

An Economic Index of Riskiness

Robert J. Aumann

Hebrew University of Jerusalem

Roberto Serrano

Brown University and Instituto Madrileño de Estudios Avanzados

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 that, roughly speaking, asserts that less risk-averse individuals accept riskier gambles. The index is positively homogeneous, continuous, and subadditive; respects first- and second-order stochastic dominance; and for normally distributed gambles is half of variance/mean. Examples are calculated, additional properties are derived, and the index is compared with others.

I. 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-

The index proposed here first surfaced in conversations between Ignacio Palacios-Huerta, Oscar Volij, and the authors during a visit of Aumann to Brown University in March 2004; it was applied in a working paper generated later that spring (Palacios-Huerta, Serrano, and Volij 2004). At that time it still had somewhat of an ad hoc flavor; among other things, there was as yet no axiomatic characterization. We are grateful to Palacios-Huerta and Volij for their input. We thank Kenneth Arrow, Zvi Artstein, Antonio Cabrales, Dean Foster, Sergiu Hart, Haim Levy, Petros Milionis, Amy Serrano, and Ivo Welch for comments, for pointing out useful references, and for research assistance and Yechezkel Zilber for pointing out an error in a previous version. We are also grateful to Nancy Stokey and three anonymous referees for helpful comments. Serrano thanks Centro de Estudios Monetarios y Financieros in Madrid for its hospitality.

[*Journal of Political Economy*, 2008, vol. 116, no. 5]
© 2008 by The University of Chicago. All rights reserved. 0022-3808/2008/11605-0001\$10.00

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 measure riskiness *objectively*—independently of the person or entity taking the risk?

Conceptually, whether or not a person takes a gamble depends on two distinct considerations (Diamond and Stiglitz 1974): (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 classic contributions of Arrow (1965, 1971) and Pratt (1964) address item ii by defining *risk aversion*, which is a personal, subjective concept, 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 (“3 hours”) or about heat or cold aversion (“it’s too cold in here”) without an objective measure of temperature (“20 degrees Fahrenheit”).

This paper addresses item i: it develops an index 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-avertter is averse. So on the whole, we expect individuals who are less risk averse to take riskier gambles. As Machina and Rothschild (2008, 7:193) put it, “risk is what risk-avertters hate.”

Unlike some other riskiness indices that have been proposed in disciplines such as finance, statistics, and psychology (Sec. VIII), ours is based on economic, decision-theoretic ideas, such as the duality principle roughly enunciated above and respect for first- and second-order stochastic dominance (Sec. V.C). Clearly, riskiness is related to dispersion, so a good riskiness measure should be monotonic with respect to second-order stochastic dominance. Less well understood, perhaps, is that riskiness should also relate to location and thus be monotonic with respect to first-order stochastic dominance, in particular, that a gamble that is sure to yield more than another should be considered less risky. Both stochastic dominance criteria are uncontroversial, and one advantage of our index is that it completes the partial ordering on gambles that they induce.

The paper is organized as follows: Section II discusses the purpose and potential uses of the proposed index—what it is we are seeking. Section III is devoted to the basic axiomatic definition of the index and its numerical characterization. Section IV relates our index to Arrow-Pratt risk aversion. Specifically, it carefully discusses our basic axiom, “duality,” in its own right as well as in relation to Arrow-Pratt (Sec. IV.A);

characterizes our index in terms of Arrow-Pratt constant absolute risk aversion, as outlined in the abstract (Sec. IV.B); and relates it to Arrow-Pratt relative risk aversion (Sec. IV.C). Section V sets forth some desirable properties of the index (in addition to the axioms) such as continuity; respect for stochastic dominance; subadditivity; its dimension (dollars); its behavior for normal gambles, for independent gambles, for “diluted” gambles, and for compound gambles; and the relatively greater weight it puts on losses vis-à-vis gains and, finally, its behavior for financial instruments (i.e., “multiplicative” gambles). Section VI characterizes the index ordinally. Section VII adduces some numerical examples, meant to give an intuitive feel for the index. Section VIII reviews the literature. Section IX is devoted to discussion and Section X to proofs. Section XI presents conclusions. Throughout, assertions that are not proved on the spot and are not immediate are proved in Section X.

II. The Concept and Its Uses

As remarked above, the concept of “riskiness” is ubiquitous in financial discussions. Investors are told that one investment may hold an opportunity for high returns but be “risky,” whereas another may be “safer” but yield lower returns. Mutual funds are characterized as “safe” or “venture capital” or “blue-chip” or “volatile”; bonds are rated AAA, AA, and so forth; and so on. We repeatedly hear that an investment that is appropriate for one investor may be “too risky” for another or that a pension fund makes “unduly risky” investments.

Here we propose to *quantify* riskiness—describe it with numbers rather than adjectives or letter “ratings.” The main purpose of such a quantification is the same as that of the adjectives and the letter ratings—to help investors and other decision makers make their decisions. For example, the investments of pension funds could be required not to exceed a stated level of riskiness. Or an investor, on being told the riskiness index of an investment, could say “well, that’s too risky for me,” or “that’s a little risky but I’ll go for it,” or “hey, that sounds just right for me.” Or an advisor could say, if you’re living on a pension you should not accept gambles that exceed such and such a riskiness, but if you’re young and have plenty of opportunities, you could up that by so-and-so much.

From this viewpoint it is clear that if the gamble g is sure to yield more than h , it cannot be considered riskier. We are considering risk-averse decision makers—those for whom risks are undesirable—who, “all other things being equal,” prefer less risky alternatives.

But riskiness and desirability are not opposites; a less risky gamble is not always more desirable. That depends on the decision maker and

on other parameters in addition to riskiness, such as the mean, maximum loss, opportunities for gain, and so on, indeed, on the whole distribution. Desirability is subjective, depending on the decision maker; one may prefer gamble g to gamble h , whereas another prefers h to g . Riskiness, however, is objective: it is the same for all individuals. Given two gambles, a more risk-averse individual may well prefer the less risky gamble, whereas a less risk-averse individual may find that the opportunities afforded by the riskier gamble outweigh the risk involved.

Like any index or summary statistic—the Gini index of inequality, parameters of distributions (mean, median, variance), the Shapley value of a game, market indices (Dow Jones, Standard and Poor's 500), cost-of-living indices, difficulty ratings of rock climbs (3, 4, 5.1–5.13; I, II, III) and ski runs (green, blue, red, black), and so on—the riskiness index summarizes a complex, high-dimensional object by a single number. Needless to say, no index captures all the relevant aspects of the situation being summarized. But once accepted, it takes on a life of its own; its “consumers” internalize its content through repeated use.

In addition to these practical uses, a riskiness index could also be a useful research tool. For example, Rabin (2000) asserts that most people would reject a gamble yielding +\$105 or −\$100 with half-half probabilities. While this sounds plausible on its face, it is difficult to verify empirically (as opposed to “experimentally”), since such gambles are not readily available in the real world. What one can ask is, Do people accept gambles with a “similar” level of riskiness? Once one has a measure of riskiness, one can approach that question by looking at real-life gambles, for example, insurance contracts.¹

Early attempts to quantify riskiness were based on mean and variance only (see Machina and Rothschild 2008). Defending this approach, Tobin (1969, 14) wrote that its critics “owe us more than demonstrations that it rests on restrictive assumptions. They need to show us how a more general and less vulnerable approach will yield the kind of comparative-static results that economists are interested in.” That is what our index aims to do.

III. Axiomatic Characterization

In this paper, a *utility function* is a von Neumann–Morgenstern utility function for money; is strictly monotonic, strictly concave,² and twice continuously differentiable; and defined over the entire real line. A

¹ Rejecting insurance is like accepting a gamble. Since insurance usually has negative expectation for the purchaser, rejecting it has positive expectation.

² Strict monotonicity means that the individual likes money; strict concavity, that he is risk averse—prefers the expected value of a gamble over the gamble itself.

gamble g is a random variable with real values³—interpreted as dollar amounts—some of which are negative, and that has positive expectation.

Say that an agent with utility function u accepts a gamble g at wealth w if $Eu(w + g) > u(w)$, where E stands for “expectation,” that is, if he prefers taking the gamble at w to refusing it. Otherwise, he rejects it. Call agent i *uniformly no less risk-averse than* agent j (written $i \geq j$) if whenever i accepts a gamble at some wealth, j accepts that gamble at any wealth. Call i *uniformly*⁴ more risk-averse than j (written $i \triangleright j$) if $i \geq j$ and $j \not\geq i$.

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.⁵ If $i \triangleright j$, i accepts g at w , and $Q(g) > Q(h)$, then j accepts h at w .

Roughly,⁶ duality says that if the more risk-averse of two agents accepts the riskier of two gambles, then a fortiori the less risk-averse agent accepts the less risky gamble.

POSITIVE HOMOGENEITY. $Q(tg) = tQ(g)$ for all positive numbers t .

Positive 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 is now as follows.

THEOREM A. For each gamble g , there is a unique positive number $R(g)$ with

$$Ee^{-g/R(g)} = 1. \quad (3.1)$$

The index R thus defined satisfies duality and positive homogeneity, 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: omitting either admits indices that are not positive multiples of R . But duality is by far the more central: together only with certain weak conditions of continuity and monotonicity—but not positive homogeneity—it already implies that the index is ordinally equivalent to R (see theorem D in Sec. VI).

³ For simplicity, we assume for now that it takes finitely many values, each with positive probability. This assumption will be relaxed later.

⁴ See Sec. IV.A for a discussion of this terminology.

⁵ Throughout, 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 .”

⁶ To make this precise (but awkward), one should preface “more” and “less” by “uniformly.”

IV. Relation with Arrow-Pratt

A. Risk Aversion and Duality

To understand the concept of uniform comparative risk aversion (Sec. III) that underlies our treatment, recall first that Arrow (1965, 1971) and Pratt (1964) define the coefficient of *absolute risk aversion* (ARA) of an agent i with utility function u_i and wealth w as

$$\rho_i(w) := \rho(w, u_i) := -u_i''(w)/u_i'(w).$$

Now, call i *no less risk-averse* than j if, at any given wealth, j accepts any gamble that i accepts.⁷ Then

$$(4.1.1) \quad i \text{ is no less risk averse than } j \text{ if and only if } \rho_i(w) \geq \rho_j(w) \text{ for all } w.$$

Our concept of $i \geq j$ —that i is *uniformly* no less risk-averse than j —is much stronger. It says that if i accepts a gamble at some wealth, j also accepts it—not only at that given wealth, but at any wealth. Parallel to (4.1.1), we then have

$$(4.1.2) \quad i \text{ is uniformly no less risk averse than } j \text{ if and only if } \rho_i(w_i) \geq \rho_j(w_j) \text{ for all } w_i \text{ and } w_j \text{ (i.e., } \min_w \rho_i(w) \geq \max_w \rho_j(w)).$$

Arrow-Pratt risk aversion is a “local” concept in that it concerns i ’s attitude toward infinitesimally small gambles at a specified wealth only; in contrast, our two concepts of comparative risk aversion are “global” in two senses: (i) they apply to gambles of arbitrary finite size, which (ii) may be taken at any wealth. Thus our concepts seem more direct, straightforward, and natural; no limiting process is involved: one deals directly with real gambles. However, we get only partial orders, whereas Arrow and Pratt define a numerical index (and so a total order). The three concepts are related by (4.1.1) and (4.1.2).

For one agent to be uniformly more risk averse than another— $i \triangleright j$ —is a very strong requirement. 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.

⁷ Closely related—in view of (4.1.1)—is the concept of Diamond and Stiglitz (1974, 346), who call i more risk averse than j if $\rho_i(w) > \rho_j(w)$ for all w . But this has no straightforward equivalent in terms of finite gambles.

B. CARA

An agent i is said to have *constant absolute risk aversion* (CARA) if his ARA is a constant α that does not depend on his wealth. In that case, i is called a CARA agent, and his utility u a CARA utility, both with parameter α . There is an essentially⁸ unique CARA utility with parameter α , given by $u(w) = -e^{-\alpha w}$. While defined in terms of the local concept of risk aversion, CARA may in fact be characterized (or, equivalently, defined) in global terms, as follows:

(4.2.1) An agent i has CARA if and only if for any gamble g and any two wealth levels, i either accepts g at both levels or rejects g at both 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; in particular, one can speak of CARA agents “accepting” or “rejecting” a gamble without specifying the wealth. This kind of “wealth-free environment” is, of course, precisely what we want when seeking an objective riskiness measure. We then have

(4.2.2) If a CARA agent accepts a gamble, then any CARA agent with a smaller parameter also accepts the gamble. Equivalently, if a CARA agent rejects a gamble, then any CARA agent with a larger parameter also rejects the gamble.

From (4.2.2) it follows that for each gamble g , there is precisely one “cutoff” value of the parameter, such that g is accepted by CARA agents with a smaller parameter and rejected by CARA agents with a larger parameter. The larger the parameter, the more risk averse the agent; so the duality principle—that less risk-averse agents accept riskier gambles—indicates that this cutoff might be a good inverse measure of riskiness. And indeed, we have the following theorem.

THEOREM B. The riskiness $R(g)$ of a gamble g is the reciprocal of the number α such that a CARA person with parameter α is indifferent between taking and not taking the gamble.

Proof. Follows from (3.1) 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.⁹

⁸ Up to additive and positive multiplicative constants.

⁹ See also Palacios-Huerta and Serrano (2006): what we call α here, referred to as a^* in that paper, plays a major role in their discussion of Rabin (2000).

C. CRRA

An agent i is said to have *constant relative risk aversion* (CRRA) if his ARA is inversely proportional to his wealth w , that is, if $w\rho_i(w)$ is a constant, called the CRRA parameter. CRRA expresses the idea that wealthier people are less risk averse. Here wealth is assumed positive, so in contrast to the rest of this paper, in this subsection we discuss utility functions defined on the positive reals only.

As with CARA, there is an essentially unique CRRA utility with a given parameter. For parameter 1, it is the classic logarithmic utility, $\log w$, originally proposed by Daniel Bernoulli (1738).¹⁰ In terms of this utility, one can lend operational meaning to the riskiness $R(g)$ of a gamble g as follows:¹¹

An agent with logarithmic utility and initial wealth w accepts a gamble g if $w + \min g > R(g)$ and rejects it if $w + \max g < R(g)$.

That is, g is accepted if taking the gamble necessarily results in a wealth greater than $R(g)$ and rejected if it necessarily results in a wealth smaller than $R(g)$. Thus when the range of g is small compared to its riskiness, the riskiness represents an approximate cutoff; the gamble is accepted if the initial wealth is considerably greater than the riskiness and is rejected if it is considerably less.

More generally, we have

(4.3.1) A CRRA agent with parameter γ and initial wealth w accepts a gamble g if $w + \min g > \gamma R(g)$ and rejects it if $w + \max g < \gamma R(g)$.

That is, for a CRRA agent with parameter γ , the approximate cutoff wealth level to accept the gamble is γ times its riskiness. Assertion (4.3.1) is an immediate consequence of the following proposition, which is of interest in its own right and does not assume CRRA:

(4.3.2) If $\rho_i(x) < 1/R(g)$ for all x between $w + \min g$ and $w + \max g$, then i accepts g at w ; if $\rho_i(x) > 1/R(g)$ for all such x , then i rejects g at w .

Next, let $w_\gamma(g)$ denote the cutoff wealth, whenever it is well defined, of a CRRA agent with parameter γ : that at which he is precisely indif-

¹⁰ Alternatively characterized by marginal utility being inversely proportional to wealth.

¹¹ The original idea of lending an operational meaning to the dollar amount of the riskiness is due to Foster and Hart (2007).

ferent between taking and not taking the gamble g .¹² Then, for $\gamma \geq 1$, $w_\gamma(g)$ is a continuous nondecreasing function of γ , and (4.3.1) yields the following theorem.

THEOREM C. $\lim_{\gamma \rightarrow \infty} w_\gamma(g)/\gamma = R(g)$.

Thus for very risk-averse CRRA agents, the cutoff wealth is proportional to the parameter, the constant of proportionality being the riskiness.¹³ These results come close to an “operational” characterization of riskiness in terms of the well-known and widely applied CRRA concept.

V. Some Properties of Riskiness

A. 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.

B. Dimension

Riskiness is measured in dollars. For an “operational” interpretation of the dollar amount, see Section IV.C.

C. Monotonicity with Respect to Stochastic Dominance

The most uncontroversial, widely accepted notions of riskiness are provided by the concepts of *stochastic dominance* (Hadar and Russell 1969; Hanoch and Levy 1969; Rothschild and Stiglitz 1970). 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 distributed like g that dominates g_* (in that sense).

An index Q is called *first- (second-) order monotonic* (M-FOD and M-SOD for short) 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, that is, to be both first- and second-order monotonic. And indeed, *the riskiness index* R is monotonic in both senses.

¹² We note that $w_1(g)$ is precisely the operational measure proposed in Foster and Hart (2007). Also, for $\gamma < 1$, w_γ may not be well defined: for instance, $w_{1/2}$ is not well defined for a half-half gamble between 7 and -1 .

¹³ Theorem C was independently discovered by and Dean Foster and Sergiu Hart (private communication).

D. Continuity

Call an index Q *continuous* if $Q(g_n) \rightarrow Q(g)$ whenever the g_n are uniformly bounded and converge to g in probability.¹⁴ With this definition, the riskiness index R is continuous; in words, when two gambles are likely to be close, their riskiness levels are close. Therefore, it is also continuous in weaker senses; for example, $R(g_n) \rightarrow R(g)$ whenever the g_n converge to g uniformly.¹⁵

E. Diluted Gambles

If g is a gamble, p a number strictly between zero and one, and g^p a compound gamble that yields g with probability p and zero 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 .

F. Compound Gambles

If two gambles g and h have the same riskiness r , then a compound gamble yielding g with probability p and h with probability $1 - p$ also has riskiness r . More generally,

(5.6.1) The riskiness of a compound of two gambles lies between their riskiness levels.

G. Normal Gambles

If the gamble g has a normal distribution,¹⁶ then $R(g) = \text{Var } g / 2Eg$, where Var stands for “variance.” Indeed, set $\text{Var } g =: \sigma^2$ and $Eg =: \mu$. The density of g ’s distribution is $e^{-(x-\mu)^2/2\sigma^2} / \sigma\sqrt{2\pi}$, so

¹⁴ That is, for every $\varepsilon > 0$, there is an N such that $\text{Prob}\{|g_n - g| > \varepsilon\} < \varepsilon$ for all $n > N$.

¹⁵ That is, for every $\varepsilon > 0$, there is an N such that $\sup |g_n - g| < \varepsilon$ for all $n > N$. In words, when two gambles are always close, their riskiness levels are close.

¹⁶ As defined in Sec. III, 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 $Ee^{-\alpha g}$ exists for all positive α .

$$\begin{aligned}
\mathbf{E}e^{-g/(\sigma^2/2\mu)} &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2} e^{-x/(\sigma^2/2\mu)} dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-[(x^2-2\mu x+\mu^2)+(4\mu x)]/2\sigma^2} dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x+\mu)^2/2\sigma^2} dx = 1.
\end{aligned}$$

So (3.1) holds with $R(g) := \sigma^2/2\mu$, so that is indeed the riskiness of g .

H. Sums of Gambles

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

$$1 = \mathbf{E}e^{-g/r}\mathbf{E}e^{-h/r} = \mathbf{E}(e^{-g/r}e^{-h/r}) = \mathbf{E}(e^{-(g+h)/r}),$$

so $R(g+h) = r$.

It follows that the sum of n i.i.d. gambles has the same riskiness as each one separately. This contrasts with the expectation—and the variance—of such a sum, which is n times the corresponding quantity for a single gamble. In the case of riskiness, one might say that location and dispersion considerations, which act in opposite directions, cancel each other out, and the result is that the riskiness stays the same.

More generally,

(5.8.1) If g and h are independent, then the riskiness of $g+h$ lies between those of g and h .

An interesting consequence is that a person or entity such as a pension fund that does not want its portfolio to exceed a certain level of riskiness need only see to it that each of the independent investments it makes does not exceed that level.

Even without independence, we still have *subadditivity*.¹⁷

(5.8.2) $R(g+h) \leq R(g) + R(h)$ for any gambles g and h .

Moreover, equality in (5.8.2) obtains when g is a positive multiple of h (that follows from homogeneity), and only then. We thus get a spectrum of circumstances, which is most transparent when the two gambles are

¹⁷ We thank Sergiu Hart for this observation and for its proof.

identically distributed and so have the same riskiness r : When the gambles are “totally” positively correlated (i.e., equal), the risks reinforce each other, and the sum has riskiness precisely $2r$. When they are independent, the risks neither reinforce nor hinder each other, and the sum has the same riskiness r as each of the gambles separately. When they are “totally” negatively correlated, the risk is minimal but need not vanish.

I. Extending the Domain

So far, riskiness is defined on the domain of “gambles”: random variables g with some negative values and $Eg > 0$. On this domain, the range of the riskiness levels is the positive reals, that is, strictly between zero and ∞ . Outside of this domain, the basic relation that determines riskiness—equation (3.1)—has no solution. The domain may be extended by defining $R(g) := 0$ when there are no negative values and $R(g) := \infty$ when $Eg \leq 0$. Intuitively, this makes good sense: When there are no negative values, there is no risk; and when $Eg \leq 0$, no risk-averse agent will accept g . When g vanishes identically, we have a “singular point,” where the riskiness remains undefined.

With these definitions, the index’s properties continue to apply. Thus it still respects first- and second-order stochastic dominance, though now only weakly.¹⁸ It is also “continuous,” under the usual meaning of “ $\rightarrow \infty$.” The other properties also apply, *mutatis mutandis*.

J. Emphasis on Losses

As we shall see in Section VII, the riskiness index R is much more sensitive to the loss side of a gamble than to its gain side. Technically, the reason is that the exponential on the right side of (3.1) 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 Secs. VIII.C–VIII.E) also emphasize loss. But there, the emphasis is built in; the definitions explicitly put more weight on the loss side. With 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 emerges naturally from the analysis.

¹⁸When g FOD g_* or g SOD g_* , one can now conclude only the weak inequality $R(g) \leq R(g_*)$.

K. *Financial Instruments as Multiplicative Gambles*

One may think of a financial instrument as a *multiplicative* gamble h , defined as a positive-valued random variable with expectation greater than one and some values that are less than one. An agent with wealth w who invests a capital of c in such an instrument¹⁹ will wind up with $w - c + hc = w + (h - 1)c$; he is thus taking a gamble g —in the ordinary additive sense considered heretofore—defined by $g := (h - 1)c$. By the homogeneity axiom, the riskiness of this gamble, or investment, is $R(g) = cR(h - 1)$. Thus when c dollars are invested, the riskiness per dollar of investment of the instrument h is $cR(h - 1)/c = R(h - 1)$, which is independent of the amount c invested. It is thus reasonable to define the *riskiness* of h as the dimensionless quantity $R(h - 1)$.

VI. Ordinality

If we are looking only for an ordinal index—that is, wish to define “riskier” without saying how much riskier—then we can replace the homogeneity axiom by conditions of monotonicity and continuity. An index Q for which $Q(g) > Q(h)$ if and only if $R(g) > R(h)$ is called *ordinally equivalent* to R . We have already seen that the riskiness index R satisfies the duality axiom (theorem A), is continuous (Sec. V.D), and is both first- and second-order monotonic (Sec. V.C). In the opposite direction, we have the following theorem.

THEOREM D. Any continuous and first-order monotonic index that satisfies the duality axiom is ordinally equivalent to R .

Moreover,

(6.1) Continuity, monotonicity, and duality are essential for theorem D.

Without any one of them, it fails.

VII. Some Numerical Examples

A. *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 $\lim_{M \rightarrow \infty} R(g_{M,l}) = l$.

By positive homogeneity, one may think of this as “calibrating” the

¹⁹ For example, IBM stock is a financial instrument, whereas buying 5,000 shares of IBM is an investment.

unit of riskiness: any gamble with riskiness \$1 is “as risky” as one in which the possible loss is \$1 and the possible gain is “very large,” where the loss probability is $1/e$ —the probability of “no success” in a Poisson distribution with mean one.

B. *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 ∞ to \$1), the riskiness goes up again, but not by much—only to \$1.64. If the gain decreases to \$1.10—so the expectation is only \$0.05—the riskiness jumps to \$11.01. As the gain approaches \$1—that is, the expectation approaches zero—the riskiness approaches ∞ . The riskiness of a half-half gamble yielding $-\$100$ or $\$105$ (Rabin 2000) is \$2,100.

C. *Insurance*

To buy insurance is to reject a gamble. For example, suppose that you insure a risk of losing \$20,000 with probability 0.001 for a premium of \$100, as when buying loss damage waiver in a car rental. Thus you end up with $-\$100$ for sure. If you decline the insurance, you are faced with a gamble that yields $-\$20,000$ with probability 0.001 and \$0 with probability 0.999. If we normalize²⁰ so that rejecting the gamble is worth \$0, then the gamble yields $-\$19,900$ with probability 0.001 and \$100 with probability 0.999. The riskiness of this gamble is \$7,491.

D. *Riskiness, Desirability, and Acceptance*

A riskier gamble need not be less desirable, even when both gambles have the same mean. For example, let g be a half-half gamble yielding $-\$3$ or $\$5$, and let h be a $7/8$ - $1/8$ gamble yielding $-\$1$ or $\$15$. The respective riskiness levels of g and h —both of which have mean one—are \$7.70 and \$9.20, but a CARA agent i with a sufficiently high parameter α will prefer the riskier gamble h ; he will essentially disregard the gains in both gambles and will prefer a loss that, though more likely, is smaller in magnitude. Indeed, i 's utilities for g and h are, respectively,

²⁰ 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 $w - \$100$.

$-\frac{1}{2}e^{3\alpha}[1 + o(1)]$ and $-\frac{7}{8}e^\alpha[1 + o(1)]$; for sufficiently high α , the second is higher than the first.

Moreover, there are even agents who accept the riskier gamble h and reject the less risky one g . For example, that is so at wealth zero for an agent with utility function $u(x) := \min(2x, x)$. To be sure, the function u is not twice continuously differentiable; but it can easily be modified so that it will be, without substantially affecting the example.

However, such an agent cannot have CARA. Indeed, as we have seen in (4.2.2), if a CARA agent rejects a gamble, then he rejects any riskier gamble.

VIII. The Literature

This section reviews other indices and compares them to ours. A prominent feature of many is that they are not monotonic with respect to first-order dominance; indeed, they may rate a gamble g riskier than h even though h is sure to yield more than g . The review is not exhaustive; we content ourselves with discussing some of the indices and briefly mentioning some others.

A. Measures of Dispersion

Pure measures of dispersion such as standard deviation, variance, mean absolute deviation ($E|g - E g|$), and interquartile range²¹ have been suggested as indices of riskiness; see the survey of Machina and Rothschild (2008). 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 rates $g + c$ precisely as risky as g , in spite of its being sure to yield more than g . According to Machina and Rothschild, an even stranger index is entropy,²² 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 obvious that such measures of dispersion cannot embody the economic, decision-making notion of riskiness set forth in Section II. As Hanoch and Levy (1970, 344) put it, "The identification of riskiness with variance, or with any other single measure of dispersion, is clearly unsound. There are many obvious cases where more dispersion is desirable, if it

²¹ Interquartile range is the difference between the first and third quartiles of the gamble's distribution. So, if g yields $-\$100$, $-\$1$, $\$2$, and $\$1,000$ with probability one-fourth each, then the interquartile range is $\$3$.

²² The formula for entropy is defined as $-\sum_k p_k \log_2 p_k$, where the p_k range over the probabilities of the gamble's different values.

is accompanied by an upward shift in the locations of the distribution, or by an increasing positive asymmetry.”

B. 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, e.g., Bodie, Kane, and Marcus 2002; Welch 2008). Specifically, any portfolio is associated with a gamble g ; the Sharpe ratio of the portfolio is defined²³ as μ/σ , where μ is the mean of g and σ its standard deviation. Portfolios with a smaller Sharpe ratio are considered riskier, so σ/μ —the reciprocal of the Sharpe ratio—might be considered an index of riskiness of the portfolio.

This index violates M-FOD. Indeed, 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. Then g has $\mu = 0.96$ and $\sigma = 0.28$, 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 more risky than g , though h FOD g . Moreover, when ε is positive but small, $h + \varepsilon$ is sure to yield more than g but is nevertheless rated riskier.

A final remark regarding normal gambles is of interest. As we said, the Sharpe ratio is viewed as a measure of risk-adjusted returns. If one takes the ratio of the mean μ to the riskiness index—which in this case equals $\sigma^2/2\mu$, by Section V.G—the result is $2\mu^2/\sigma^2$, which is ordinarily equivalent to the Sharpe ratio. Thus, the Sharpe ratio ranks normal gambles by their riskiness-adjusted expected returns. Matters are different for nonnormal gambles.

C. 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 percent confidence level, the VaR of a gamble g is the absolute value of its fifth percentile when that is nonpositive and zero otherwise. In words, it is the greatest possible loss, if we ignore losses with probability less than 5 percent. 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 percent VaR of $\$1$, and so does the gamble yielding $-\$1$ and $\$100,000$ with 0.06 and 0.94 probabilities.

This index depends on a parameter—the confidence level—whose “appropriate” value is not clear. Also, it ignores completely the gain side

²³ The standard definition looks more complicated but boils down to this.

of the gamble; in particular, it violates M-FOD. And even on the loss side, it concentrates only on that loss that “hits” the confidence level.

D. “Coherent” Measures of Risk

Artzner et al. (1999) call an index Q *coherent* if it satisfies five axioms: (i) positive homogeneity, (ii) subadditivity, (iii) weak first-order monotonicity, (iv) “relevance,” and (v) “translation invariance.” Axiom i is as in our Section III; ii is our (5.8.2) (except that with us it follows from the axioms, whereas they assume it). Their iii says that if $g \geq h$ identically, then $Q(g) \leq Q(h)$, which for us follows from first-order monotonicity (Sec. V.C). Thus our index obeys their first three axioms. Their iv concerns “gambles” with no positive values, which we exclude.²⁴ Their axiom v says that if c is a constant, then $Q(g + c) = Q(g) - c$, which is not so for our index.

Like ours, their indices measure risk in dollars. But their five axioms are very far from determining an index. Indeed, for any family of probability measures μ on the underlying probability space, the supremum of $E_\mu(-g)$ over the family is a coherent index. One example is $|\min g|$, which violates first- and second-order monotonicity and also continuity; but there are very many others. All these indices violate our duality axiom.

E. Additional Indices

Brachinger and Weber (1997) and Brachinger (2002) are good surveys of the psychological literature. Like VaR and the “coherent” measures, their measures of perceived risk take the form of families rather than proposing a single index. The studies include Coombs (1969), Pollatsek and Tversky (1970), Fishburn (1977, 1982, 1984), Luce (1980), Sarin (1987), Luce and Weber (1988), and Jia, Dyer, and Butler (1999).

Of all these, Sarin’s measure $S(g) := Ee^{-g}$ is the closest to our index R . This is monotonic with respect to FOD, so it must violate duality. 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 that S violates duality, set $\alpha := 1/R(g)$. By (4.2.2), a CARA agent i with parameter $\frac{5}{6}\alpha$ accepts g ; whereas a CARA agent j with parameter $\frac{2}{3}\alpha$ —who is less risk averse than i —rejects $2g$, which is rated less risky than g by S . So S violates duality. It also violates positive homogeneity.

Finally, we mention the index recently proposed by Foster and Hart (2007), which is closely related to ours but is based on somewhat dif-

²⁴ Unless the domain is extended (Sec. V.I), when the riskiness is $+\infty$.

ferent ideas. While appealing in many ways and connecting with our operational concerns of Section IV.C, it is not continuous: the riskiness it ascribes to a gamble is at least equal to the maximum possible loss under the gamble, no matter how small the probability of that loss might be.

IX. Discussion

A. *Axiomatics*

We have used axioms—chiefly, the duality axiom—to characterize the proposed index. Of course, this does not imply that the resulting index is in any sense the only possible or reasonable one. On the contrary, the idea is to capture the “content” of the definition in conceptual terms rather than just writing down a formula; different concepts lead to different indices.²⁵

Already at the outset of our research, in 2004, we felt “in our guts” that the concept we were seeking could be based on CARA (see theorem B in Sec. IV.B), roughly, for the reasons set forth below (Sec. IX.B). But to pin this down—express our gut feeling precisely—we sought an axiomatic characterization.

B. *Intuition*

An important underlying intuition that we were trying to capture is that the riskiness of a gamble should be independent of the person considering that gamble. That is precisely the point made in the introduction: that whether or not a person accepts a gamble depends on both (i) the attributes of the gamble—in particular, its riskiness—and (ii) the attributes of the person—in particular, his aversion to risk. What we are trying to do here is to separate these two. So one can say “hey, I measured the riskiness of this gamble; it’s too risky for me” or the other attitudes expressed in Section II.

In particular, the index should have nothing to do with wealth. It is widely assumed in economics that wealthier people are less risk averse; that is the idea behind CRRA (see Sec. IV.C). So wealth is related to risk aversion—item ii above—which we want to avoid considering in measuring riskiness—item i.

However, our approach is based on the idea that riskiness is the “dual” of risk aversion—“what risk-averters hate” (Machina and Rothschild 2008). So risk aversion—specifically, utilities—must somehow enter the analysis after all. Still, as we have seen, wealth should be irrelevant.

²⁵ For example, Foster and Hart have axiomatized their (2007) index (personal communication); see also Sec. VIII.D above.

That wealth is “irrelevant” for a decision maker means that any gamble is either accepted by him at all wealth levels or rejected by him at all wealth levels. Now that is the case if and only if the decision maker has CARA; see (4.2.1). Thus, as explained in Section IV.B, CARA provides an appropriate tool to express the idea that riskiness depends on the gamble only, not on the agent considering the gamble, in particular, not on his wealth.

So far, these are merely intuitions. To bring them to fruition—achieve a coherent development—one needs a rigorous definition of comparative risk aversion, which also encapsulates the above ideas. This is accomplished by the \succcurlyeq and \succ relations. The uniform element of these relations expresses the irrelevance of wealth discussed above.

Practically speaking, the \succcurlyeq relation is characterized by (4.1.2). In particular, it is in general not reflexive; indeed, $i \succcurlyeq i$ if and only if i has CARA.

C. Further Research

Several avenues for further research present themselves. For one thing, we have calculated the riskiness of normally distributed gambles (Sec. V.G); it would be desirable to investigate the riskiness of gambles whose distributions have various other widely used forms.

Another avenue of research is to try to extend the approach to “gambles” with negative expected value, which would apply to risk lovers. This could shed light on gambling behavior, for instance, and on high-risk “venture capital.”

X. Proofs

A. Preliminaries

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

$$u_i(0) = u_j(0) = 0 \quad \text{and} \quad u'_i(0) = u'_j(0) = 1. \quad (1)$$

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),

$$\begin{aligned} \log u'_i(y) &= \log u'_i(y) - \log u'_i(0) = \int_0^y [\log u'_i(z)]' dz = \int_0^y \left[\frac{u''_i(z)}{u'_i(z)} \right] dz \\ &= \int_0^y -\rho_i(z) dz < \int_0^y -\rho_j(z) dz = \log u'_j(y). \end{aligned}$$

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

$$u_i(w) = - \int_0^{|w|} u'_i(y) dy < - \int_0^{|w|} u'_j(y) dy = u_j(w).$$

QED

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. Similar to that of lemma 2, with i and j interchanged, strict inequalities replaced by weak inequalities, and the restriction to $|w| < \delta$ eliminated. QED

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. Without loss of generality,²⁶ $w_i = w_j = 0$, 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,

$$u_i(w) < u_j(w) \text{ whenever } |w| < \delta \text{ and } w \neq 0. \tag{5}$$

Choose ε 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),

$$f_i(x) < f_j(x) \text{ for all } x. \tag{6}$$

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 be-

²⁶ For arbitrary w_i and w_j , define $u_i^*(x) := [u_i(x + w_i) - u_i(w_i)]/u'_i(w_i)$ and u_j^* similarly, and apply the current reasoning to u_i^* and u_j^* .

tween zero and ε , since f_i is continuous. So by (6), $f_j(y) > 0$. 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 $Eu_j(g) = f_j(y - \eta) > 0 > f_i(y - \eta) = Eu_i(g)$. So j accepts g whereas i rejects it. QED

B. *Proof of²⁷ (4.1.1)*

“Only if”: Assume that i is no less risk averse than j ; we must show

$$\rho_i(w) \geq \rho_j(w) \quad \text{for all wealth levels } w. \quad (7)$$

If not, then there is a w with $\rho_i(w) < \rho_j(w)$. So by lemma 4, there is a gamble that i accepts at w and j rejects at w , contradicting i being less risk averse than j . So (7) is proved.

“If”: Assume (7); we must show that for each wealth level w and gamble g , if i accepts g at w , then j accepts g at w . Without loss of generality, $w = 0$, so we must show that

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

From (1), (7), and corollary 3 (with i and j reversed), we conclude that $u_j(w) \geq u_i(w)$ for each w . So $Eu_j(g) \geq Eu_i(g)$, which yields (8). QED

C. *Proof of (4.1.2)*

Statement (4.1.2) follows from (4.1.1) by shifting the independent variable on one of the utilities to make $w_i = w_j$. QED

D. *Proof of Theorem A*

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

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

To see that there is a unique $R(g) > 0$ satisfying (3.1), set $f(\alpha) := Ee^{-\alpha g} - 1$, and note that f is convex, $f(0) = 0$, $f'(0) = -Eg < 0$, and $f(M) > 0$ for M sufficiently large. So there is a unique $\gamma > 0$ with $f(\gamma) = 0$, and we set $R(g) := 1/\gamma$.

To see that R satisfies the duality axiom, let i, j, g, h , and w be as in

²⁷ Statements (4.1.1) and (4.1.2) are needed in the proof of theorem A, so we prove them first.

the hypothesis of that axiom; without loss of generality, $w = 0$. Set $\gamma := 1/R(g)$, $\eta := 1/R(h)$, $\alpha_i := \inf \rho_p$, and $\alpha_j := \sup \rho_j$. Thus

$$Eu_\gamma(g) = \frac{1 - Ee^{-\gamma g}}{\gamma} = 0 \quad \text{and} \quad Eu_\eta(h) = \frac{1 - Ee^{-\eta h}}{\eta} = 0. \quad (10)$$

By hypothesis, $R(g) > R(h)$, so $\eta > \gamma$. By corollary 3,

$$u_i(x) \leq u_{\alpha_i}(x) \quad \text{and} \quad u_{\alpha_j}(x) \leq u_j(x) \quad \text{for all } x. \quad (11)$$

Now assume $Eu_i(g) > 0$; we must prove that $Eu_j(h) > 0$. From $Eu_i(g) > 0$ and (11), it follows that $Eu_{\alpha_i}(g) > 0$. So by (10), $Eu_\gamma(g) = 0 < Eu_{\alpha_i}(g)$. So by (9), $\gamma > \alpha_i$. By (4.1.2), $\alpha_i \geq \alpha_j$, so $\eta > \gamma$ yields $\alpha_j < \eta$. Then (10), (9), and (11) yield $0 = Eu_\eta(h) < Eu_{\alpha_j}(h) < Eu_j(h)$, so indeed R satisfies the duality axiom. That R is positively homogeneous 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

$$Q \text{ is ordinally equivalent to } R. \quad (12)$$

If this is not true, then there must exist g and h that are ordered differently by Q and R . This means either that the respective orderings are reversed, that is,

$$Q(g) > Q(h) \quad \text{and} \quad R(g) < R(h), \quad (13)$$

or that equality holds for exactly one of the two indices, that is,

$$Q(g) > Q(h) \quad \text{and} \quad R(g) = R(h) \quad (14)$$

or

$$Q(g) = Q(h) \quad \text{and} \quad R(g) > R(h). \quad (15)$$

If either (14) or (15) holds, then by homogeneity, replacing g by $(1 - \varepsilon)g$ for sufficiently small positive ε leads to reversed inequalities. So without loss of generality we may assume (13).

Now let $\gamma := 1/R(g)$ and $\eta := 1/R(h)$; then (10) holds. By (13), $\gamma > \eta$. Choose μ and ν 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), $Eu_\mu(g) > Eu_\gamma(g) = 0$ and $Eu_\nu(h) < Eu_\eta(h) = 0$. So if i and j have utility functions u_μ and u_ν , respectively, then i accepts g and j rejects h . But from $\mu > \nu$ and (4.1.2), it follows that $i \ni j$, 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. QED

Needless to say, both duality and positive homogeneity are essential to theorem A. Thus, the mean Eg is positively homogeneous but violates duality; and the index $[R(g)]$, where $[x]$ denotes the integer part of x , satisfies duality but is not positively homogeneous. Neither Eg nor $[R(g)]$ is even ordinally equivalent to R .

E. Proof of (4.2.1)

“Only if”: All CARA utility functions have the form $-e^{-\alpha x}$. Thus i accepts g at wealth w if and only if $-Ee^{-\alpha(g+w)} > -e^{-\alpha w}$, that is, if and only if $Ee^{-\alpha g} < 1$; and this condition does not depend on w .²⁸

“If”: Suppose that i 's Arrow-Pratt index of ARA 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 $\rho(w) = \lim_{\delta \rightarrow 0} [p_\delta(w) - \frac{1}{2}]/\delta$,²⁹ that is, since even-money half-half bets are always rejected by risk-averse utility maximizers, the Arrow-Pratt index is the probability premium over one-half, per dollar, that is needed for i to be indifferent between taking and not taking a small even-money gamble. So, if δ is sufficiently small, $q - \frac{1}{2}$ lies halfway 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. QED

F. Proof of (4.2.2)

Let g_1 be a gamble and g_2 a riskier gamble. For $\alpha \geq 0$ and $i = 1, 2$, set $f_i(\alpha) := Ee^{-\alpha g_i} - 1$. We saw (near the start of the proof of theorem A) that $f_i(0) = 0$ and $f_i(\alpha) < 0$ when $0 < \alpha < 1/R(g_i)$, $f_i(1/R(g_i)) = 0$, and $f_i(\alpha) > 0$ when $\alpha > 1/R(g_i)$. So a CARA agent with parameter α accepts g_i if and only if $f_i(\alpha) < 0$, that is, if and only if $\alpha \in (0, 1/R(g_i))$. Since $1/R(g_2) < 1/R(g_1)$, it follows that if the agent rejects g_1 , then he also rejects g_2 , as was to be proved. This proves the second sentence and so the whole assertion. QED

G. Proof of (4.3.2)

To prove the first sentence, let u_i be i 's utility, and define a utility u_j as follows: when x is between $w + \min g$ and $w + \max g$, define $u_j(x) :=$

²⁸ Pratt (1964, 130) makes a similar argument for preferences between gambles.

²⁹ See, e.g., Aumann and Kurz (1977, sec. 6); but there may well be earlier sources.

$u_i(x)$; when $x \leq w + \min g$, define $u_j(x)$ to equal a CARA utility with parameter $\rho_i(w + \min g)$ and $u_j(w + \min g) = u_i(w + \min g)$ and $u'_j(w + \min g) = u'_i(w + \min g)$; when $x \geq w + \max g$, define $u_j(x)$ to equal a CARA utility with parameter $\rho_i(w + \max g)$ and $u_j(w + \max g) = u_i(w + \max g)$ and $u'_j(w + \max g) = u'_i(w + \max g)$. Let u_k be a CARA utility with parameter $[1/R(g)] - \varepsilon$. Then

$$\min_x \rho_k(x) > \max_x \rho_j(x) \tag{16}$$

for positive ε sufficiently small. By theorem B (Sec. IV.B), a CARA person with parameter $1/R(g)$ is indifferent between taking and not taking g . Therefore, k , who is less risk averse, accepts g . So by (16) and (4.2.2), j also accepts g . But between the minimum and maximum of $w + g$, the utilities of i and j are the same. So i accepts g at w . This proves the first sentence of (4.3.2); the proof of the second sentence is similar. QED

H. Proof of the Claims in Section V.C

For $\alpha \geq 0$, set $f(\alpha) := Ee^{-\alpha g}$ and $f_*(\alpha) := Ee^{-\alpha g_*}$. If g FOD g_* , then $f(\alpha) < f_*(\alpha)$ whenever $\alpha > 0$. From this and the proof that (3.1) has a unique positive root,³⁰ 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 SOD 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. QED

I. Proof of the Claim in Section V.D

For $\alpha \geq 0$, set $f(\alpha) := Ee^{-\alpha g}$ and $f_n(\alpha) := Ee^{-\alpha g_n}$, denote the unique positive root of $f = 1$ by γ and that of $f_n = 1$ by γ_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 γ_n have a limit point $\gamma_* \neq \gamma$; arguing by contradiction, we may assume without loss of generality 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 $f_n(\gamma_n) = 1$. QED

J. Proof of (5.6.1)

Denote by $g^p \oplus h^{1-p}$ the compound gamble that yields g with probability p and h with probability $1 - p$. By theorem A, the riskiness

³⁰ Near the beginning of the proof of theorem A.

$R(g^p \oplus h^{1-p})$ is the reciprocal of the unique positive root of $f = 1$, where

$$f(\alpha) := \mathbb{E}e^{-\alpha(g^p \oplus h^{1-p})} = p\mathbb{E}e^{-\alpha g} + (1-p)\mathbb{E}e^{-\alpha h}.$$

So if $f(\alpha) = 1$, then it cannot be that both $\mathbb{E}e^{-\alpha g} > 1$ and $\mathbb{E}e^{-\alpha h} > 1$, and it cannot be that both $\mathbb{E}e^{-\alpha g} < 1$ and $\mathbb{E}e^{-\alpha h} < 1$. So $\mathbb{E}e^{-\alpha g} \leq 1$ and $\mathbb{E}e^{-\alpha h} \geq 1$, say. So $1/R(g^p \oplus h^{1-p}) = \alpha \leq 1/R(g)$ and similarly $1/R(g^p \oplus h^{1-p}) = \alpha \geq 1/R(h)$. Thus $R(g) \leq R(g^p \oplus h^{1-p}) \leq R(h)$, as asserted. QED

K. Proof of (5.8.1)

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

L. Proof of (5.8.2)

Set $r := R(g)$, $r' := R(h)$, and $\lambda := r/(r+r') \in (0, 1)$. Then $(g+h)/(r+r') = \lambda(g/r) + (1-\lambda)(h/r')$; so from (3.1) and the convexity of the exponential, we get $\mathbb{E}e^{-(g+h)/(r+r')} \leq \lambda \mathbb{E}e^{-g/r} + (1-\lambda)\mathbb{E}e^{-h/r'} = 1$, so $r+r' \leq R(g+h)$ (see the second paragraph in the proof of theorem A), as asserted. QED

M. Proof of Theorem D

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 ε leads to reversed inequalities; this follows from first-order monotonicity and continuity. The remainder of the proof of (12) is as above. QED

N. Proof of (6.1)

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 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 nonempty 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 VIII, Sarin’s index $S(g)$ is continuous and first-order monotonic but violates duality. QED

XI. Conclusion

We have defined a numerical index of the *riskiness* of a gamble with stated dollar outcomes and stated probabilities. It is denominated in dollars, is monotonic with respect to first- and second-order stochastic dominance, is continuous in about any sense one wishes, is positively homogeneous, and satisfies a duality condition that says, roughly, that agents who are uniformly more risk averse are less likely to accept gambles that are riskier. It is the only index satisfying these conditions. Moreover, the index may be characterized in terms of constant Arrow-Pratt risk aversion—both absolute (CARA) and relative (CRRA)—and it may be used to define a (dimensionless) index of riskiness for financial instruments such as stocks or bonds.

References

- Arrow, Kenneth J. 1965. *Aspects of the Theory of Risk-Bearing*. Helsinki: Yrjö Jahns-sonin Säätiö.
- . 1971. *Essays in the Theory of Risk Bearing*. Chicago: Markham.
- Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. “Coherent Measures of Risk.” *Math. Finance* 9:203–28.
- Aumann, Robert J., and Mordecai Kurz. 1977. “Power and Taxes.” *Econometrica* 45:1137–61.
- Bernoulli, Daniel. [1738] 1954. “Exposition of a New Theory on the Measurement of Risk.” *Econometrica* 22:23–36.
- Bodie, Zvi, Alex Kane, and Alan J. Marcus. 2002. *Investments*. 5th ed. New York: McGraw-Hill.
- Brachinger, Hans W. 2002. “Measurement of Risk.” In *Optimization and Operations Research*, edited by U. Derigs. Encyclopedia of Life Support Systems (EOLSS), UNESCO. Oxford: EOLSS Pubs.

- Brachinger, Hans W., and Martin Weber. 1997. "Risk as a Primitive: A Survey of Measures of Perceived Risk." *Operations Res. Spectrum* 19:235–50.
- Coombs, Clyde H. 1969. "Portfolio Theory: A Theory of Risky Decision Making." In *La decision*, edited by G. T. Guilbaud. Paris: Centre National de la Recherche Scientifique.
- Diamond, Peter A., and Joseph E. Stiglitz. 1974. "Increases in Risk and Risk Aversion." *J. Econ. Theory* 8:337–60.
- Fishburn, Peter C. 1977. "Mean-Risk Analysis with Risk Associated with Below Target Returns." *A.E.R.* 67:116–26.
- . 1982. "Foundations of Risk Measurement (II): Effects of Gains on Risk." *J. Math. Psychology* 25:226–42.
- . 1984. "Foundations of Risk Measurement (I): Risk as Probable Loss." *Management Sci.* 30:396–406.
- Foster, Dean P., and Sergiu Hart. 2007. "An Operational Measure of Riskiness." Discussion Paper no. 454, Center Study Rationality, Hebrew Univ. Jerusalem.
- Hadar, Josef, and William R. Russell. 1969. "Rules for Ordering Uncertain Prospects." *A.E.R.* 59:25–34.
- Hanoch, Giora, and Haim Levy. 1969. "The Efficiency Analysis of Choices Involving Risk." *Rev. Econ. Studies* 36:335–46.
- . 1970. "Relative Effectiveness of Efficiency Criteria for Portfolio Selection." *J. Financial and Quantitative Analysis* 5:63–76.
- Jia, Jianmin, James Dyer, and John Butler. 1999. "Measures of Perceived Risk." *Management Sci.* 45:519–32.
- Luce, R. Duncan. 1980. "Several Possible Measures of Risk." *Theory and Decision* 12:217–28. (Correction: *Theory and Decision* 13 [1981]: 381.)
- Luce, R. Duncan, and Elke U. Weber. 1988. "An Axiomatic Theory of Conjoint Expected Risk." *J. Math. Psychology* 30:188–205.
- Machina, Mark, and Michael Rothschild. 2008. "Risk." In *The New Palgrave Dictionary of Economics*, 2nd ed., edited by Steven N. Durlauf and Lawrence E. Blume. London: Macmillan.
- Palacios-Huerta, Ignacio, and Roberto Serrano. 2006. "Rejecting Small Gambles under Expected Utility." *Econ. Letters* 91:250–59.
- Palacios-Huerta, Ignacio, Roberto Serrano, and Oscar Volij. 2004. "Rejecting Small Gambles under Expected Utility." Manuscript, Dept. Econ., Brown Univ.
- Pollatsek, Alexander, and Amos Tversky. 1970. "A Theory of Risk." *J. Math. Psychology* 7:540–53.
- Pratt, John. 1964. "Risk Aversion in the Small and in the Large." *Econometrica* 32:122–36.
- Rabin, Matthew. 2000. "Risk Aversion and Expected Utility: A Calibration Theorem." *Econometrica* 68:1281–92.
- Rothschild, Michael, and Joseph E. Stiglitz. 1970. "Increasing Risk: I. A Definition." *J. Econ. Theory* 2:225–43.
- Sarin, Rakesh K. 1987. "Some Extensions of Luce's Measures of Risk." *Theory and Decision* 22:125–41.
- Tobin, James. 1969. "Comment on Borch and Feldstein." *Rev. Econ. Studies* 36: 13–14.
- Welch, Ivo. 2008. *Corporate Finance: An Introduction*. Upper Saddle River, NJ: Prentice-Hall.