



Calhoun: The NPS Institutional Archive

Faculty and Researcher Publications

Faculty and Researcher Publications

2014

Summarizing Risk Using Risk Measures and Risk Indices

Mackenzie, Cameron A.

<http://hdl.handle.net/10945/39503>



Calhoun is a project of the Dudley Knox Library at NPS, furthering the precepts and goals of open government and government transparency. All information contained herein has been approved for release by the NPS Public Affairs Officer.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

<http://www.nps.edu/library>

This is an author's manuscript that has been submitted for publication in the journal *Risk Analysis*. (February 9, 2014).

Summarizing Risk Using Risk Measures and Risk Indices

Cameron A. MacKenzie

Defense Resources Management Institute

Naval Postgraduate School

Monterey, California

ABSTRACT

Our society is fascinated with risk in many different areas and disciplines. One of the main ways to describe and communicate the level of risk is through risk indices, which summarize risk using numbers or categories such as words, letters, or colors. These indices are used to communicate risks to the public, understand how risk is changing over time, compare among different risks, and support decision making. Given the different methods to construct risk indices, including flawed methods such as risk matrices, this paper develops specific steps that analysts can follow to create a risk index. This paper emphasizes the importance of describing risk with a probability distribution, developing a numerical risk measure that summarizes the probability distribution, and finally translating the risk measure to an index. Measuring the risk is the most difficult part and requires the analyst to summarize a probability distribution into one or possibly a few numbers. The risk measure can then be transformed to a numerical or categorical index. We apply the method outlined in this paper to construct a risk index that compares the risk of fatalities in aviation and highway transportation.

Keywords: risk measure, risk index, transportation, risk communication

Acknowledgements: The author would like to thank Elisabeth Pate-Cornell of Stanford University (who first suggested the topic) and Kash Barker of the University of Oklahoma for their helpful comments on this paper. The majority of this research was conducted while the author was a student at Stanford.

1. INTRODUCTION

Our society is preoccupied with the subject of risk. From financial risk to the risk of a terrorist attack to the risk of disease, individuals want to understand how a disastrous event may affect them and how they can avoid it. With so many different types of risks present in modern society, risk indices have increasingly become an important way to communicate the seriousness of a risk. A search for the term “risk index” in Google Scholar returns 980 results for the year 2001 compared to 4,460 results for the year 2011 even though a search for the term “risk” returns more than three times as many results for 2001 as for 2011. It would appear that risk indices are outpacing the risks that they measure. The increasing popularity of risk indices demands a careful study of risk indices and the best ways to construct a risk index.

Three well-known examples of risk indices, which summarize risk via a number or category, demonstrate their wide variety. The American Society of Civil Engineers⁽¹⁾ uses a letter grade to rate the quality of U.S. infrastructure (where the letter grades explain the vulnerability or risk of failure of infrastructure); value at risk is a single number that describes how much a portfolio could lose at a specified probability level; and the now-defunct Homeland Security Advisory System used a color and word categorization to describe the national security threat level. First, each index summarizes risk differently: one uses letters, another uses real numbers, and the third uses colors with word descriptors. Second, each risk index covers a different discipline: engineering, financial, and security. Third, the purposes of the indices vary: the letter grades are used to motivate investment and attention on U.S. infrastructure, an investment firm usually calculates value at risk for its own internal benefit, and the Homeland Security Advisory System was designed to communicate the risk of terrorism to an entire nation.

As these three examples show, risk indices can have varied purposes. A risk index

usually attempts to fulfill at least one of the following objectives.

- Describe the risk of an event accurately
- Communicate the level of risk
- Compare among different risks
- Identify the most serious risk
- Translate how risk changes over time
- Measure the effectiveness of risk-reduction strategies
- Recommend actions for a given level of risk

Despite the different goals, audiences, and means of constructing risk indices, this paper offers recommendations to aid in the construction of a risk index for any discipline. The specific goals and intended audience of a risk index should determine the final form of the index. Because risk indices summarize risks into a single or possibly two or three metrics, careful decision making where risk is present should eschew the index in favor of a more complete description of risk. As Cox⁽²⁾ demonstrates, selecting among different risk reduction strategies should rely on both the cost and effectiveness of each strategy.

This paper seeks to assist practitioners who want to construct a risk index by outlining issues that should be considered and suggesting concrete steps to construct an index. In order to accomplish that task, the paper makes the following contributions. First, the paper defines a risk index and distinguishes it from a risk measure (Section 2), which to our knowledge has not been done in the literature. Clearly defining and distinguishing between these terms is necessary to help analysts and practitioners who are measuring and communicating risk. Second, Section 3 outlines the different choices for a numerical risk measure, discusses advantages and disadvantages of each type of risk measure, and encourages using measures composed of

multiple numbers rather than resorting to a single number. Third, the paper describes options for translating a numerical measure to either a numerical or categorical risk index but emphasizes that even categorical indices should be based on sound numerical analysis (Section 4). Fourth, the guidelines are applied to an illustrative example in Section 5 focusing on the risk of transportation, and this example demonstrates how the selection of a risk measure and the type of index impacts the interpretation and perception of risks. Finally, the conclusion details how the above objectives can help determine the most appropriate type of risk index.

2. STEPS TO CREATE A RISK INDEX

The term risk index is frequently used but rarely defined. In economics, where the subject of indices has been explored both theoretically and practically, indices are “statistical summaries used to compare different situations, times, conditions, or objects.”⁽³⁾ Similarly, a risk index summarizes the risk of an event or situation by using numbers or categorical values for the purpose of identifying and comparing risks. Summarizing information appeals to all of us who want to feel informed but do not have time to fully digest the entirety of the information.

Two of the most popular ways to summarize the risk of an event into an index form include color-coded risk matrices and additive scoring methods, and people seem to gravitate to these methods because they appear objective and are relatively easy to understand and implement.^(4,5) These methods do not account for uncertainty in a mathematically rigorous manner, rely on ambiguous meanings of words such as “very frequent” and “high impact,” and create unintended consequences based on arbitrary scoring and ranking systems.⁽⁶⁾ By developing concrete steps based on principles of decision and risk analysis, this paper strives to move people away from methods like risk matrices and additive scoring methods when creating

risk indices.

To start, one must understand the most complete way to describe the risk of an event. In the tradition of probabilistic risk analysis, risk is best defined as a combination of the probability of an unwanted event occurring and the severity or consequences if that event occurs.^(6,7) Fully describing risk involves calculating a probability distribution over the range of possible consequences for an event. A risk index is a summary of the probability distribution over the range of possible consequences. Other factors, such as the voluntariness, controllability, dread, and familiarity of the risk, also influence people's perception of risk.⁽⁸⁾ Although a risk analyst may incorporate the public's reactions to a disruption within his or her consequence assessment, risk analysis and risk communication generally should seek to quantify and explain the risk and either instruct people about how to handle or mitigate the risk or empower them to make their own decisions.⁽⁹⁾ Because this paper assumes the risk index is normative with the goal of accurately informing people about the size of a risk, measuring a risk based on these other factors departs from that goal.

The terms risk measure and risk index are often used synonymously, and both attempt to summarize risk through a single value. This paper makes an important distinction between the two. A risk measure is a numerical summary of risk, and that number is a real number. As will be discussed later, we also allow the measure to be a vector of real numbers. A risk index can also be a real number or a vector of real numbers, but it is often composed of ordinal numbers, letters, words, or colors, thus widening the means of communicating risk to the user. We view risk measures to be a subset of risk indices.

To construct a risk index, an analyst should first develop a probability distribution over the range of consequences that fully describe the risk of an event, investment, or project. Next,

the analyst should calculate a numerical measure that accurately summarizes the probability distribution. Finally, the risk measure can be translated to a different scale (like a color or ordinal number) to communicate the risk to a larger audience. Choosing a risk measure and translating the measure to a numerical or categorical scale require subjective judgments by the analyst or decision maker. Many experts^(10,11,12,13) have described how to conduct probabilistic risk analysis, and this paper assumes that the analyst is able to assess a probability distribution over the consequences. The distribution could be based on data and rely on a frequentist interpretation of probability, be based on expert elicitation and a subjective Bayesian interpretation, or combine these two views of probability. Given the assumption of a probability distribution over the range of consequences, this paper focuses on the last two steps of developing a risk index. The next section discusses and comments on some of the commonly used approaches for calculating a risk measure.

3. CHOICES FOR RISK MEASURES

Actuarial science and financial mathematics researchers have extensively studied and evaluated several types of risk measures. Financial risk measures are often interpreted as the amount of capital to be held in reserve for a risky portfolio or investment. Much of the literature^(14,15,16) lists desirable axioms or properties for financial risk measures that encourage desirable investor behavior, and risk measures are evaluated against those axioms. This axiomatic approach to financial risk measurement have led to several risk measures that serve as useful quantitative tools for making decisions about investments and calculating insurance premiums.^(17,18,19) Unfortunately, that same type of sophistication has generally not been translated to risk measures outside of investment science. When the properties of more general

risk measures have been explored, the risk measures are often developed in order to describe how people naturally perceive risk.^(20,21,22,23)

Because an important step to constructing a risk index is the accurate measurement of risk, we seek to explore the choice of risk measures from a practical viewpoint that can apply to a wide variety of areas in order to help inform and education people about risk. This brief survey of risk measures is not exhaustive, but it should help risk analysts and risk managers in areas where users want to see risk measures and indices to think more systematically about how to choose a risk measure as a foundation for the index. Some of the more well-known financial risk measures are discussed in order to illustrate how ideas from financial risk could be applied to other areas, but a rich literature already exists explaining the merits and drawbacks of various financial risk measures (see e.g., Albrecht⁽²⁴⁾ and Dowd and Blake⁽²⁵⁾).

The different types of risk measures in use today can be classified into four general categories: moments of a probability distribution, quantiles of a distribution, disutility functions, and combined factors. Each of these risk measures has positive and negative aspects. A risk measure is often a single number, such as the mean or a quantile, but a risk measure can also be comprised of multiple numbers (or a vector of numbers), such as the first two moments of a distribution or three different quantiles. We first survey single-number risk measures based on the four general categories before discussing the benefits and drawbacks of using multiple-number measures.

This paper uses a random variable X to represent the consequences of an event whose risk we are interested in measuring. $X < 0$ implies negative or adverse consequences (e.g., money lost on an investment, deaths resulting from an event); $X > 0$ implies positive consequences; and $x_1 < x_2$ indicates that x_1 has worse consequences than x_2 .

3.1. Moments of a Distribution

Statistical moments are one of the most popular ways to summarize a probability distribution, and they can serve as good measures of risk. The mean value is used frequently, and the risk of an event may be stated as the mean, expected value, or expected loss.^(2,26,27) Standard deviation can also serve as a measure of risk, especially in finance, and more variation generally means more risk. Relying on just one of these measures can ignore important information, however. The mean fails to convey information about the spread of the distribution, and only relying on standard deviation as a risk measure is problematic because it gives no indication of the central tendency. Using standard deviation as a measure of risk could lead to situations where a decision maker prefers a certain outcome of no gain or loss to a highly variable situation where only positive outcomes are possible.⁽²⁸⁾ The mean and standard deviation may also fail to capture information about other properties in a distribution such as the number of modes or the broadness of the tails. Higher-order moments, such as skewness or kurtosis, can be used to model the rate of tail decay, but estimating these moments from data is highly sensitive to the sample. For heavy-tailed distributions whose higher-order moments may be infinite (e.g., the Pareto distribution), the tail index (see Beirlant et al.⁽²⁹⁾) can better measure the rate of tail decay.

A risk measure may combine the mean and standard deviation or variance into a single number in order to incorporate both the location and spread of the probability distribution. The mean-variance theory developed by Markowitz⁽³⁰⁾ argues that both metrics are important: investment portfolios should be selected to maximize the expected return and minimize variance. In insurance, the standard deviation principle defines the risk premium as the expected amount to pay in claims plus a constant term multiplied by the variance or standard deviation.^(17,31,32) The

expected number of fatalities plus the product of a risk aversion parameter and the standard deviation of fatalities has also been proposed has a measure to determine if a risk is socially acceptable.⁽³³⁾ A measure combining the first two moments $M(X)$ is presented in Eq. (1) as a weighted sum of the mean $\mu(X)$ and standard deviation $\sigma(X)$ where k is a risk aversion parameter. Because the risk measure is a function of the negative of the mean, the measure increases as consequences worsen. (Recall that $X < 0$ implies negative consequences and $X > 0$ indicates positive consequences.)

$$M(X) = -\mu(X) + k\sigma(X) \quad (1)$$

Risk neutrality implies $k = 0$, and risk aversion implies that $k > 0$. The k parameter can be tailored to reflect an individual's willingness to trade off between the mean and the standard deviation.

Identifying a single probability distribution that describes the range of potential consequences can be very difficult. Even an extensive risk analysis may be unable to choose among multiple probability distributions, especially if different scenarios or "states of the world" are possible where each scenario generates a different distribution. Each distribution can generate a different mean and standard deviation. A risk measure could be a weighted sum of the expected value according to the most likely distribution (i.e., the most probable scenario) and the smallest expected value from all possible distributions.⁽³⁴⁾ Eq. (2) depicts this risk measure designed to deal with uncertainty in the probability distributions, where $\mu_{ML}(X)$ is the most likely expected value, $\mu_{min}(X)$ is the worst-case expected value, and $\rho \in [0,1]$ is the analyst's or decision maker's degree of confidence in the most likely distribution.

$$M(X) = -\rho\mu_{ML}(X) - (1 - \rho)\mu_{min}(X) \quad (2)$$

Alternatively, a hierarchical model could assess a probability distribution over the range of

possible scenarios and a risk measure, such as the expected value, can be calculated using the probability of each scenario.

3.2. Quantiles of a Distribution

Risk analysis often focuses on extreme or high-consequence events. If these extreme events occur infrequently or have a small probability of occurring, quantiles can be a useful tool to measure these risks. The q -quantile of a distribution, $q \in (0,1)$, is the largest value x_q (technically, the supremum) such that the probability that the consequences are worse than x_q is less than or equal to q , or $\sup\{x_q: P(X \leq x_q) \leq q\}$.⁽³⁵⁾ Perhaps the best example of using a quantile as a risk measure is value at risk. With origins dating to the early twentieth century but popularized in the 1990s, value at risk can be defined as the minimum amount of capital that would need to be added to a risky investment such that the probability of loss does not exceed a specified probability q .⁽³⁶⁾ The functional form is $\text{VAR}_q(X) = -x_q$ where x_q is the q -quantile. For example, if $\text{VAR}_{0.05}(X) = \$1000$, there is a 5% chance that a portfolio will lose at least \$1000, or alternatively, if \$1000 is added to the value of the portfolio, there is a 5% chance the portfolio will lose money. The value-at-risk measure does not, however, describe the maximum possible loss or take into account losses that exceed value at risk (Figure 1). Because the value at risk of a weighted sum of two investments X and Y , $\text{VAR}_q(\alpha X + (1 - \alpha)Y)$ where $\alpha \in [0,1]$, is in general not less than or equal to the weighted sum of each value-at-risk measure, $\alpha \text{VAR}_q(X) + (1 - \alpha) \text{VAR}_q(Y)$, this measure does not encourage diversification of risky investments, which is a desirable characteristic of financial risk measures.^(14,36) One solution to these limitations of value at risk is to use the weighted averages of the quantiles for different values of q as a risk measure.⁽¹⁶⁾

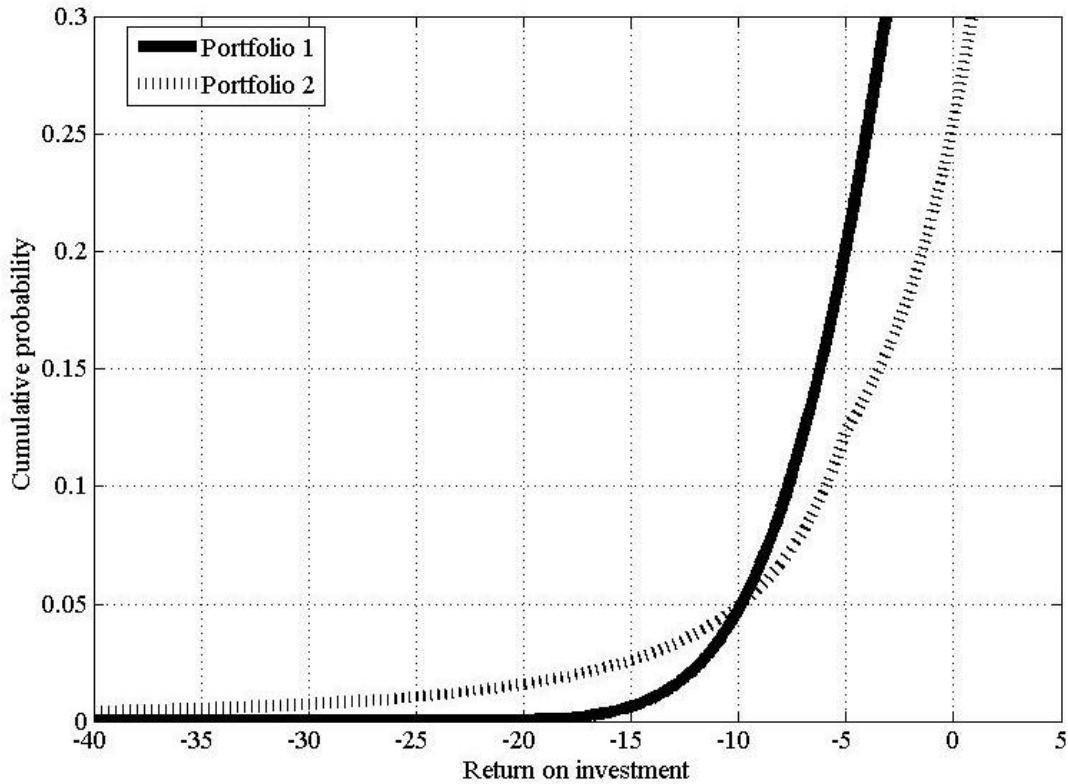


Fig. 1. Two portfolios have identical values at risk at $q = 0.05$: $VAR_{0.05} = 10$; but the probability of very extreme losses (e.g., less than -20) is larger for portfolio 2 than portfolio 1.

One example of using weighted averages of the different quantiles is conditional value at risk. Conditional value at risk calculates the expected value given that losses exceed or are equal to value at risk, and it better captures the extreme tails and encourages portfolio diversification.^(35,36) Although conditional value at risk was introduced and remains most popular as a financial risk measure, it has also been proposed to measure the risk of transporting hazardous materials,⁽³⁷⁾ industrial production under uncertain demand,⁽³⁸⁾ and wait times at hospitals.⁽³⁹⁾ This measure has been suggested as a useful tool for analyzing engineering risk as part of the partitioned multiobjective risk method which seeks to find Pareto optimal sets of alternatives based on different measures of risk.^(13,40)

Although measures based on a quantile of a distribution can describe the risk of low

probability, high-consequence events to a greater degree than mean or variance, they do not account for risks that are less extreme than the chosen quantile. Two distributions could be identical for all values less than the value corresponding to q but differ for less extreme values, and a risk measure based on quantiles would be the same for these two distributions. Losses that do not exceed value at risk may be of relatively little concern in finance and insurance because an organization will likely have enough money to cover these smaller losses. However, less extreme consequences may be important to consider for when making decisions about non-monetary risks. For example, military generals preparing for war may want to measure the risk of fatalities due to war, and a risk measure that fails to distinguish between fatalities less extreme than the 5% quantile may judge two strategies to be equally risky even though one strategy may lead to more fatalities at the 10%, 20%, and 50% quantiles. Using a quantile as a risk measure seems most suitable if it is used in concert with a measure representing the average or median risk—as an investor might do by maximizing the expected return of a portfolio subject to a value-at-risk constraint⁽⁴¹⁾—or if the analyst is primarily concerned with low-probability, high-consequence events and can ignore less extreme risk scenarios.

3.3. Disutility Functions

Because utility functions provide a solid foundation for making decisions under uncertainty, measuring risk with utility functions is a natural extension.^(42,43) In insurance, the zero-utility premium is a risk measure that describes the premium an individual or organization requires in order to insure a risky event. The expected utility of the premium and the uncertain risky event equals the utility without either the premium or event. Eq. (3) depicts two mathematical expressions for the zero-utility premium, depending on whether or not initial

wealth is included, where w is the initial wealth and $U(\cdot)$ is the utility function.^(17,18)

$$E[U(w + M(X) + X)] = U(w) \text{ or } E[U(M(X) + X)] = 0 \quad (3)$$

The risk measure $M(X)$ is the premium necessary to make an individual indifferent between insuring the risk and not insuring the risk.

Applying the zero-utility premium to non-monetary risks may pose difficulties because calculating a monetary risk premium requires assessing a tradeoff between money and potential consequences. Another utility-based risk measure uses disutility or loss functions. Disutility functions $V(\cdot)$ are the opposite of utility functions and describe the disutility an individual feels due to negative consequences. Mathematically, $V(x) = -U(x)$, and $V(x)$ increases as x decreases (i.e., as the consequences worsen). Similar to how expected utility can be deployed to choose between two alternatives, expected disutility, as given in Eq. (4), can serve as measure of risk, where $f_X(x)$ is a probability density function over the potential outcomes.

$$M(X) = E[V(X)] = \int_{-\infty}^{\infty} V(x)f_X(x)dx \quad (4)$$

If the risk measure is designed to only capture potential losses (downside risk), the upper limit of integration should be zero instead of positive infinity.

In actuarial science, disutility functions are typically called loss functions or utility-based shortfall risk if the loss function is convex with respect to X .⁽⁴⁴⁾ Utility-based shortfall risk measures provide a consistent measure of risk over time⁽⁴⁴⁾ and may be more sensitive to extreme risks and large losses than quantile-based measures.⁽³⁶⁾ Föllmer and Schied^(15,45) use a loss function to calculate the amount of capital necessary for an investor to accept a risky position X and extend the concept to develop robust risk measures for multiple scenarios.

Although many different disutility functions may be appropriate, this paper will focus on exponential disutility and the α - t model. The exponential disutility function is given in Eq. (5)

where γ reflects an individual's tolerance for risk.

$$\text{If risk averse, } V(x) = e^{-\gamma x} - 1, \gamma > 0$$

$$\text{If risk neutral, } V(x) = -x \tag{5}$$

$$\text{If risk seeking, } V(x) = 1 - e^{-\gamma x}, \gamma < 0$$

All of these functions are positive when $x < 0$, and the disutility function equals 0 if $x = 0$.

Another possible disutility function defines risk as a failure to meet a target outcome or objective. Fishburn⁽⁴⁶⁾ generalized this viewpoint in the α - t model, which is defined in Eq. (6) where t is a fixed upper bound so that only consequences below this upper bound are included in the risk measure, and α characterizes the risk attitude.

$$M(X) = \int_{-\infty}^t (t - x)^\alpha f_X(x) dx \tag{6}$$

The parameter $\alpha < 1$ implies risk seeking, $\alpha = 1$ implies risk neutrality, and $\alpha > 1$ implies risk aversion. A typical value for t is 0, which indicates that only outcomes where $x < 0$ will be included in the risk measure.

Both the exponential disutility and the α - t model can be transformed so that the units of the risk measure are in the same units as the consequences x . This can be accomplished by taking the inverse function of the disutility function: for exponential disutility, $V^{-1}(M(X))$ where $V^{-1}(\cdot)$ is the inverse function of Eq. (3) and $t - M(X)^{1/\alpha}$ for the α - t model. In insurance, the measure resulting from the inverse of expected disutility or loss is known as the mean value principle and enables the risk measure or premium to be expressed in monetary units.⁽¹⁷⁾

Estimating the risk attitude parameter poses the most difficult obstacle for using a risk measure based on disutility. If multiple individuals are going to use the risk measure, as is usually the case, determining a single risk attitude parameter that can be used for the entire group may be impossible. Even if everybody's risk attitude could be assessed, individuals have

different tolerances for risk,^(47,48) and one measure will likely not accurately reflect each individual's attitude toward risk.

3.4. Functions of Indicators or Factors

The fourth category of risk measures are those based on factors or indicators which are aggregated together. Several factors that are related to the risk of the event are aggregated together, often through a weighted linear combination of the factors.⁽⁴⁹⁾ Some risk measures or indices based on factors describe the risk of earthquakes,⁽⁵⁰⁾ hurricanes,⁽⁵¹⁾ climate change,⁽⁵²⁾ and terrorist activity.⁽⁵³⁾

One reason indices or measures based on aggregating different factors are popular is because these factors or indicators can often be measured or easily assessed from experts. When data are available, regression can be deployed to calculate the weighting parameters. Aggregating these factors through a functional form makes the measure appear objective, and the challenging task of estimating a probability distribution is usually not necessary.

As deterministic functions, these measures frequently fail to account for the uncertainty that is inherent in risk. Even when they incorporate uncertainty, they treat uncertainty or the frequency of the event as another indicator or factor rather than using a true probability distribution. For example, the WorldRiskIndex,⁽⁵⁴⁾ which has been developed to inform communities around the world about the risk of natural disasters, has four components: exposure, susceptibility, coping capacities, and adaptive capacities. Each component is expressed as a percentage between 0 and 100 percent, and exposure attempts to measure the percentage of the citizens in a country exposed to a natural disaster. An overall level of risk for a country is the product of the exposure indicator and a weighted sum of the percentages of the other three

components. Although exposure can be interpreted as the probability of a natural disaster impacting a random individual in a specified nation, none of the indicators quantify uncertainty in the consequences from a natural disaster. Multiplying exposure by the other indicators that are themselves weighted sums of other sub-categories has no true probabilistic meaning.

As another example of basing a risk measure on factors, failure mode and effect analysis is a risk analysis method used by industry to identify points of failure in a process or system. The risk measure, called the risk priority number, is the product of three factors: the frequency of a problem, the likelihood the problem will not be detected, and the severity of the problem. Each factor is measured on a ten-point ordinal scale without any translation to probabilities or measurable consequences such as cost or days behind schedule.^(55,56)

3.5. Risk Measures of More Than One Number

The risk measures thus far examined are measures composed of a single number. Even the measure in Eq. (1) based on two statistical moments (the mean and standard deviation) results in a single number. Instead of summarizing risk as a single number, a risk measure could be composed of two or three summary numbers, which conveys more information than a single number. A risk measure based on the moments of the distribution can include both the mean and standard deviation of the probability distribution. Investment science has traditionally used mean-variance analysis to evaluate and compare different investments. This measure conveys both the location and the spread of the distribution, and the closer the distribution is to a normal distribution, the more the mean and standard deviation fully describe the distribution.

In order to emphasize low-probability events without ignoring more likely scenarios, a three-number measure may be composed of different quantiles, like the 0.05, 0.50, and 0.95

quantiles. Such a description could communicate a fairly accurate picture of the probability distribution, and users could compare the risks of two different events on a quantile-by-quantile basis.⁽⁵⁷⁾ As mentioned earlier, the expected return of an investment portfolio and value at risk for that portfolio can be used to select the mix of assets for that portfolio.⁽⁴¹⁾ Jouini et al.⁽⁵⁸⁾ extend a class of popular financial risk measures, which includes conditional value at risk, to a multidimensional framework.

Because the major difficulty with a disutility-based measure is determining a risk attitude parameter appropriate for several individuals, a risk measure could be composed of multiple values, each of which corresponds to a different risk attitude. For example, one number could be expected disutility for a moderately risk seeking person, another number for a risk neutral person, and the final number for a moderately risk averse person. A person using this risk measure could use the measure that best corresponds to his or her risk preference.

One difficulty with using a risk measure composed of multiple numbers is that the measure does not instruct a user how to rank or compare risks. For example, if a risk measure is composed of the mean and standard deviation, anyone using the measure must still decide whether the mean or standard deviation is more important. If the measure should determine that one event carries more risk by weighting the importance of the standard deviation vis-à-vis the mean, the risk measure in Eq. (1) provides an appropriate option. Alternatively, an analyst could present the two- or three-number risk measure and realize that different people will make different conclusions about the riskiness of the event based on their own beliefs.

Despite the ability to communicate more information with multiple values, most risk measures appear to be single number. It is unclear as to why a single-number measure might be preferred: because analysts prefer to summarize their research into a single number or because

users of these measures want to interpret a single value. Risk analysts should use risk measures composed of multiple numbers more frequently than currently appears to be the case because these measures may better accomplish the purpose behind analyzing and communicating the risk.

4. FROM A RISK MEASURE TO A RISK INDEX

After choosing a risk measure, the analyst may want to transform the measure to a risk index. As this paper has defined a measure and index, a risk measure can be considered a risk index without any further calculations or operations. Because financial risk measures are usually in monetary units, rescaling these measures is usually unnecessary. Transforming the scale of non-financial risk measures may help facilitate understanding, especially for measures that are functions of indicators and factors because these functions often do not have any units like number of fatalities. Even a measure with real-world units may be difficult to interpret. For example, an expected value measure such as 220,000 barrels of oil might be difficult to interpret without context about the geographic area or the maximum number of barrels. If that measure is transformed to a number such as 80 on a scale from 0 to 100, it may be easier to understand that the number represents a serious risk. If a measure is a vector of numbers, each number in the measure could be mapped to a different scale in the index, and the risk index would also be composed of multiple numbers. Whether a measure is composed of a single or multiple numbers, the new scale can be composed of continuous real numbers (a numerical index) or discrete categories or numbers (a categorical index).

4.1. Numerical Index

A continuous numerical scale is appropriate if the risk index is going to be used to

compare different risks that are quite similar or to see how risk is changing over time because such a scale reveals subtle changes and differences. If decision makers are going to use the measurement or index as a basis for decisions, a numerical scale is preferable to a categorical scale because the numerical scale can provide greater detail and more levels of differentiation than a general category.

A 100-point or 10-point scale is probably the most natural type of scale for a numerical index, as humankind has relied on a decimal numerical system for millennia. A linear transformation of the risk measure to a new scale requires identifying an upper bound and lower bound for the risk measure— M_{UB} and M_{LB} —which correspond to 0 and 100 on the 100-point scale. These bounds may not be true maximum and minimum values but may be arbitrarily chosen as the maximum acceptable level of risk, M_{UB} , and the most desirable level of risk, M_{LB} , which is similar to assessing high and low values for single-attribute value functions in decision problems.⁽⁵⁹⁾ Eq. (7) demonstrates how the risk measure can be mapped to a 100-point risk index $I(X)$.

$$I(X) = \frac{M(X) - M_{LB}}{M_{UB} - M_{LB}} * 100 \quad (7)$$

Determining a value for M_{UB} often presents a greater challenge than selecting M_{LB} , which will usually equal 0. If the probability distribution underlying the risk measure is unbounded, the risk measure may not have any true maximum or upper bound. Instead, one could choose a value for M_{UB} that represents the minimum value of risk that is considered unacceptable. Some governmental agencies have developed thresholds for individual fatality risks⁽⁶⁰⁾ which can help guide the selection of M_{UB} .

Humans have a natural tendency to compare numbers using basic mathematical operations such as multiplication and addition. Consequently, a numerical risk index designed

for public consumption may raise questions about whether the intervals between numbers on the index have any meaning. For example, a risk index evaluating safety risks in buildings for the city of London uses a 100-point scale,⁽⁶¹⁾ and the study's authors warn its readers about misinterpreting the distance between numbers because the index is based on a logarithmic scale. Other papers use words such as “twice as risky” to compare risks with numerical measures or on a risk index.^(62,63,64) Risk priority numbers calculated using failure mode and effects analysis for several risks within an organization are sometimes added together to evaluate the overall risk for that organization even though adding risk priority numbers—which are based on ordinal scales—really has no mathematical justification.^(65,66) Hubbard⁽⁶⁾ criticizes organizations that arbitrarily add and multiply risk scores together to calculate an overall risk measure or index without understanding whether scores should be added or multiplied.

Because a numerical risk index could prompt people and organizations to add or multiply numbers in the index, understanding whether an index has certain mathematical properties (e.g., positive homogeneity, additivity, translation invariance) is a useful exercise. Financial risk measures are often evaluated based on whether these measures follow these and other axioms. In actuarial science, properties like positive homogeneity,⁽⁶⁾ translation invariance, and convexity explain the amount of capital required to cover a risk and encourage diversification in portfolios. Although these properties are less important for non-financial risk measures and indices, the tendency for people to add, subtract, and multiply numbers makes it desirable to investigate if a risk index has any of these properties. The users of the index should also be informed about the appropriateness of adding or multiplying numbers in the index and about the meaning, if any, of the numerical distance between values in the index.

Whether an index number twice as large as another number on the index means that the

former is twice as risky can be answered by exploring if the index is positive homogenous. An index with positive homogeneity means that if all the consequences worsen by a factor of $\lambda > 0$, the index number is multiplied by λ , as Eq. (8) shows.

$$I(\lambda X) = \lambda I(X) \quad (8)$$

Positive homogeneity implies that an event with an index value of 40 is twice as risky as an event with an index value of 20. If $M_{LB} = 0$ and the risk measure is positive homogeneous, the numerical index obeys this property. The numerical consequences that underlie the risk measure should also be measured on a ratio scale with a meaningful zero value.

An additive index implies that if two events are combined, the risk of these two events equals the sum of the risks of each event, as depicted in Eq. (9), where X and Y are random variables representing two events.

$$I(X + Y) = I(X) + I(Y) \quad (9)$$

For the risk index to be additive, M_{LB} must equal 0. The appealing feature of this property is that an organization can separately measure multiple risks and add the risks together to obtain a numerical measure of the organization's risk. When an organization considers multiple risks, it is preferable to model the joint distribution of $X + Y$ rather than modeling each distribution separately and adding the summary risk measures or index numbers together. However, the joint modeling exercise could pose difficulties especially if more than two risks are being considered, and using an additive risk measure or index may be an appropriate substitute.

Finally, a translation invariant risk measure means that if consequences improve by some real number β , the riskiness of the event decreases by that same amount, as given in Eq. (10).

$$M(X + \beta) = M(X) - \beta \quad (10)$$

A translation invariant risk measure means that the numerical difference between two measures

can be interpreted using the units that measure the consequences. If the risk measure is translation invariant, mapping the measure to a 100-point scale as shown in Eq. (5) implies the relationship in Eq. (11).

$$I(X) - I(X + \beta) = \frac{100\beta}{M_{UB} - M_{LB}} \quad (11)$$

The risk index is translation invariant if $M_{UB} - M_{LB} = 100$, which seems extremely unlikely as the risk index is usually based on a different scale than the consequences X . If the risk measure is translation invariant, the difference between two values on the risk index can still be used for comparisons among index values. For example, a difference of five between two scores means the same thing whether the scores are in the twenties or the eighties.

The mean value is the only risk measure considered that satisfies the positive homogenous, additive, and translation invariant properties. Value at risk, conditional value at risk, and the measure combining mean and standard deviation as in Eq. (1) are positive homogeneous and translation invariant⁽³⁶⁾ but not additive. Standard deviation and the α - t model are positive homogeneous but neither translation invariant nor additive. Exponential disutility is additive if X and Y are independent, and it is translation invariant but not positive homogeneous. (For any of these properties to hold for the two disutility models, the inverse function of the disutility model must be used.)

An analyst may determine that these properties are not important for the risk index because he or she does not foresee that the index will be interpreted in such a manner. In that case, whether or not the risk measure or index has any of these properties will not enter the decision on the type of measure to choose. However, if the analyst does believe the users will add or multiply numbers in the index, then these properties may influence the type of risk measure chosen. If separate risks are going to be combined into an overall risk measure, the

additive property may be necessary, which limits the risk measures to expected values or expected exponential disutility. Some measures, such as conditional value at risk and standard deviation are subadditive, which means that $M(X + Y) \leq M(X) + M(Y)$. This property ensures that the summation of two risk measures provides an upper bound or a worst-case risk measure if the two risks are combined. If risks are going to be compared against each other, a positive homogeneous measure might be better because saying something like one risk is twice as risky as another event has a clear mathematical meaning. If the analyst believes that all three properties are crucial for the index or measure, using the mean as the measure might be the best choice.

4.2. Categorical Index

If the analyst is uncomfortable about communicating specific numbers or is worried that these numbers might give an inaccurate impression, he or she might want to transform the measure to a categorical rather than a numerical risk index. A categorical risk index is most appropriate when the goal is to give people a general sense of the level of risk. Especially if a large number of people are going to be using a risk index, using categories can facilitate broader communication because categories can be understood more easily than a numerical scale.

A categorical risk index provides an ordering relationship among the categories, which can be composed of numbers, words, letters, or colors. Using words as categories like “severe” or “moderate” has an explicit communication purpose, and using letters implies a grading system in the United States. Another popular type of risk index is a color-coded chart. If a risk is classified in the “red” category, it sends a clear signal that the risk is very serious and may automatically trigger mitigating actions. If the risk measure is composed of multiple numbers, a

single risk may fall into multiple categories, one for each number in the risk measure. This may create confusion because people may not understand why the same risk overlaps multiple categories. If people are carefully informed about the meanings of each categorization, however, a multiple categorization may allow each individual to rank risks based on what he or she believes is more important. For example, a single risk may be classified in a less severe category for risk neutral or moderately risk averse individuals but may be catalogued in a more severe category for very risk averse individuals.

Regardless of the categorical scale chosen, the index should still rely on a careful analysis of the risks, and the level of a category should be based on a numerical risk measure. After determining the type of category, the analyst would choose the number of categories. Two objectives for determining how many categories are to minimize the number of categories and have enough categories to so that distinct risks are categorized differently.

The next step is to decide how to translate the numerical risk measure to a categorical scale. The simplest method is to divide the range of measures into equally divided intervals. Another method is to choose the categories based on the actual calculated risk measures. For example, 10 different events may have the following measures: 2.0, 2.5, 6.4, 7.0, 7.3, 9.2, 10.0, 11.0, 13.0, and 13.5. The analyst may want to put the first two events in the same risk category, the next three events in another category, the next three events in one or possibly two categories, and the final two events in another category. This categorization method may have problems if new risks are later added to the index or if new information arrives that changes the calculations for the risk measures.

Another method of creating categories is based on where a distinction between categories will most likely be noticed. Gustav Fechner, a German psychologist, postulated the intensity of a

sensation increases as the logarithm of the stimulus increases, and several risk scales—including the Richter scale for earthquakes and the Torino scale for near-Earth objects like asteroids—use a logarithmic scale.⁽⁶⁷⁾ Dividing the categories according to the base-10 logarithm of the risk measure (i.e., 10^{-2} , 10^{-1} , 0, 10) may follow most closely with how people perceive changing intensities.⁽⁶⁸⁾

One benefit of using a categorical scale is that recommended actions can be attached to each category. If the risk falls into a certain category, the public can be advised to take certain actions. For example, the National Weather Service issues tornado watches and warnings. Each advisory has a specific meaning and carries recommendations for people who receive a watch or a warning.⁽⁶⁹⁾ Because these advisories have been used for a long time and are issued frequently in areas where tornadoes appear, people understand what these advisories mean. A person subconsciously attaches his or her own risk attitude to these advisories when deciding whether to follow the recommended actions.

5. ILLUSTRATIVE EXAMPLE: TRANSPORTATION RISK

An illustrative example comparing the risk of fatal accidents between aviation and highway demonstrates the construction and interpretation of risk measures and indices. Although this example will lead to conclusions about the riskiness of aviation versus highway, the most important insights from this example relate to what judgments need to be made about risk measures and whether and how to translate the measure to a risk index.

Approximately 35,000 people die in transportation-related incidents in the United States each year, and the vast majority of fatalities result from accidents on the road.⁽⁷⁰⁾ Whether driving or flying is safer depends on the assumptions used to calculate the probability of a fatal

accident.^(71,72) Communicating the risk of fatalities of different transportation alternatives may help people make more informed choices, and policy makers are interested in reducing transportation-related fatalities.^(73,74,75,76) Risk measures and indices for aviation and highway risk can inform both the general public and policy makers.

Consequences of a highway or aviation incident can include fatalities, injured persons, and the monetary cost of accidents. Because our interest is to demonstrate how consequences can be incorporated into a risk measure and index as opposed to generating a complete picture of transportation risk, the consequences in this example are composed solely of fatalities. Analyzing the risk of fatalities from transportation could serve different purposes: (1) to help public officials determine which mode carries more risk or (2) inform an individual about the riskiness of flying or driving. As will be discussed, each of these purposes may inform the type of risk measure selected.

Aviation risk is divided into two categories based on U.S. Federal Aviation Regulations: air carriers or commercial aviation, and air taxi and commuters or on-demand and chartered flights. The data reveal a large difference in the risk of fatalities between these two categories of aviation. Most of the flying public use air carriers, and the risk of a fatal accident from an air carrier may be independent of the fatal risk from air taxis and commuters. The number of fatalities from each fatal accident in aviation and highway travel in the United States is recorded from 2005 through 2009 in Tables I – III. The data are derived from the Bureau of Transportation Statistics,^(77,78,79,80) the National Transportation Safety Board,⁽⁸¹⁾ and the National Highway Traffic Safety Administration.⁽⁸²⁾

Table I. Fatalities from air carrier incidents in the U.S., 2005-2009.

Fatalities per incident	Number of incidents
1	4
20	1
49	1
50	1
4.007x10 ¹⁰ air carrier miles traveled, 2005-2009	

Table II. Fatalities from air taxi and commuter incidents in the U.S., 2005-2009.

Fatalities per incident	Number of incidents
1	18
2	14
3	7
4	7
5	4
6	2
7	2
8	2
9	1
2.487x10 ⁸ air taxi and commuter miles traveled, 2005-2009	

Table III. Fatalities from highway incidents in the U.S., 2005-2009.

Fatalities per incident	Number of incidents
1	165,614
2	12,177
3	1,858
4	509
5	144
6	38
7	14
8	5
9	4
10	2
11	1
12	1
17	1
23	1
1.497x10 ¹³ highway miles traveled, 2005-2009	

We use the data to develop probability distributions for each of the three transportation modes. Normalizing the number of incidents, we assume that the number of fatalities given that

at least one fatality occurs follows a “shifted” geometric distribution: $P(y) = p * (1 - p)^{y-1}$ where $y = 1,2,3, \dots$ is the number of fatalities; $p = 1/\bar{y}$; and \bar{y} is the average number of fatalities given a fatal incident occurs. From the data in the tables, $\bar{y} = 17.6$ for air carriers, 2.9 for air taxi and commuters, and 1.1 for highway, which translate to $p = 0.057, 0.343,$ and 0.907 for the three transportation modes, respectively. Figure 2 depicts the cumulative probability distributions and the empirical probabilities from Tables I – III. The shifted geometric distribution provides a good fit for air taxi and commuter and highway, but the theoretical distribution for air carrier underestimates the probability of an incident with 1 fatality. Because the probability distribution for air carrier is based on only four data points, this distribution could change significantly if just one or two additional fatal accidents occurred.

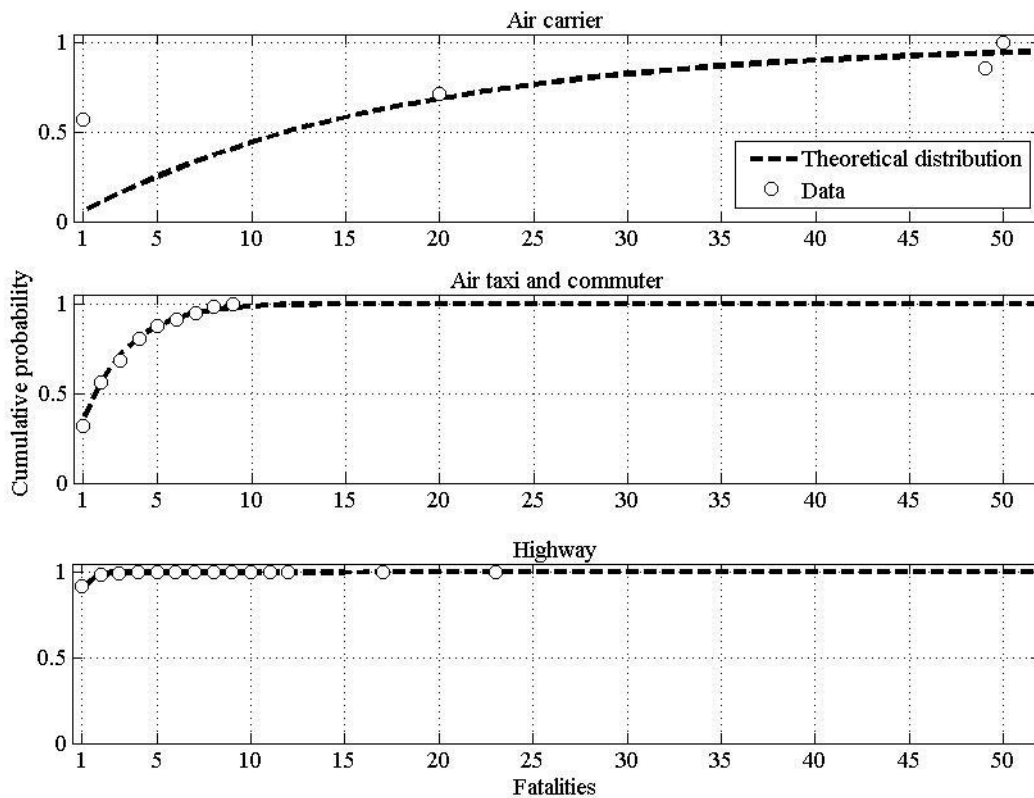


Fig. 2. Cumulative probability distributions for number of fatalities given a fatal accident occurs.

Because the distributions are calculated given a fatal incident occurs, we calculate the

probability of fatal incident by dividing the total number of fatal incidents by the total number of miles traveled in the United States. The probabilities of a fatal incident per 1 million miles traveled are calculated as 0.000175 for air carriers, 0.229 for air taxis and commuters, and 0.0121 for highway. The probability distribution of fatalities for a mode of transportation is the probability of a fatal incident multiplied by the shifted geometric distribution. We also rotate the distribution over the vertical axis so that fatalities are represented by negative numbers to correspond with how consequences were defined in Section 3.

The different formulas discussed in Section 3 are used to measure the risk of fatalities for the three different transportation modes using these probability distributions of fatalities. Table IV depicts the mean, standard deviation, value at risk for two different quantiles, and conditional value at risk.¹ Figure 3 displays risk measures using Eq. (1), which combines the mean and standard deviation into a single risk measure, for different values of k . Figures 4 and 5 plot risk measures for exponential disutility and the α - t model (where $t = 0$) for different risk attitudes. The risk measure for the exponential disutility function and the α - t model are transformed by taking the inverse of disutility functions so that the measure is in units of fatalities. If $\alpha < 1$, raising $M(X)$ to the power $1/\alpha$ results in an extremely small number. Consequently, the risk measure for the α - t model is not in units of fatalities for a risk seeking attitude.

Table IV. Summary risk measures for transportation

	Mean	Standard deviation	Value at risk		Conditional value at risk	
			$q = 0.05$	$q = 0.01$	$q = 0.05$	$q = 0.01$
Air carrier	0.0031	0.32	0	0	0.061	0.31
Air taxi and commuter	0.67	1.67	4	8	6.48	10.31
Highway	0.013	0.13	0	1	0.27	1.12

¹ Different calculations exist for conditional value at risk for discrete distributions. We use the expected shortfall definition given in Acerbi and Tasche⁽³⁵⁾ where $CVAR = -\frac{1}{q}(E[X|X \leq VAR] + VAR * (P[X \leq VAR] - q))$.

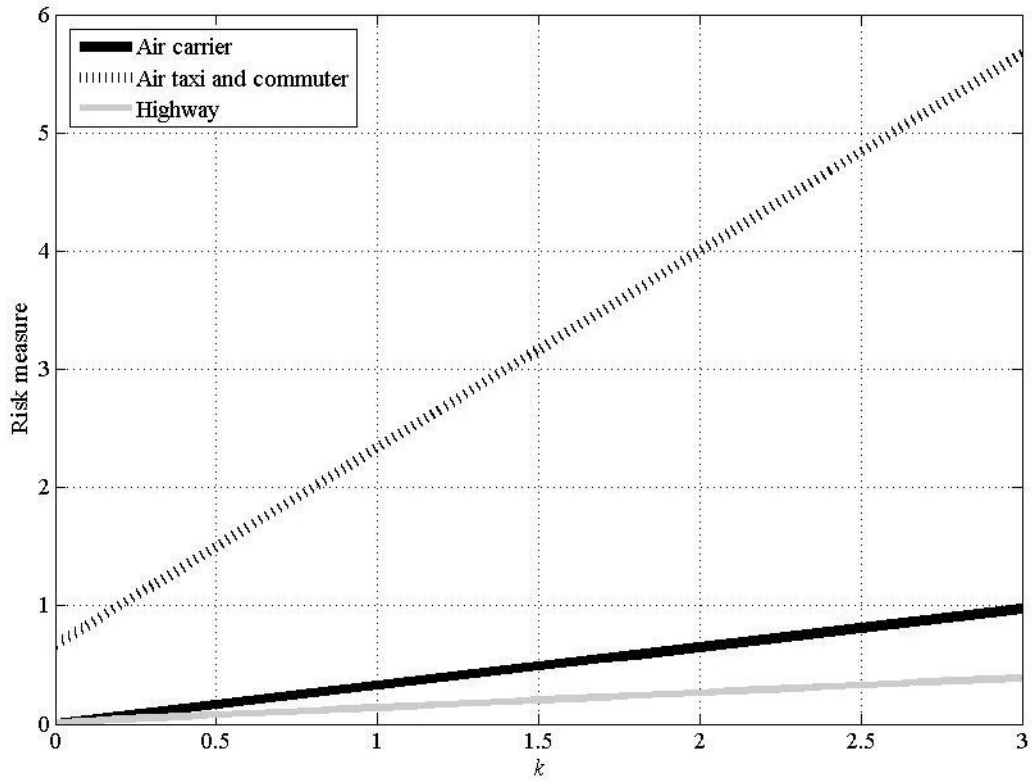


Fig. 3. Risk measure as a weighted sum between the mean and standard deviation as given in Eq. (1).

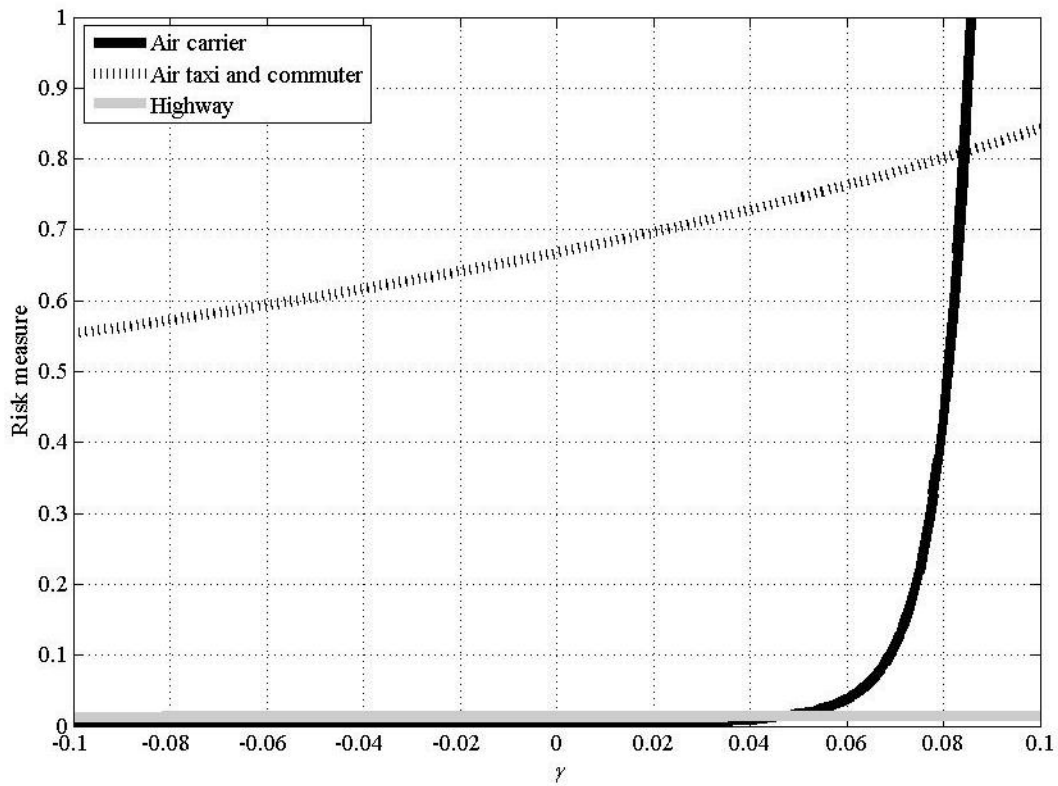


Fig. 4. Risk measure using an exponential disutility function.

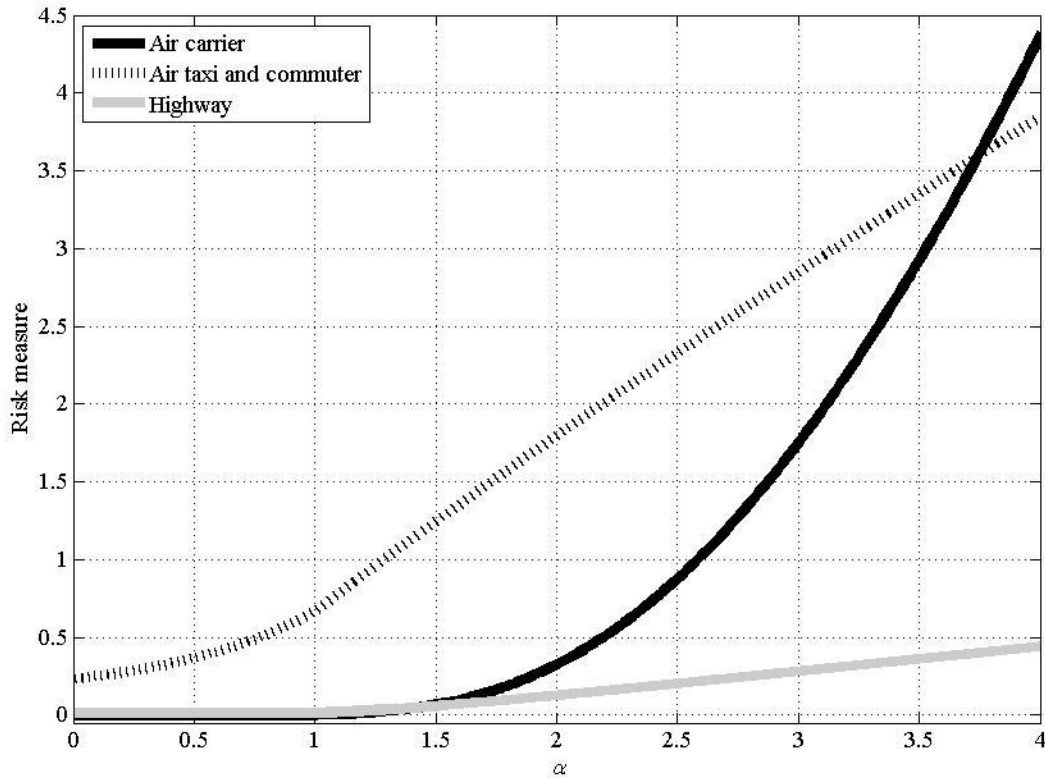


Fig. 5. Risk measure using the α - t model where $t = 0$.

Air taxi and commuter carries the largest risk of fatalities according to almost all of the risk measures under consideration. If the mean (or a risk neutral attitude) is used as a risk measure, highway transportation carries more risk than air carrier transportation. The mean number of deaths per 1 million miles traveled on the highway is 3.5 times greater than the mean number of deaths for air carrier transportation. Value at risk and conditional value at risk, both of which attempt to measure extreme values in the tails of the probability distributions, also produce risk measures in which highway carries more risk than air carriers. Because the probability of 4 or more fatalities is greater for air carriers than for highway, the former mode of transportation has a larger standard deviation and consequently, more variability. The standard deviations for both modes of transportation are much larger than their respective means, and putting a small amount of weight ($k \geq 0.1$) on the standard deviation in the measure in Eq. (1)

produces a risk measure where air carrier transportation is riskier than highway (Figure 3).

The risk of fatalities from highway transportation exceeds that from air carrier travel for moderately risk averse attitudes in the disutility functions (Figures 4 and 5), but as risk aversion increases, air carriers become riskier than highway. If $\gamma \geq 0.048$ in the exponential disutility or if $\alpha \geq 1.46$ in the α - t model, air carrier risk exceeds highway risk. The disutility risk measures for air carrier transportation are also larger than that for air taxi and commuters for very risk averse attitudes: $\gamma \geq 0.085$ in the exponential disutility or $\alpha \geq 3.75$ in the α - t model. The probability of a fatal incident for air carriers is less than the probability for both air taxis and commuters or highway, but the probability of a large number of fatalities given a fatal accident occurs for air carriers is greater than the other transportation modes. The larger values for γ and α emphasize crashes with scores of fatalities, which have occurred with air carriers.

Whether a user interested in transportation fatalities would exhibit risk aversion depends on the goal of the risk measure and the user's preferences. If the goal is to help inform policy makers about reducing transportation fatalities, any risk attitude may be justified. Debate exists within the decision and risk analysis literature over the appropriate risk attitude for fatalities in a public policy context.^(83,84,85) Many risk assessments of potential fatalities assume moderate risk aversion.^(86,87,88) People tend to prefer a situation with a large probability of a small number of fatalities to one with a smaller probability of a larger number of fatalities even if the expected number of fatalities is identical in both situations.^(89,90) However, Slovic *et al.*⁽⁸⁴⁾ question the risk aversion assumption due to an experiment in which people chose to reduce the number of single-fatality incidents rather than reduce the probability of a catastrophic incident with a large number of fatalities. Another study⁽⁹¹⁾ reveals that people may exhibit both risk seeking and risk averse attitudes, depending on how the situation is framed. Because incidents with a large number of

fatalities receive more media and public attention, a policy maker may want to focus on minimizing the number of accidents with many fatalities.

The risk measure may also be used to inform an individual whether he or she is more likely to die traveling in an airplane or in a motor vehicle, which suggests that the probability of a fatal incident is the best measure. The probability of a fatal incident per one million miles is 0.00018 for air carriers, 0.23 for air taxis and commuters, and 0.012 for highway transportation. *Ceteris paribus*, an individual should travel with air carriers to minimize his or her chances of being killed while traveling.

Translating the above risk measures to a different numerical scale may help with communication and give people a more intuitive sense of risk. A 100-point numerical scale is created where a value of 0 on the scale is equivalent to no fatalities ($M_{LB} = 0$) and a value of 100 is equivalent to exactly one fatality per one million miles ($M_{UB} = 1$). Table V depicts risk indices using this scale for the measures based on the mean values, probabilities of a fatal incident, exponential disutility where $\gamma = 0.05$, and α - t model where $\alpha = 1.25$. (Decimals are rounded to the nearest integer for these four indices.) The logarithmic index in the last column of Table V uses a different transformation and will be analyzed shortly.

Table V. Transportation risk index

	<i>100-point scale</i>				Logarithmic 10-point scale
Mean	Probability of fatal incident	Exponential disutility $\gamma = 0.05$	α - t model $\alpha = 1.25$		
Air carrier	0	0	2	2	7.5
Air taxi and commuter	67	23	74	96	9.8
Highway	1	1	1	3	8.1

The severity of the risk as communicated by the index differs according to the risk

measure. The index value of 96 based on the risk averse α - t model makes air taxi and commuter risk appear extremely severe, but the index value of 23 based on the probability of a fatal incident makes the risk appear much more acceptable. This example illustrates how the type of risk measure selected heavily influences the riskiness of an event as determined and communicated by an index. When compared with air taxi and commuter, the indices for air carrier and highway are much less severe.

The logarithmic index is the most unique index depicted in Table V. This index is calculated by adding 10 to the base-10 logarithm of the expected number of fatalities. Thus, one fatality per one million miles corresponds to an index value of 10, and 10^{-10} expected fatalities corresponds to 0. (If 10^{-10} expected fatalities is too small to serve as an effective lower bound for the index, the base of the logarithm can be changed to reflect a different lower bound.) Taking the logarithm of the expected values as opposed to linear scaling—as in the other indices—compresses the differences between the expected number of fatalities. The risks of fatalities from each of the three modes of transportation appear more similar with a logarithmic index, and this index may more accurately reflect how people naturally perceive the differences among these risks.⁽⁹²⁾

The majority of risk indices and risk measures demonstrate that air carrier transportation is the least risky with respect to fatalities and that air taxis and commuters carry the most risk. As risk aversion increases, air carrier transportation carries a greater risk than highway transportation and even becomes riskier than air taxis and commuters. The risk averse disutility functions emphasize incidents with a large number fatalities, which is more probable with air carrier transportation.

These conclusions about the riskiness of each mode of transportation depend on the

probability of fatalities for each mode. Because the probability of a fatal incident for air carrier transportation is so small (0.000175), only seven fatal incidents for air carriers occurred between 2005 and 2009. If more fatal incidents occur, the probability distribution given a fatal incident in air carrier transportation would change, but the risk measure or risk index may not be much different as long as the probability of a fatal incident remains small. Epistemic uncertainty may also exist which may require more sophisticated modeling. For example, the likelihood of fatal incidents increases in severe weather, which changes the risk for all three modes and potentially the ordering of these risks. If the goal of the analysis and risk index is to inform actions during severe weather, the probabilities should perhaps be assessed using data under severe weather conditions, by conducting a probabilistic analysis of possible initiating events leading to fatal incidents, and/or through simulation techniques.

Analysts and decision makers could (and should!) make identical conclusions about the riskiness of each transportation mode by studying the fatality data in Tables I – III, and the most complete way to incorporate risk into a decision is to examine probability distributions. However, many people are not comfortable with probability distributions and prefer a single numerical risk measure or a categorization of risk rather than studying a table of raw data. This example demonstrates how the guidelines in Sections 2 and 3 can be applied to translate raw data to a summary risk measure and index. If the analyst needs to summarize the data or probability distribution, his or her choices about the type of measure (e.g., mean, standard deviation, quantile, disutility functions) and parameters, such as the risk averse parameter in disutility functions or the value of M_{UB} , can impact how risky an event is and the ranking of different risks. If a decision maker wants a single numerical measure describing the risk of transportation fatalities or if the government wants to communicate this risk to the public with a 100-point scale

risk index, exploring the issues with these summary measures and how choices of an index can change the interpretation and ranking of risks should be part of the process.

The index should be selected based on which index best meets the goal of the risk assessment. If the goal is to inform the public about the risk of each mode of transportation, the risk index based on the mean or probability of a fatal incident or the logarithmic index may be most appropriate. An index based on either of the disutility functions or on conditional value at risk may best communicate how risk is changing over time to a policy maker that wants to emphasize incidents with a large number of fatalities. One of the risk measures from Table IV (maybe the mean or a measure composed of two or three values) can help a decision maker identify if more safety measures are needed for one of the transportation modes; however, that decision should also incorporate the cost and effectiveness of the proposed safety measure.⁽²⁾

6. CONCLUSION

This paper has outlined steps to create a risk index with the goal of unifying the construction of risk indices and in order to correct flawed methods such as those resulting in risk matrices. Several issues have been raised for analysts and decision makers to consider when thinking about the type of risk index to use. A risk index should be based on a numerical analysis of the risk using probabilities and consequences that is summarized by a risk measure. Risk measures can generally be classified into one of four categories: moments, quantiles, disutility functions, and weighted indicators or factors. Risk measures can also be composed of two or three values, and decision makers who want to rely on risk measures should consider a measure composed of multiple numbers. As demonstrated by the transportation example, the choice of risk measure can impact the riskiness of an event and the ordering of multiple risks.

Once the risk measure has been calculated, the analyst can map that measure to a numerical or categorical index, or the measure itself can serve as a risk index. Recall the list of objectives for risk indices outlined in the introduction. The most important goals for a situation should determine the type and structure of the risk index. If the most important goal is to describe the risk of an event in the most accurate way possible, the analyst should use the probability distribution or rely on a measure with multiple numbers. If a single number is desired, using a disutility function that incorporates the decision maker's preferences to the greatest possible extent is best. A numerical measure is also most appropriate when the goal is to assess how a mitigating action impacts the risk. The measure can provide a metric to judge the effectiveness of different mitigation strategies if it is known how these mitigation strategies impact the probability distribution. In this case, the risk is first measured under the assumption of no mitigation strategy and then assessed assuming the mitigation strategy was enacted. The difference between the two measures describes the benefit of the mitigation strategy.

A numerical risk index can best achieve the objectives of comparing between different risks, determining the most serious risk, and understanding how a risk changes over time. A numerical risk index can enable a user to clearly see which risk is most serious and how close the risk is to the maximum level of risk, although determining the upper bound for risk can pose serious challenges.

A categorical scale can be most useful when the primary goal is to communicate risk to a large group of people and recommend actions if the risk falls into a certain category. Simple and clear messages can be most effective, especially in moments of crises, and a categorical scale that uses colors or words may be the clearest communication tool. If more than one goal is important as is usually the case, the analyst or policy maker should choose the index that best

meets those goals and the situation.

The audience must also be considered when choosing the type of risk measure or index. If the audience is limited to a group of people who understand probability distributions, the analyst may forgo a risk measure and present the group with a probability distribution over the consequences. If the audience is composed of people who regularly work with numbers but who do not have the expertise or the time to work with probability distributions, a risk measure composed of multiple numbers may be the most appropriate. If the audience is an entire nation, a categorical index that is simple but informative may be ideal.

In all of these cases, an explanation should accompany the risk index. The analyst needs to explain what the number or category means, what it means if the risk increases or decreases, what it means if two risks have different risk numbers or are categorized differently, and the limitations of the risk index or measure. It takes time for people to trust and learn how to react to a new risk index. If the categories are specific and people understand why a risk is categorized at a given level, the index can provide useful information about the risk of an event to the users.

REFERENCES

-
1. American Society of Civil Engineers. America's infrastructure report card. Reston, VA, March 2013. Available at: <http://www.infrastructurereportcard.org/>, Accessed on February 1, 2014.
 2. Cox LA Jr. Evaluating and improving risk formulas for allocating limited budgets to expensive risk-reduction opportunities. *Risk Analysis*, 2012; 32(7):1244-1252.
 3. Hand DJ. *Measurement Theory and Practice: The World through Quantification*. New York: Oxford University Press, 2004.

-
4. Cox LA Jr. What's wrong with risk matrices? *Risk Analysis*. 2008; 28(2):497-512.
 5. Cox LA Jr. Some limitations of "risk=threat x vulnerability x consequence" for risk analysis of terrorist attacks. *Risk Analysis*. 2008; 28(6):1749-61.
 6. Hubbard DW. *The Failure of Risk Management: Why It's Broken and How to Fix It*. Hoboken, NJ: John Wiley & Sons, 2009.
 7. Kaplan S, Garrick BJ. On the quantitative definition of risk. *Risk Analysis*, 1981; 1(1):11-27.
 8. Slovic P. Perception of risk. *Science*. 1987; 237(4799):280-285.
 9. Fisher A. Risk communication challenges. *Risk Analysis*. 1991; 11(2):173-179.
 10. Paté-Cornell ME. Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering and System Safety*, 1996; 54(2-3):95-111.
 11. Bedford T, Cooke RM. *Probabilistic Risk Analysis: Foundations and Methods*. Cambridge: Cambridge University Press, 2001.
 12. Cox LA Jr. *Risk Analysis of Complex and Uncertain Systems*, International Series in Operations Research and Management Science. New York: Springer, 2009.
 13. Haimes YY. *Risk Modeling, Assessment, and Management*. 3rd ed. Hoboken NJ: John Wiley & Sons, 2009.
 14. Artzner P, Delbaen F, Eber J-M, Heath D. Coherent measures of risk. *Mathematical Finance*, 1999; 9(3):203-228.
 15. Föllmer, H, Schied A. Convex measures of risk and trading constraints. *Finance and Stochastics*, 2002; 6:429-447.
 16. Acerbi C. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*. 2002; 26:1505-1518.

-
17. Goovaerts MJ, de Vylder F, Haezendonck J. Insurance Premiums: Theory and Applications. Amsterdam: Elsevier Science, 1984.
 18. Denuit M, Dhaene J, Goovaerts M, Kaas R. Actuarial Theory for Dependent Risks: Measures, Orders and Models. Chichester, England: John Wiley & Sons, 2005.
 19. Tsanaks A, Desli E. Measurement and Pricing of Risk in Insurance Markets. Risk Analysis. 2005; 25(6):1653-1668.
 20. Luce RD. Several possible measures of risk. Theory and Decision. 1980; 12:217-228.
 21. Fishburn PC. Foundations of measurement. II. Effects of gains on risk. Journal of Mathematical Psychology. 1982; 25:226-242.
 22. Fishburn, PC. Foundations of risk measurement. I. Risk as probable loss. Management Science, 1984; 30(4):396-406.
 23. Luce RD, Weber EU. An axiomatic theory of conjoint, expected risk. Journal of Mathematical Psychology. 1986; 30:188-205.
 24. Albrecht P. Risk measures. In: Sundt B, Jozef T, editors. Encyclopedia of Actuarial Science. Hoboken, NJ: John Wiley & Sons; 2004. p. 1493-1501.
 25. Dowd K, Blake D. After VaR: The theory, estimation, and insurance applications of quantile-based risk measures. The Journal of Risk and Insurance. 2006; 73(2):193-229.
 26. Paté-Cornell ME. Fire risks in oil refineries: Economic analysis of camera monitoring. Risk Analysis, 1985; 5(4):277-288.
 27. Willis HH. Guiding resource allocations based on terrorism risk. Risk Analysis, 2007; 27(3):597-606.
 28. Cox LA Jr. Why risk is not variance: An expository note. Risk Analysis, 2008; 28(4):925-928.

-
29. Beirlant J, Vynckier P, Teugels JL. Tail index estimation, Pareto quantile plots, and regression diagnostics. *Journal of the American Statistical Association*. 1996; 91(436):1659-67.
 30. Markowitz H. Portfolio selection. *The Journal of Finance*. 1952; 7(1):77-91.
 31. Buhlmann H. Premium calculation from top down. *Astin Bulletin*. 1985; 13(2):89-101.
 32. Deprez O, Gerber HU. On convex principles of premium calculation. *Insurance: Mathematics and Economics*. 1985; 4:179-189.
 33. Vrijling JK, van Hengel W, Houben, RJ. A framework for risk evaluation. *Journal of Hazardous Materials*, 1995; 43(3):245-261.
 34. Ellsberg D. Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*. 1961; 75(4):643-669.
 35. Acerbi C, Tasche D. Expected shortfall: A natural coherent alternative to value at risk. *Economic Notes*, 2002; 31(2):379-388.
 36. Giesecke K, Schmidt T, Weber S. Measuring the risk of large losses. *Journal of Investment Management*, 2008; 6(4):1-15.
 37. Toumazis I, Kwon C, Batta R. Value-at-risk and conditional value-at-risk minimization for hazardous materials routing. In: Batta R, Kwon C, editors. *Handbook of OR/MS models in hazardous materials transportation*. Springer; forthcoming.
 38. Verderame PM, Floudas C. Operational planning of large-scale industrial batch plants under demand due date and amount uncertainty: II. Conditional value-at-risk framework. *Industrial & Engineering Chemistry Research*, 2010; 49(1):260-275.

-
39. Dehlendorff C, Kulahci M, Merser S, Andersen KK. Conditional value at risk as a measure for waiting time in simulations of hospital units. *Quality Technology & Quantitative Management*, 2010; 7(3):321-326.
 40. Asbeck EL, Haimes YY. The partitioned multiobjective risk method (PMRM). *Large Scale Systems* 1984; 6:13-38.
 41. Tasche D. Risk contributions and performance measurement. Working paper, Munich University of Technology, 2000. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.68.9393&rep=rep1&type=pdf>, Accessed on September 20, 2013.
 42. Bell DE. Risk, return, and utility. *Management Science*, 1995; 41(1):23-30.
 43. Jia J, Dyer JS. A standard measure of risk and risk-value models. *Management Science*, 1996; 42(12):1691-1705.
 44. Weber S. Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, 2006; 16(2):419-441.
 45. Föllmer, H, Schied A. Robust preferences and convex measures of risk. In: Sandmann K, Schönbucher PJ, editors. *Advances in finance and stochastics: essays in honour of Dieter Sondermann*. New York: Springer-Verlag; 2002. p. 39-56.
 46. Fishburn, PC. Mean-risk analysis with risk associated with below-target returns. *The American Economic Review*, 1977; 67(2):116-126.
 47. Donkers B, Melenberg B, van Soest A. Estimating risk attitudes using lotteries: A large sample approach. *Journal of Risk and Uncertainty*, 2003; 22(2):165-195.
 48. Hallahan TA, Faff RW, McKenzie MD. An empirical investigation of personal financial risk tolerance. *Financial Services Review*, 2004; 13:57-78.

-
49. Fischhoff B, Watson SR, Hope C. Defining risk. *Policy Sciences*, 1984; 17:123-139.
 50. Davidson RA. An urban earthquake disaster risk index [dissertation] Stanford, CA: Stanford University; 1997. 280 p.
 51. Davidson RA, Lambert KB. Comparing the hurricane disaster risk of U.S. coastal cities. *Natural Hazards Review*, 2001; 2(3):132-142.
 52. Anemüller S, Monreal S, Bals C. *Global Climate Risk Index 2006: Weather-Related Loss Events and Their Impacts on Countries in 2004 and in a Long-Term Comparison*. Bonn: Germanwatch e.V., 2006. Available at: www.germanwatch.org/klak/cr06.pdf, Accessed on September 25, 2012.
 53. Piegorsch WW, Cutter SL, Hardisty F. Benchmark analysis for quantifying urban vulnerability to terrorist incidents. *Risk Analysis*, 2007; 27(6):1411-1425.
 54. Beck MW, Shepard CC, Birkmann J, Rhyner J, Welle T, Witting M, Wofertz J, Marens J, Maurer K, Mucke P, Radtke K. *WorldRiskReport 2012*. Berlin: Bündnis Entwicklung Hilft Alliance Development Works, 2012. Available at: <http://www.ehs.unu.edu/article/read/worldriskreport-2012>, Accessed on February 7, 2014.
 55. Gilchrist W. Modelling failure modes and effects analysis. *The International Journal of Quality & Reliability Management*, 1993; 10(5):16-23.
 56. Ben-Daya M, Raouf A. A revised failure mode and effects analysis model. *The International Journal of Quality & Reliability Management*, 1996; 13(1):43-47.
 57. Long J, Fischhoff B. Setting risk priorities: A formal model. *Risk Analysis*, 2000; 20(3):339-351.
 58. Jouini E, Meddeb M, Touzi N. Vector-valued coherent risk measures. *Finance and Stochastics*, 2004; 8:531-552.

-
59. Kirkwood CW. Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets. Belmont, CA: Brooks/Cole; 1997.
 60. Paté-Cornell ME. Quantitative safety goals for risk management of industrial facilities. *Structural Safety*, 1994; 13(3):145-157.
 61. Department for Communities and Local Government. Review of Health and Safety Risk Drivers: BD 2518, London, 2008. Available at: <http://webarchive.nationalarchives.gov.uk/20120919132719/www.communities.gov.uk/publications/planningandbuilding/reviewhealthsafety>, Accessed on September 20, 2013.
 62. Markese J. Treasury bond funds: A risky riskless investment? *The AAIJ Journal*. 1998:January. Available at: <http://www.aaij.com/journal/article/treasury-bond-funds-a-risky-riskless-investment->, Accessed on September 20, 2013.
 63. Härmä M, Sallinen M, Ranta R, Mutanen P, Müller K. The effect of an irregular shift system on sleepiness at work in train drivers and railway traffic controllers. *Journal of Sleep Research*. 2002; 11:141-151.
 64. The International Bank for Reconstruction and Development. *Natural Hazards, UnNatural Disasters: The Economics of Effective Prevention*. Washington, DC: The World Bank, 2010. Available at: <https://www.gfdrr.org/NHUD-home>, Accessed on September 20, 2013.
 65. Bonfant G, Belfanti P, Paternoster G, Gabrielli D, Gaiter AM, Manes M, Molina A, Pellu V, Ponzetti C, Farino M, Nebiolo PE. Clinical risk analysis with failure mode and effect analysis (FMEA) model in a dialysis unit. *Journal of Nephrology*. 2010; 23(1):111-118.
 66. Institute for Healthcare Improvement. *Failure modes and effects analysis tool*, 2013. Available at: <http://app.ihp.org/Workspace/tools/fmea/>, Accessed on September 20, 2013.

-
67. Plum MM. Rational and development of a scale to communicate environmental and other community risks. In: Linkov I, Ferguson E, Magar VS, editors. Real-time and deliberative decision making, Dordrecht, The Netherlands: Springer; 2009. p. 103-126.
 68. Paulos JA. Innumeracy: Mathematical Illiteracy and Its Consequences. New York: Hill and Wang, 1988.
 69. Novy, CH, Edwards R, Imy D, Goss S. SPC and Its Products, March 2010. Available at: [http://www.spc.noaa.gov/misc/about.html#Severe Weather Watches](http://www.spc.noaa.gov/misc/about.html#Severe%20Weather%20Watches), Accessed on September 25, 2012.
 70. Bureau of Transportation Statistics (US), Research and Innovative Technology Administration. Table 2-1: Transportation fatalities by mode. In: National transportation statistics, Chapter 2 - transportation safety, section A – multimodal, April 2012. Available at: http://www.bts.gov/publications/national_transportation_statistics/html/table_02_01.html, Accessed on September 25, 2012.
 71. Evans L, Frick MC, Schwing RC. Is it safer to fly or drive? Risk Analysis, 1990; 10(2):239-246.
 72. Sivak M, Weintraub DJ, Flannagan M. Nonstop flying is safer than driving. Risk Analysis, 1991; 11(1):145-148.
 73. McCarthy PS. Public policy and highway safety: A city-wide perspective. Regional Science and Urban Economics, 1999; 29:231-244.
 74. Horn MET, Fulton N, Westcott M. Measures of societal risk and their potential use in civil aviation. Risk Analysis, 2008; 28(6):1711-1726.
 75. Li G, Baker SP. Crash risk in general aviation. JAMA, 2007; 297(14):1596-1598.

-
76. Bhalla K, Ezzati M, Mahal A, Salomon J, Reich M. A risk-based method for modeling traffic fatalities. *Risk Analysis*, 2007; 27 (1):125-136.
 77. Bureau of Transportation Statistics (US), Research and Innovative Technology Administration. Table 2-9: U.S. air carrier safety data. In: National transportation statistics, Chapter 2 - transportation safety, section B – air, July 2011. Available at: http://www.bts.gov/publications/national_transportation_statistics/html/table_02_09.html, Accessed on June 27, 2012.
 78. Bureau of Transportation Statistics (US), Research and Innovative Technology Administration. Table 2-10: U.S. commuter air carrier safety data.” In: National transportation statistics, Chapter 2 - transportation safety, section B – air, July 2011. Available at: http://www.bts.gov/publications/national_transportation_statistics/html/table_02_10.html, Accessed on June 27, 2012.
 79. Bureau of Transportation Statistics (US), Research and Innovative Technology Administration. Table 2-13: U.S. on-demand air taxi safety data.” In: National transportation statistics, Chapter 2 - transportation safety, section B – air, October 2011. Available at: http://www.bts.gov/publications/national_transportation_statistics/html/table_02_13.html, Accessed on June 27, 2012.
 80. Bureau of Transportation Statistics (US), Research and Innovative Technology Administration. Table 2-18: Motor vehicle fatalities, vehicle-miles, and associated rates by highway functional system.” In: National transportation statistics, Chapter 2 - transportation safety, section C – highway, July 2012. Available at:

http://www.bts.gov/publications/national_transportation_statistics/html/table_02_18.html,
Accessed on September 25, 2012.

81. National Transportation Safety Board (US). Aviation: Accident Database & Synopses, 2012. Available at: <http://www.nts.gov/aviationquery/index.aspx>, Accessed on June 28, 2012.
82. National Highway Traffic Safety Administration (US). Fatality Analysis Reporting System (FARS) Encyclopedia. NCSA Data Resource Website, 2012. Available at: <http://www-fars.nhtsa.dot.gov/Main/index.aspx>, Accessed on June 25, 2012.
83. Fischhoff B, Lichtenstein S, Slovic P, Derby SL, Keeney RL. Acceptable Risk. Cambridge: Cambridge University Press; 1981.
84. Slovic P, Lichtenstein S, Fischhoff B. Modeling the societal impact of fatal accidents. *Management Science*, 1984; 30(4):464-474.
85. Keller LR, Wang Y, Feng T. Measures of risk equity. In: JJ Cochran, editor. *Wiley encyclopedia of operations research and management science*. Hoboken, NJ: John Wiley & Sons; 2011.
86. Griesmeyer JM, Okrent D. Risk management and decision rules for light water reactors. *Risk Analysis*, 1981; 1(2):121-136.
87. Vatn J. A discussion of the acceptable risk problem. *Reliability Engineering and System Safety*, 1998; 61:11-19.
88. Vrijling JK, van Hengel W, Houben RJ. Acceptable risk as a basis for design. *Reliability Engineering and System Safety*, 1998; 59(1):141-150.
89. Ferreira J Jr, Slesin L. Observations on the social impact of large accidents. Massachusetts Institute of Technology; October 1976. 46 p. Technical Report No. 122.

-
90. Keeney RL. Equity and public risk. *Operations Research*, 1980; 28(3):527-534.
 91. Fischhoff B, Slovic P, Lichtenstein S. Knowing what you want: Measuring labile values.
In: Wallsten TS, editor. *Cognitive processes in choice and decision behavior*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1980. p. 117-141.
 92. Calman KC, Royston GHD. Risk language and dialects. *British Medical Journal*, 1997; 315:939-942.