derby, and R. L. Keeney, *Acceptable Risk*, Cambridge University Press, Cambridge, U.K., 1981.

——, S. Watson, and C. Hope, ''Defining Risk,'' *Policy Sci.*, 17 (1984), 123–139.

Gaba, A. and R. L. Winkler, ''Implications of Errors in Survey Data: A Bayesian Model,'' *Management Sci.*, 7 (1992), 913–925.

Givon, M. M. and Z. Shapira, ''Response to Rating Scales: A Theoretical Model and its Application to the Number of Categories Problem,'' *J. Marketing Res.*, 21 (1984), 410–419.

Greene, W. H., *Econometric Analysis*, Second Ed., Macmillan, NY, 1990.

Hogarth, R. M. (ed.), *Insights in Decision Making: A Tribute to Hillel J. Einhorn*, University of Chicago Press, Chicago, IL, 1990.

Kahneman, D., P. Slovic, and A. Tversky (Eds.), *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, U.K., 1982.

Kunreuther, H., ''Limited Knowledge and Insurance Protection,'' *Public Policy*, 24, 2 (1976), 227–261.

——, H. Ginsberg, L. Miller, P. Sagi, P. Slovic, B. Borkan, and N. Katz, *Disaster Insurance Protection: Public Policy Lessons*, Wiley, NY, 1978.

Lichtenstein, S., B. Fischhoff, and L. D. Phillips, ''Calibration of Probabilities: The State of the Art to 1980,'' in D. Kahneman, P. Slovic, and A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, U.K., 1982, 306–334.

——, P. Slovic, B. Fischhoff, M. Layman, and B. Combs, ''Judged Frequency of Lethal Events,'' *J. Experimental Psychology: Human Learning and Memory*, 4 (1978), 551–578.

Lindley, D. V. and L. D. Phillips, ''Inference for a Bernoulli Process (A Bayesian View),'' *American Statistician*, 30 (1976), 112–119.

Machina, M. J., ''Decision Making in the Presence of Risk,'' *Science*, 23b (1987), 537–543.

MacCrimmon, K. R. and D. A. Wehrung, ''Characteristics of Risk Taking Executives,'' *Management Sci.*, 36 (1990), 422–435.

Morrison, D. G., ''Purchase Intentions and Purchase Behavior,'' *J. Marketing*, 43 (1979), 65–74.

—— and G. Brockway, ''A Modified Beta Binomial Model with Applications to Multiple Choice and Taste Tests,'' *Psychometrika*, 44 (1979), 427–442.

Murphy, A. H., ''The Value of Climatological, Categorical and Probabilistic Forecasts in the Cost-Loss Ratio Situation,'' *Monthly Weather Rev.*, 105 (1977), 803–816.

Pearson, K., *Tables of the Incomplete Beta Function*, Cambridge University Press, Cambridge, U.K., 1934, 1968.

Slovic, P., and S. Lichtenstein, ''The Relative Importance of Probabilities and Payoffs in Risk Taking,'' *J. Experimental Psychology Monograph Suppl.*, 78, Part 2 (1968), 1–18.

Spetzler, C. A. and C.-A. S. Staël von Holstein, ''Probability Encoding in Decision Analysis,'' *Management Sci.*, 22 (1975), 340–358.

U.S. Bureau of Alcohol, Tobacco, and Firearms, *Final Report of the Research Study of Public Opinion Concerning Warning Labels on Containers of Alcoholic Beverages*, Vol. 1, Table 2, U.S. Govt. Printing Office, Washington, DC, December 1988.

U.S. Department of Health and Human Services, *Reducing the Health Consequences of Smoking: 25 Years of Progress*, A Report to the Surgeon General, U.S. Government Printing Office, Washington, DC, 1989.

Viscusi, W. K., *Employment Hazards: An Investigation of Market Performance*, Harvard University Press, Cambridge, MA, 1987.

—— and W. Evans, ''Utility Functions That Depend on Health Status: Estimates and Economic Implications,'' *American Econ. Rev.*, 80, 3 (1990), 353–374.

—— and C. O'Connor, ''Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers?,'' *American Econ. Rev.*, 74 (1984), 942–956.

—— and W. A. Magat, *Learning about Risk*, Harvard University Press, Cambridge, MA, 1987.

Accepted by L. Robin Keller; received February 7, 1996. This paper has been with the authors 3 months for 2 revisions.