An Economic Index of Riskiness

Robert J. Aumann† Roberto Serrano‡

October 2006

Abstract

Define the riskiness of a gamble as the reciprocal of the absolute risk aversion (ARA) of an individual with constant ARA who is indifferent between taking and not taking that gamble. We characterize this index by axioms, chief among them a “duality” axiom which, roughly speaking, asserts that less risk-averse individuals accept riskier gambles. The index is homogeneous of degree 1, monotonic with respect to first and second order stochastic dominance, and for gambles with normal distributions, is half of variance/mean. Examples are calculated, additional properties derived, and the index is compared with others in the literature.

JEL classification: C00, C43, D00, D80, D81, E44, G00.

Keywords: riskiness; risk aversion; expected utility; decision making under uncertainty; portfolio choice; Sharpe ratio; variance-mean ratio; value at risk.

1 Introduction

On March 21, 2004, an article on the front page of the New York Times presented a picture of allegedly questionable practices in some state-run pension funds. Among the allegations were that these funds often make unduly risky investments, recommended by consultants who are interested parties. The concept of “risky investment” is commonplace in financial discussions, and seems to have clear conceptual content. But when one thinks about it carefully and tries to pin it down, it is elusive. Can one give a clear, precise definition of riskiness, one that is independent of the person or entity making the investment?

Conceptually, whether or not a person takes a gamble depends on two distinct considerations:

*We thank Kenneth Arrow, Zvi Artstein, Dean Foster, Sergiu Hart, Chaim Levy, Petros Millimis and Ivo Welch for comments, for pointing out useful references, and for research assistance. The index proposed here first surfaced during a visit of Aumann to Brown University in March of 2004, and was applied in a working paper generated later that spring (Palacios-Huerta, Serrano and Volij, 2004). At that time, though, it still had somewhat of an ad hoc flavor; inter alia, there was as yet no axiomatic characterization.
†Center for Rationality, Hebrew University of Jerusalem; raumann@math.huji.ac.il
‡Department of Economics, Brown University; roberto_serrano@brown.edu
(i) the attributes of the gamble, and in particular, how risky it is; and
(ii) the attributes of the person, and in particular, how averse he is to risk.

The famous contributions of Arrow (1965, 1971) and Pratt (1964) address item (ii) by defining absolute and relative risk aversion, which are personal, subjective concepts, depending on the utility function of the person in question. But they do not define riskiness; they do not address item (i). It is like speaking about subjective time perception (“this movie was too long”) without having an objective measure of time (“it was two hours long”), or talking about heat or cold aversion (“it’s too cold in here”) without having in mind an objective concept of temperature (“it is 20 degrees”).

This paper addresses item (i); it develops a measure of riskiness of gambles. The concept is based on that of risk aversion: We think of riskiness as a kind of “dual” to risk aversion—specifically, as that aspect of a gamble to which a risk-averter is averse. So on the whole, we expect individuals who are less risk averse to take riskier gambles.

The gambles treated here yield gains or losses, measured in stated dollar amounts, with stated probabilities. Needless to say, many real-life gambles are not of that kind. For one thing, the outcomes may be non-monetary; getting married (or divorced), adopting a child, quitting one job for another, choosing a ski resort, deciding on an operation—all are gambles that do not fit the current framework easily, or indeed at all. Even gambles that are monetary are often not that well defined, numerically speaking. When we invest in stocks or bonds, we have at best a rough concept of the probabilities involved; likewise for many forms of insurance (such as those in which one’s own behavior is an important parameter, like in automobile accident insurance).

But, to start with, it is important to define riskiness in principle. The real-life problems we have mentioned are akin to measurement problems in physics. That there exists a physically precise definition of temperature does not imply that one can always tell just how hot or cold it is in a given place at a given moment. But before one can even ask that question, one does need to have the definition. Similarly, the main purpose of the research described here is to define riskiness; once one has the definition, one can address the problem of determining—measuring—riskiness in the applications.

One final point, essential to understanding our index, is in order: Riskiness is not the opposite of desirability; a less risky gamble need not be more desirable. The two concepts are “orthogonal.” Desirability is subjective, depending on the individual; one individual may prefer gamble g to gamble h, while another prefers h to g. Riskiness, on the other hand, is objective; it is the same for all individuals. Given two gambles, a more risk-averse individual may well prefer the less risky gamble, while a less risk-averse individual may actually prefer the riskier gamble.

The riskiness index proposed here is not the first; others have been proposed, in disciplines such as finance, statistics and psychology. We discuss some of those in Section 7, and compare them with the one proposed here. Basically, what sets our index apart from these others is that ours is based on economic, decision-
theoretic considerations, such as the duality principle roughly enunciated above. An even more basic requirement of this nature is monotonicity with respect to (w.r.t.) first-order stochastic dominance: if the outcome of a gamble $g$ is sure to be no better than that of $h$, and is with positive probability actually worse, or more generally, if each loss or its probability is higher under $g$ than under $h$, and each gain or its probability is lower under $g$ than under $h$, then $g$ should be riskier than $h$. Our index satisfies this elementary requirement, but surprisingly, very few in the literature do.

The plan of the paper is as follows: Section 2 is devoted to the basic axiomatic definition of the index, and its numerical characterization. Section 3 discusses the index conceptually, and relates it to the Arrow-Pratt coefficient of risk-aversion. In Section 4, the index is characterized in terms of constant absolute risk aversion, as outlined in the abstract. Section 5 sets forth some of the basic properties of the index, including its dimension (dollars), its monotonicity w.r.t. first- and second-order stochastic dominance, its continuity, its behavior for “diluted” gambles, normal gambles and sums of independent gambles, and its ordinal characterization. Section 6 provides numerical examples, meant to give the reader the beginnings of a quantitative “feel” for the index. Section 7 discusses the literature, and Section 8 is devoted to proofs. Section 9 concludes.

2 Axiomatic Characterization

In this paper, a utility function is a Bernoulli utility function for money, strictly monotonic, concave,\(^1\) twice continuously differentiable, and defined over the entire real line. A gamble $g$ is a random variable with real values—interpreted as dollar amounts—some of which are negative, and that has positive expectation. Let agent $i$ have utility function $u_i$. Let $w$ be a real number, interpreted as a wealth level. Say that $i$ accepts $g$ at $w$ if $E u_i(w + g) > u_i(w)$, where $E$ stands for “expectation;” i.e., if $i$ prefers taking the gamble to refusing it. Otherwise, $i$ rejects $g$ at $w$. Call $i$ at least as risk-averse as $j$ (written $i \succeq j$) if for all levels $w_i$ and $w_j$ of wealth, $j$ accepts at $w_j$ any gamble that $i$ accepts at $w_i$. Call $i$ more risk-averse than $j$ (written $i \succ j$) if $i \succeq j$ and $j \not\succeq i$.

Define an index as a positive real-valued function on gambles (to be thought of as measuring riskiness). Given an index $Q$, say that “gamble $g$ is riskier than gamble $h$” if $Q(g) > Q(h)$. We consider two axioms for $Q$, the first of which posits a kind of “duality” between riskiness and risk aversion; roughly, that less risk-averse agents accept riskier gambles. The axioms are as follows:

Duality:\(^3\) If $i \succ j$, $i$ accepts $g$ at $w_i$, and $Q(g) > Q(h)$, then $j$ accepts $h$ at $w_j$.

---

\(^1\)Monotonicity means that the individual likes money; concavity, that he is risk-averse—weakly prefers the expected value of a gamble over the gamble itself.

\(^2\)For simplicity, we assume for now that it takes finitely many values, each with positive probability. This assumption will be relaxed in the sequel.

\(^3\)Throughout this paper, the universal quantifier applies to variables that are not explicitly quantified otherwise. For example, the duality axiom should be understood as being prefaced by: “For all gambles $g$, $h$, agents $i,j$, and wealth levels $w_i, w_j$.”


In words, duality says that if the more risk-averse agent accepts the riskier gamble, then a fortiori the less risk-averse agent accepts the less risky gamble.

**Homogeneity of Degree 1:** $Q(tg) = tQ(g)$ for all positive numbers $t$.

Homogeneity embodies the *cardinal* nature of riskiness. If $g$ is a gamble, it makes sense to say that $2g$ is “twice as” risky as $g$, not just “more” risky. Similarly, $tg$ is $t$ times as risky as $g$. Our main result can now be stated as follows:

**Theorem A:** For each gamble $g$, there is a unique positive number $R(g)$ with

$$E e^{-g/R(g)} = 1.$$ 

The index $R$ thus defined satisfies duality and homogeneity of degree 1; and, any index satisfying these two axioms is a positive multiple of $R$.

We call $R(g)$ the *riskiness* of $g$. Both axioms are essential to the result: dropping either of them admits indices that are not positive multiples of $R$.

### 3 Discussion

#### 3.1 Emphasis on Losses

As we shall see in Section 6, the riskiness index $R$ is much more sensitive to the loss side of a gamble than to its gain side. Technically, that is because the exponential on the right side of (2.1) has a positive exponent if and only if the value of $g$ is negative. Conceptually, too, the idea of “risk” is usually associated with possible losses rather than with gains; one speaks more of risking losses than of risking smaller gains.

Many of the indices discussed in the literature (see Section 7.4 and 7.5) also emphasize losses. But in those cases, the emphasis on losses is built in; the definitions explicitly put more weight on the loss side. In the case of the index $R$, the definition as such does not distinguish between losses and gains, and indeed there is no sharp division between them. The distinction that we do observe emerges naturally from the analysis; it is not entered artificially.

#### 3.2 Risk Aversion and Duality

For one agent to be more risk-averse than another in our sense—$i \succ j$—is a very strong requirement. That is because $j$ must accept any gamble that $i$ does, quite independent of the gamble and the respective wealth levels. It is precisely this strength that makes the duality axiom highly acceptable: Since this strong requirement appears in the hypothesis of the axiom, the axiom as a whole calls for very little, and what it does call for is eminently reasonable.

#### 3.3 Relation with Arrow-Pratt

Arrow (1965, 1971) and Pratt (1964) define the coefficient of *absolute risk aversion* of an agent $i$ with utility function $u_i$ and wealth $w$ as $\rho_i(w) := \rho_i(w, u_i) :=$
This concept is “local,” in that it concerns i’s attitude towards infinitesimally small gambles only; in contrast, the concept $\succeq$ of comparative risk-aversion defined above is “global,” in that it applies to gambles of arbitrary size. In this way, the concept $\succeq$ seems more direct, straightforward, and natural; no limiting process is involved, one deals directly with real gambles.

Another distinction is that the Arrow-Pratt coefficient is defined for a particular wealth level $w$ only, whereas the concept $\succeq$ abstracts away from wealth, deals simultaneously with all wealth levels. This fits our purposes well: We seek a notion of riskiness that depends only on the gamble in question; current wealth should not matter.

On the other hand, the relation $\succeq$ is only a partial order, whereas Arrow and Pratt define a numerical index (and a fortiori, a total order). The two notions are related by the following:

**Proposition 3.1:** $i \succeq j$ if and only if $\rho_i(w_i) \geq \rho_j(w_j)$ for all $w_i$ and $w_j$.

### 4 Characterization in Terms of CARA

An agent $i$ is said to have **constant absolute risk aversion** (CARA) if his Arrow-Pratt coefficient $\rho_i(w)$ is a constant $\alpha$ that does not depend on the wealth $w$. In that case, $i$ is called a CARA agent, and his utility $u$ a CARA utility, both with parameter $\alpha$. There is in fact an essentially

unique CARA utility with parameter $\alpha$, given by $u(w) = -e^{-\alpha w}$. While defined in terms of the Arrow-Pratt coefficient, which is a local concept, CARA may in fact be characterized (or equivalently, defined) in global terms. Indeed, we have

**Proposition 4.1:** An agent $i$ has CARA if and only if for any gamble $g$ and any two wealth levels, $i$ either accepts $g$ at both wealth levels, or rejects $g$ at both wealth levels.

In words, whether or not $i$ accepts a gamble $g$ depends only on $g$, not on the wealth level. CARA utility functions thus constitute a kind of medium or context in which gambles may be evaluated “on their own,” without reference to wealth.

**Theorem B:** The riskiness $R(g)$ of a gamble $g$ is the reciprocal of the number $\alpha$ such that a CARA person with parameter $\alpha$ is indifferent between taking and not taking the gamble.

**Proof:** It follows from (2.1) and the form of CARA utilities.

Note that Theorem B goes a little beyond Theorem A in characterizing riskiness; it actually fixes the index numerically, not just within a positive constant.

---

4 Up to an additive and a positive multiplicative constant.
5 Properties of Riskiness

The properties below are proved on the spot, follow immediately from (2.1), or are proved in Section 8.

5.1 The Parameters of Riskiness

The riskiness of a gamble depends on the gamble only—indeed, on its distribution only—and not on any other parameters, such as the utility function of the decision maker or his wealth.

5.2 Dimension

Riskiness is measured in dollars.

5.3 Monotonicity w.r.t. Stochastic Dominance

The most uncontroversial, widely accepted notions of riskiness are provided by the concept of stochastic dominance (Hadar and Russell (1969), Hanoch and Levy (1969), Rothschild and Stiglitz (1970); see also Mas-Colell, Whinston and Green (1995, pp.195-197)). Say that a gamble \( g \) first-order dominates \( (FOD) \) \( g^* \) if \( g \geq g^* \) for sure, and \( g > g^* \) with positive probability; and \( g \) second-order dominates \( (SOD) \) \( g^* \) if \( g^* \) may be obtained from \( g \) by "mean-preserving spreads"—by replacing some of \( g \)'s values with random variables whose mean is that value. Say that \( g \) stochastically dominates \( g^* \) (in either sense) if there is a gamble that is distributed like \( g \) and that dominates \( g^* \) (in that sense).

An index \( Q \) is called first- (second-) order monotonic if \( Q(g) < Q(g^*) \) whenever \( g \) F(S)OD \( g^* \). First- and second-order dominance constitute partial orders. One would certainly expect any reasonable notion of riskiness to extend these partial orders—i.e., to be both first- and second-order monotonic. And indeed, the riskiness index \( R \) is monotonic in both senses.

5.4 Continuity

An index \( Q \) is called continuous if it is continuous in the topology of uniform convergence; i.e., if \( Q(g_n) \rightarrow Q(g) \) whenever \( g_n \rightarrow g \) uniformly.\(^5\) With this definition, the riskiness index \( R \) is continuous. It is also continuous if we adopt more demanding definitions, for example if we replace uniform convergence by convergence in probability, as long as the \( g_n \) are uniformly bounded. In either case, the continuity is not uniform, because as \( Eg \) approaches 0, the riskiness \( R(g) \) may approach \( \infty \).

\(^5\)Equivalently, if for every gamble \( g \) and \( \varepsilon > 0 \), there is a \( \delta > 0 \) such that \( |Q(g^*) - Q(g)| < \varepsilon \) whenever \( |g^* - g| < \delta \) for each of their values.
5.5 Diluted Gambles

If \( g \) is a gamble, \( p \) a number strictly between 0 and 1, and \( g^p \) a compound gamble that yields \( g \) with probability \( p \) and 0 with probability \( 1 - p \), then \( R(g^p) = R(g) \).

Though at first this may sound counterintuitive, on closer examination it is very reasonable; indeed, any expected utility maximizer—risk averse or not—accepts \( g^p \) if and only if he accepts \( g \).

5.6 Normal Gambles

If the gamble \( g \) has a normal distribution, then \( R(g) = \text{Var}(g) / 2E(g) \), where \( \text{Var} \) stands for “variance.” Indeed, set \( \text{Var}(g) =: \sigma^2 \) and \( E(g) =: \mu \). The density of \( g \)’s distribution is \( e^{-(x-\mu)^2/2\sigma^2} / \sigma \sqrt{2\pi} \), so

\[
E(e^{-g/(\sigma^2/2\mu)}) = \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} e^{-x/(\sigma^2/2\mu)} dx
\]

\[
= \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-[(x^2-2\mu x+m^2)+(4\mu x)]/2\sigma^2} dx
\]

\[
= \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x+m)^2/2\sigma^2} dx = 1.
\]

So (2.1) holds with \( R(g) := \sigma^2 / 2\mu \), so that is indeed the riskiness of \( g \).

In the finance literature, Variance/Mean is sometimes used to measure riskiness. We shall see below (Section 7.3) that this is in general not reasonable. But for normal gambles, it is, as we have just seen.

5.7 Sums of i.i.d. Gambles

If \( g \) and \( h \) are independent identically distributed (i.i.d.) gambles with riskiness \( \nu \), then \( g + h \) also has riskiness \( \nu \). Indeed, the hypothesis yields \( E(e^{-g/\nu}) = E(e^{-h/\nu}) = 1 \). Since \( g \) and \( h \) are independent, so are \( e^{-g/\nu} \) and \( e^{-h/\nu} \), so \( 1 = E(e^{-g/\nu} e^{-h/\nu}) = E(e^{-(g+h)/\nu}) \), so \( R(g + h) = \nu \).

5.8 Sums of Independent Gambles

The previous result may be generalized as follows: If \( g \) and \( h \) are independent, then the riskiness of \( g + h \) lies between those of \( g \) and \( h \).

5.9 Ordinality

If we are looking only for an ordinal index—i.e., wish to define “riskier,” without saying how much riskier—then we can replace the homogeneity axiom by conditions of monotonicity and continuity.

\(^6\) As defined in Section 2, a gamble has only finitely many values; so strictly speaking, its distribution cannot be normal. We therefore redefine a “gamble” as a random variable \( g \) (Borel-measurable function on a probability space) for which \( E(e^{-\alpha g}) \) exists for all positive \( \alpha \).
An index $Q$ for which $Q(g) > Q(h)$ if and only if $R(g) > R(h)$ is called \textit{ordinally equivalent} to $R$. We have already seen that the riskiness index $R$ satisfies the duality axiom (Theorem A), is continuous (5.4), and is both first- and second-order monotonic (5.3). In the opposite direction, any \textit{continuous and first-order monotonic index that satisfies the duality axiom is ordinally equivalent to} $R$. Moreover, \textit{continuity, monotonicity and duality are essential for this result}; without any one of them, it fails.

6 Some Numerical Examples

6.1 A Benchmark

A gamble that results in a loss of $l$ with probability $1/e$, and a “very large” gain with the remaining probability, has riskiness $l$. Formally, if $g_{M,l}$ yields $-l$ and $M$ with probabilities $1/e$ and $1 - (1/e)$ respectively; then $\lim_{M \to \infty} R(g_{M,l}) = l$.

The probability $1/e$ is that of “no success” in a Poisson distribution with expectation 1.

6.2 Some Half-Half Gambles

We have just seen that the riskiness of a gamble yielding a loss of $1$ with probability $1/e$, and a large gain with the remaining probability, is close to $1$. If the probabilities are half-half, the riskiness goes up to $1/\log 2 \approx 1.44$, where “log” denotes the natural logarithm (i.e., to base $e$). If the gain decreases to $3$ (so the expectation decreases from $\infty$ to $1$), the riskiness goes up again, but not by much—only to $1.64$. If the gain decreases to $1.1$—so the expectation is only $0.05$—the riskiness jumps to $11.01$. As the gain approaches $1$—i.e., the expectation approaches $0$—the riskiness approaches $\infty$. The riskiness of a half-half gamble yielding -$100 or $105 (cf. Rabin (2000)) is $2,100.

6.3 Insurance

To buy insurance is to reject a gamble. For example, suppose you insure a risk of losing $20,000 with probability $0.001$ for a premium of $100$—like when buying loss damage waiver in a car rental. That means that you will end up with $-100$ for sure. If you decline the insurance, you are faced with a gamble that yields $-20,000$ with probability .001, and 0 with probability 0.999. If we normalize\textsuperscript{7} so that rejecting the gamble is worth 0, then the gamble yields $-19,900$ with probability .001, and $100$ with probability 0.999. The riskiness of this gamble is about $2,750.

\textsuperscript{7}You cannot “stay where you are;” you must either pay the premium, which means moving to your current wealth $w$ less $100$, or decline the insurance, which means moving to $w - 100$ plus the gamble $g$ described in this sentence. That is like choosing between $g$ and $0$, from what your vantage point would be if your current wealth were $w - 100$. 

8
7 The Literature

There exist other indices in the literature that purport to measure riskiness. All those of which we know suffer from serious deficiencies,\(^8\) prominent among which is that they violate the elementary condition of monotonicity w.r.t. first order dominance (M-FOD); indeed, they may rate a gamble \(g\) riskier than \(h\) even though \(h\) is sure to yield more than \(g\). We will not conduct an exhaustive review of the literature, but content ourselves with discussing some of the more prominent indices, and briefly mentioning some others.

7.1 Measures of Dispersion

Pure measures of dispersion like standard deviation, variance, mean absolute deviation (E\(|g−Eg|\)), and interquartile range\(^9\) have been suggested as indices of riskiness; see the survey of Machina and Rothschild (1987). That seems bizarre, as these indices measure only dispersion, taking little account of the gamble’s actual values. Thus if \(g\) and \(g+c\) are gambles, where \(c\) is a positive constant, then any of these indices rate \(g+c\) precisely as risky as \(g\), in spite of the fact that it is sure to yield more than \(g\). An even stranger index (op. cit.) is entropy,\(^10\) which totally disregards the values of the gamble, taking into account only their probabilities; thus a gamble with three equally probable (but different) values has entropy \(\log_2 3\), no matter what its values are. It seems clear that people who suggest such measures of dispersion as indices of riskiness are not thinking of riskiness in the economic, decision-making sense that we are trying to capture here.

There are other indices that use measures of dispersion, but also factor in some measure of the gamble’s magnitude, most prominently its mean. We now discuss two of these; it turns out that they, too, violate M-FOD.

7.2 Standard Deviation/Mean

Standard deviation/mean is related to the Sharpe Ratio, a measure of “risk-adjusted returns” frequently used to evaluate portfolio selection; see, for example, Bodie, Kane and Marcus (2002) and Welch (2005). An odd feature of this index is that it is homogeneous of degree zero: A half-half gamble yielding $2 or -$1 is rated exactly as risky as one yielding $2,000,000 and -$1,000,000. Worse, though, is that the index violates M-FOD; indeed, a gamble that is sure to yield higher returns than another may nevertheless be rated riskier.

Let \(g\) be a gamble yielding \(-1\) with probability 0.02 and 1 with probability 0.98, and \(h\) a gamble that yields \(-1\) with probability 0.02, yields 1 with probability 0.49, and yields 2 with probability 0.49. Note that \(h\) never yields less

\(^8\)Indeed, by (5.9), an index that is not ordinally equivalent to our index \(R\) must violate continuity, or monotonicity, or duality.

\(^9\)The difference between the first and third quartiles of the gamble’s distribution. So, if a gamble yields \(-100, -1, 2,\) and \(1000\) with probability 1/4 each, then the interquartile range is 3.\(^{10}\)

\(^{10}\)\(−\sum_k p_k \log_2 p_k\), where the \(p_k\) range over the probabilities of the gamble’s different values.
than $g$, and yields more with probability almost half. The gamble $g$ has mean $\mu = 0.96$ and s.d. (standard deviation) $\sigma = 0.28$, so $\sigma/\mu = 7/24 \approx 0.29$. For $h$, the numbers are $\mu = 1.45$ and $\sigma = 7\sqrt{3}/20$, so $\sigma/\mu = 7\sqrt{3}/29 \approx 0.42$. Thus $h$ is rated considerably more risky than $g$, which is patently absurd. Moreover, for positive $\varepsilon$, the gamble $h + \varepsilon$ is sure to yield more than $g$; but if $\varepsilon$ is small enough, it will nevertheless be rated riskier.

### 7.3 Variance/Mean

Variance/mean $(\sigma^2/\mu)$ is another index that may be used to evaluate risks. This does have the “right” dimension—it is homogeneous of degree one—but like $\sigma/\mu$, it violates M-FOD. Indeed, the above example works here too. Even simpler is the following example: let $g$ and $h$ be half-half gambles yielding, respectively, $-2$ or $4$, and $-1$ or $17$. Then $h$ yields more than $g$ for sure; but $\sigma^2/\mu = 9$ for $g$, and $= 9^2/8 > 9$ for $h$. So $h$ is rated riskier than $g$—an absurdity.

In Section 5.6, we showed that in the case of normal gambles, our riskiness index $R$ does yield precisely $\sigma^2/2\mu$. But many gambles, especially those in finance, are very far from normal. For example, one may expect an investment in high tech either to fall flat on its face or to be wildly successful, with little inbetween. For such gambles, $\sigma^2/2\mu$ may be far from the riskiness. On the other hand, the normal approximation does make sense for a well-diversified portfolio, so in that case $\sigma^2/2\mu$ is a good measure of riskiness.

### 7.4 Value at Risk

Another index used extensively by banks and finance professionals in portfolio risk management is value at risk (VaR). This depends on a parameter called a confidence level. At a 95% confidence level, the VaR of a gamble $g$ is the absolute value of its fifth percentile, when that is non-positive, and 0 otherwise. In words, it is the greatest possible loss, ignoring losses with probability less than 5%. Thus a gamble yielding -$1,000,000, -$1, and $100,000 with respective probabilities of 0.04, 0.02, and 0.94 has a 95% VaR of $1, and so does the gamble yielding -$1 and $100,000 with 0.06 and 0.94 probabilities.

This index has various troubles. To start with, it depends on a parameter—the confidence level—whose “appropriate” value is not clear. Also, this index ignores completely the gain side of the gamble. In particular, it violates M-FOD. And even on the loss side, it concentrates only on that loss that “hits” the 95% level, as the above examples show.

### 7.5 Additional Indices

Brachinger (2002) and Brachinger and Weber (1997) are good surveys of some of the more recent literature. In psychology, measures of perceived risk have been proposed; early studies are Coombs (1969), and Pollatsek and Tversky (1970). Our index shares some of the rescaling and translation properties proposed by Luce (1980). Luce’s families of risk measures depend on a host of different
parameters, including transformations of the density of the random variable, as well as linear combinations of the conditional expectation of gains raised to some power and the conditional expectation of losses also raised to the same power. Similar comments apply to the conjoint expected risk model of Luce and Weber (1988). Fishburn (1977, 1982, 1984) generalizes many of the previous measures. He also develops indices where losses and gains are treated separately (see also Jia, Dyer and Butler (1999)).

Sarin’s (1987) risk measures improve upon Luce’s. One of Sarin’s measures, the closest to the index $R$, is $Ee^{-g} =: S(g)$. This is monotonic w.r.t. FOSD and SOSD, so it must violate duality. It also violates homogeneity of degree 1. Indeed, let $g$ be the gamble that assigns probability 0.01 to a loss of 1 and probability 0.99 to a gain of 2. Then $S(2g) = 0.09 < 0.16 = S(g)$. In contrast, $R(2g) = 2R(g) > R(g)$. To see now that the index $S$ violates duality, set $\alpha := 1/R(g)$. By (2.1), a CARA agent $i$ with parameter $\frac{2}{\alpha}$ accepts $g$, while a CARA agent $j$ with parameter $\frac{2}{\alpha}$—who is less risk-averse than $i$—rejects $2g$, which is rated less risky than $g$ by $S$. So $S$ violates duality.

Most of this literature develops families of indices rather than proposing a single index of riskiness, and it is unclear how to assign values to many of the parameters upon which the index depends.

8 Proofs

8.1 Preliminaries

In this section, agents $i$ and $j$ have utility functions $u_i$ and $u_j$, and Arrow-Pratt coefficients $\rho_i$ and $\rho_j$ of absolute risk aversion. Since utilities may be modified by additive and positive multiplicative constants, we may—and do—assume throughout the following.

(1) $u_i(0) = u_j(0) = 0$ and $u'_i(0) = u'_j(0) = 1$.

**Lemma 2:** For some $\delta > 0$, suppose that $\rho_i(w) > \rho_j(w)$ at each $w$ with $|w| < \delta$. Then $u_i(w) < u_j(w)$ whenever $|w| < \delta$ and $w \neq 0$.

**Proof:** Let $|y| < \delta$. If $y > 0$, then by (1),

$$
\log u'_i(y) = \log u'_j(y) - \log u'_i(0) = \int_0^y (\log u'_i(z))'dz = \int_0^y (u''_i(z)/u'_i(z))dz = \int_0^y -\rho_i(z)dz < \int_0^y -\rho_j(z)dz = \log u'_j(y).
$$

If $y < 0$, the reasoning is similar, but the inequality is reversed, because then $\int_0^y = -\int_0^{|y|}$. Thus $\log u'_i(y) \leq \log u'_j(y)$ when $y \geq 0$, so also $u'_i(y) \leq u'_j(y)$ when $y \geq 0$.

So if $w > 0$, then by (1), $u_i(w) = \int_0^w u'_i(y)dy < \int_0^w u'_j(y)dy = u_j(w)$, and if $w < 0$, then $u_i(w) = -\int_0^{|w|} u'_i(y)dy < -\int_0^{|w|} u'_j(y)dy = u_j(w)$, q.e.d.

**Corollary 3:** If $\rho_i(w) \leq \rho_j(w)$ for all $w$, then $u_i(w) \geq u_j(w)$ for all $w$.

**Proof:** It is similar to that of Lemma 2, with $i$ and $j$ interchanged, strict inequalities replaced by weak inequalities, and the restriction to $|w| < \delta$ eliminated.
Lemma 4: If \( \rho_i(w_i) > \rho_j(w_j) \), then there is a gamble \( g \) that \( j \) accepts at \( w_j \) and \( i \) rejects at \( w_i \).

Proof: W.l.o.g.\(^{11} \) \( w_i = w_j = 0 \), so \( \rho_i(0) > \rho_j(0) \). Since \( u_i \) and \( u_j \) are twice continuously differentiable, it follows that there is a \( \delta > 0 \) such that \( \rho_i(w) > \rho_j(w) \) at each \( w \) with \( |w| < \delta \). So by Lemma 2,

\[
(5) \quad u_i(w) < u_j(w) \text{ whenever } |w| < \delta \text{ and } w \neq 0.
\]

Choose \( \varepsilon \) with \( 0 < \varepsilon < \delta / 2 \). For \( 0 \leq x \leq \varepsilon \), and \( k = i, j \), set \( f_k(x) := \frac{1}{2}u_k(-\varepsilon + x) + \frac{1}{2}u_k(\varepsilon + x) \). By (5),

\[
(6) \quad f_i(x) < f_j(x) \text{ for all } x.
\]

By (6), concavity, and (1), \( f_i(0) < f_j(0) \leq u_j(0) = 0 \). By monotonicity of the utilities, \( f_i(\varepsilon) = \frac{1}{2}u_i(2\varepsilon) = \frac{1}{2}u_i(0) = 0 \). So \( f_i(y) = 0 \) for some \( y \) between 0 and \( \varepsilon \), since \( f_i \) is continuous. So by (6), \( f_j(y) \geq 0 \). So if \( \eta > 0 \) is sufficiently small, then \( f_j(y-\eta) > 0 > f_i(y-\eta) \). So if \( g \) is the half-half gamble yielding \(-\varepsilon + y - \eta \) or \( \varepsilon + y - \eta \), then \( \text{Eu}_j(g) = f_j(y-\eta) > 0 > f_i(y-\eta) = \text{Eu}_i(g) \). So \( j \) accepts \( g \) whereas \( i \) rejects it, q.e.d.

8.2 Proof of Proposition 3.1

This Proposition is used in the proof of Theorem A.

"Only if": Assume \( i \gtrless j \); we must show

\[
(7) \quad \rho_i(w_i) \gtrless \rho_j(w_j) \text{ for all wealth levels } w_i \text{ and } w_j.
\]

If not, then there are \( w_i \) and \( w_j \) with \( \rho_i(w_i) < \rho_j(w_j) \). So by Lemma 4, there is a gamble that \( i \) accepts and \( j \) rejects, contradicting \( i \gtrless j \). So (7) is proved.

"If": Assume (7); we must show \( i \gtrless j \), i.e., that for all wealth levels \( w_i \) and \( w_j \) and each gamble \( g \), if \( i \) accepts \( g \) at \( w_i \), then \( j \) accepts \( g \) at \( w_j \). W.l.o.g. \( w_j = w_i = 0 \), so we must show that

\[
(8) \quad \text{if } i \text{ accepts } g \text{ at } 0, \text{ then } j \text{ accepts } g \text{ at } 0.
\]

From (1), (7), and Corollary 3 (with \( i \) and \( j \) reversed), we conclude \( u_j(w) \geq u_i(w) \) for each \( w \). So \( \text{Eu}_j(g) \geq \text{Eu}_i(g) \), which yields (8), q.e.d.

8.3 Proof of Theorem A

For \( \alpha > 0 \), let \( u_\alpha(x) = (1 - e^{-\alpha x})/\alpha \); this is a CARA utility function with parameter \( \alpha \). The functions \( u_\alpha \) satisfy (1), so by Lemma 2 (with \( \delta \) arbitrarily large), their graphs are "nested;" that is,

\[
(9) \quad \text{if } \alpha > \beta, \text{ then } u_\alpha(x) < u_\beta(x) \text{ for all } x \neq 0.
\]

To see that there is a unique \( R(g) > 0 \) satisfying (2.1), set \( f(\alpha) := Ee^{-\alpha g} - 1 \), and note that \( f \) is convex, \( f(0) = 0 \), \( f'(0) < 0 \), and \( f(M) > 0 \) for \( M \) sufficiently large. So there is a unique \( \gamma > 0 \) with \( f(\gamma) = 0 \), and we set \( R(g) := 1/\gamma \).

\(^{11} \) "Without loss of generality." For arbitrary \( w_i \) and \( w_j \), define \( u_i^*(x) := u_i(x + w_i) \) and \( u_j^*(x) := u_j(x + w_j) \), and apply the current reasoning to \( u_i^* \) and \( u_j^* \).
To see that $R$ satisfies the duality axiom, let $i, j, g, h, w_i, w_j$ be as in the hypothesis of that axiom; w.l.o.g. $w_i = w_j = 0$. Set $\gamma := 1/R(g)$, $\eta := 1/R(h)$, $\alpha_i := \inf_w \rho_i(w)$, $\alpha_j := \sup_w \rho_j(w)$. Thus

(10) $E u_i(g) = (1 - E e^{-\gamma}) / \gamma = 0$ and $E u_j(h) = (1 - E e^{-\eta}) / \eta = 0$.

By hypothesis, $R(g) > R(h)$, so $\eta > \gamma$. By Corollary 3, (11) $u_i(x) \leq u_{\alpha_i}(x)$ and $u_{\alpha_j}(x) \leq u_j(x)$ for all $x$.

Now assume $E u_i(g) > 0$; we must prove that $E u_j(h) > 0$. From $E u_i(g) > 0$ and (11) it follows that $E u_{\alpha_i}(g) > 0$. So by (10), $E u_i(g) = 0 < E u_{\alpha_i}(g)$. So by (9), $\gamma > \alpha_i$. By Proposition 3.1, $\alpha_i \geq \alpha_j$, so $\eta > \gamma$ yields $\alpha_j < \eta$. Then (10), (9) and (11) yield $0 = E u_j(h) < E u_{\alpha_i}(h) < E u_j(h)$, so indeed, $R$ satisfies the duality axiom. That $R$ is homogeneous of degree 1 is immediate, so indeed, $R$ satisfies the axioms.

In the opposite direction, let $Q$ be an index that satisfies the axioms. We first show that $Q$ is ordinally equivalent to $R$. If this is not true, then there must exist $g$ and $h$ that are ordered differently by $Q$ and $R$. This means that either the respective orderings are reversed, i.e.,

(13) $Q(g) > Q(h)$ and $R(g) < R(h)$,

or that equality holds for exactly one of the two indices; i.e.,

(14) $Q(g) > Q(h)$ and $R(g) = R(h)$

or

(15) $Q(g) = Q(h)$ and $R(g) > R(h)$.

If either (14) or (15), then by homogeneity, replacing $g$ by $(1 - \varepsilon)g$ for sufficiently small positive $\varepsilon$ leads to reversed inequalities. So w.l.o.g. we may assume (13).

Now let $\gamma := 1/R(g)$, $\eta := 1/R(h)$; then (10) holds. By (13), $\gamma > \eta$. Choose $\mu$ and $\nu$ so that $\gamma > \mu > \nu > \eta$. Then $u_\gamma(x) < u_\mu(x) < u_\nu(x) < u_\eta(x)$ for all $x \neq 0$. So by (10), $E u_\mu(g) > E u_\gamma(g) = 0$ and $E u_\nu(h) < E u_\eta(h) = 0$. So if $i$ and $j$ have utility functions $u_\mu$ and $u_\nu$ respectively, then $i$ accepts $g$ and $j$ rejects $h$. But from $\mu > \nu$ and Proposition 3.1, it follows that $i \succeq j$, contradicting the duality axiom for $Q$. So (12) is proved.

To see that $Q$ is a positive multiple of $R$, let $g_0$ be an arbitrary but fixed gamble, and set $\lambda := Q(g_0)/R(g_0)$. If $g$ is any gamble, and $t := Q(g)/Q(g_0)$, then $Q(tg_0) = tQ(g_0) = Q(g)$, so $tR(g_0) = R(tg_0) = R(g)$ by the ordinal equivalence between $Q$ and $R$, so $R(g)/R(g_0) = t = Q(g)/Q(g_0)$, so $Q(g)/R(g) = Q(g_0)/R(g_0) = \lambda$, so $Q(g) = \lambda R(g)$. This completes the proof of Theorem A.

To see that both duality and homogeneity of degree 1 are essential to Theorem A, consider the following two examples. The variance-mean ratio is homogeneous of degree 1, but violates duality. The index $Q(g) = I[R(g)]$, where $I$ denotes the integer part of a real number, satisfies duality, but is not homogeneous of degree 1.
8.4 Proof of 5.3

For $\alpha \geq 0$, set $f(\alpha) := E e^{-\alpha g}$, $f_*(\alpha) := E e^{-\alpha g_*}$. If $g$ FOSD $g_*$, then $f(\alpha) < f_*(\alpha)$ whenever $\alpha > 0$. From this and the proof that (2.1) has a unique positive root,\(^{12}\) it follows that the unique positive root of $f_*=1$ is smaller than that of $f=1$, so $R(g_*) > R(g)$, as asserted.

If $g$ SOSD $g_*$, then, too, $f(\alpha) < f_*(\alpha)$, because of the strict convexity of $e^{-\alpha x}$ as a function of $x$. The remainder of the proof is as before.

8.5 Proof of 5.4

For $\alpha \geq 0$, set $f(\alpha) := E e^{-\alpha g}$, $f_n(\alpha) := E e^{-\alpha g_n}$; denote the unique positive root of $f = 1$ by $\gamma$, of $f_n = 1$ by $\gamma_n$. We have $f_n \rightarrow f$, uniformly in any finite interval. Now $f(\gamma/2) < 1$ and $f(2\gamma) > 1$. So for $n$ sufficiently large, $f_n(\gamma/2) < 1$ and $f_n(2\gamma) > 1$, so $\gamma/2 < \gamma_n < 2\gamma$. Suppose that the $\gamma_n$ have a limit point $\gamma_* \neq \gamma$; arguing by contradiction, we may assume w.l.o.g. that it is the limit. For any $\varepsilon > 0$, we have $|f_n(\gamma_n) - f(\gamma_n)| < \varepsilon$ for $n$ sufficiently large, because of the uniform convergence. Also $|f(\gamma_n) - f(\gamma_*)| < \varepsilon$, because of the continuity of $f$. So $|f_n(\gamma_n) - f(\gamma_*)| < 2\varepsilon$. So $\lim f_n(\gamma_n) = f(\gamma_*) \neq 1$, contradicting $f_n(\gamma_n) = 1$; q.e.d.

8.6 Proof of 5.8

By Theorem A, the riskiness $R(g+h)$ is the reciprocal of the unique positive root of $f = 1$, where $f(\alpha) := E e^{-\alpha (g+h)}$. Because $g$ and $h$ are independent, $f(\alpha) = E e^{-\alpha g} e^{-\alpha h} = E e^{-\alpha g} e^{-\alpha h}$. So if $f(\alpha) = 1$, then it cannot be that both $E e^{-\alpha g}$ and $E e^{-\alpha h}$ are $>1$, and it cannot be that both $E e^{-\alpha g}$ and $E e^{-\alpha h}$ are $<1$. So $E e^{-\alpha g} \leq 1$ and $E e^{-\alpha h} \geq 1$, say. So $1/R(g+h) = \alpha \leq 1/R(g)$ and similarly $1/R(g+h) = \alpha \geq 1/R(h)$. Thus $R(g) \leq R(g+h) \leq R(h)$, as asserted.

8.7 Proof of 5.9

The proof of ordinal equivalence follows that of (12) above. If either (14) or (15) holds, and $Q$ is first-order monotonic, then replacing $g$ by $g - \varepsilon$ for sufficiently small positive $\varepsilon$ leads to reversed inequalities; this follows from monotonicity and continuity. If $Q$ is second-order monotonic, then we reach the same conclusion by applying a small mean-preserving spread to $g$. The remainder of the proof of (12) is as above.

To see that first-order monotonicity is essential, define

\[
Q(g) := \begin{cases} 
R(g), & \text{when } 0 < R(g) \leq 1, \\
1, & \text{when } 1 \leq R(g) \leq 2, \\
R(g) - 1, & \text{when } 2 \leq R(g).
\end{cases}
\]

Thus $Q$ collapses the interval $[1,2]$ in the range of $R$ to a single point. It may be seen that it is continuous and satisfies the duality axiom, but is not first-order.

\(^{12}\)Near the beginning of the proof of Theorem A.
monotonic; and there are \( g \) and \( h \) (in the “collapsed” region) satisfying (15), so \( Q \) is not ordinally equivalent to \( R \).

To see that continuity is essential, let \( A \) be a non-empty proper subset of the set \( R^{-1}(1) \) of all gambles with riskiness 1. Define

\[
Q(g) := \begin{cases} 
R(g), & \text{when } R(g) < 1 \text{ or } g \in A, \\
R(g) + 1, & \text{when } R(g) > 1 \text{ or } g \in R^{-1}(1) \setminus A.
\end{cases}
\]

One may think of \( Q \) as resulting from \( R \) by “tearing” along the “seam” \( R(g) = 1 \), with the seam itself going partly to the upper fragment and partly to the lower fragment. It may be seen that \( Q \) is first-order monotonic and satisfies the duality axiom, but is not continuous; and there are \( g \) and \( h \) (on the “seam”) satisfying (15), so \( Q \) is not ordinally equivalent to \( R \).

Finally, as already argued at the end of Section 7, Sarin’s index \( S(g) \) is continuous and first-order monotonic, but it violates duality.

### 8.8 Proof of Proposition 4.1

“Only if.” All CARA utility functions have the form \(-e^{-\alpha x}\). Thus \( i \) accepts \( g \) at wealth \( w \) if and only if \(-Ee^{-\alpha(g+w)} > -e^{-\alpha w}\), i.e., if and only if \( Ee^{-\alpha g} < 1 \); and this condition does not depend on \( w \).

“If.” Suppose \( i \)’s Arrow-Pratt index of absolute risk aversion is not constant, say \( \rho(w) > \rho(w_*) \). Consider a gamble yielding \( \pm \delta \) with probabilities \( p \) and \( 1-p \) respectively, and let \( p_\delta(w) \) be that \( p \) for which \( i \) is indifferent at \( w \) between taking and not taking the gamble. Then

\[
\rho(w) = \lim_{\delta \to 0} \frac{p_\delta(w) - \frac{1}{2}}{\delta};
\]

i.e., noting that even-money \( \frac{1}{2} - \frac{1}{2} \) bets are always rejected by risk-averse utility maximizers, the Arrow-Pratt index is the probability premium over \( \frac{1}{2} \), per dollar, that is needed for \( i \) to be indifferent between taking and not taking a small even-money gamble. So, if \( \delta \) is sufficiently small, \( q - \frac{1}{2} \) lies half-way between \( \rho(w) \) and \( \rho(w_*) \), and \( g \) is an even money gamble yielding \( \pm \delta \) with probabilities \( q \) and \( 1-q \) respectively, then \( i \) accepts \( g \) at \( w_* \) and rejects it at \( w \); this proves the contrapositive of “if,” and so “if” itself.

### 9 Conclusion

We have defined a numerical index of the riskiness of a gamble with stated dollar outcomes and stated probabilities, based on economic, decision-theoretic principles. Contrary to other indices of riskiness, this index is not one of a family, but stands alone. It is denominated in dollars, monotonic w.r.t. first and second order stochastic dominance, continuous in about any sense one wishes, homogeneous of degree 1, and satisfies a duality condition that says, very roughly, that risk-aversers dislike riskier gambles. Moreover, it is the only index satisfying these conditions.

\[\text{E.g., see Aumann and Kurz (1977), Section 6; but there may well be earlier sources.}\]
REFERENCES


