Dual Moments and Risk Attitudes*

Louis R. Eeckhoudt
IÉSEG School of Management
Catholic University of Lille
and CORE
Louis.Eeckhoudt@fucam.ac.be

Roger J. A. Laeven
Amsterdam School of Economics
University of Amsterdam, EURANDOM
and CentER
R.J.A.Laeven@uva.nl

This Version: March 14, 2018

Abstract

In decision under risk, the primal moments of mean and variance play a central role to define the local index of absolute risk aversion. In this paper, we show that in canonical non-EU models dual moments have to be used instead of, or on par with, their primal counterparts to obtain an equivalent index of absolute risk aversion.

Keywords: Risk Premium; Expected Utility; Dual Theory; Rank-Dependent Utility; Local Index; Absolute Risk Aversion.

JEL Classification: D81.

OR/MS Classification: Decision analysis: Risk.

*We are very grateful to Loic Berger (discussant), Sebastian Ebert, Glenn Harrison, Richard Peter (discussant), Nicolas Treich, Ilia Tsetlin, Bob Winkler, and, in particular, to Christian Gollier for many detailed comments and suggestions and to Harris Schlesinger (†) for discussions. We are also grateful to conference and seminar participants at the CEAR/MRIC Behavioral Insurance Workshop in Munich and City University of London for their comments. This research was funded in part by the Netherlands Organization for Scientific Research under grant NWO VIDI 2009 (Laeven). Research assistance of Andrei Lalu is gratefully acknowledged. Eeckhoudt: Catholic University of Lille, IÉSEG School of Management, 3 Rue de la Digue, Lille 59000, France. Laeven: University of Amsterdam, Amsterdam School of Economics, PO Box 15867, 1001 NJ Amsterdam, The Netherlands.
1 Introduction

In their important seminal work, Pratt [22] and Arrow [2, 3] (henceforth, PA) show that under expected utility (EU) the risk premium $\pi$ associated to a small risk $\bar{\varepsilon}$ with zero mean can be approximated by the following expression:

$$\pi \approx m_2 \left( - \frac{U''(w_0)}{U'(w_0)} \right).$$  \hspace{1cm} (1.1)

Here, $m_2$ is the second moment about the mean (i.e., the variance) of $\bar{\varepsilon}$ while $U'(w_0)$ and $U''(w_0)$ are the first and second derivatives of the utility function of wealth $U$ at the initial wealth level $w_0$. In the PA-approach, the designation “small” refers to a risk that has a probability mass equal to unity but a small variance. The PA-approximation in (1.1) provides a very insightful dissection of the EU risk premium, disentangling the complex interplay between the probability distribution of the risk, the decision-maker’s risk attitude, and his initial wealth. This well-known result has led to many developments and applications within the EU model in many fields.

The aim of this paper is to show that a similar result can also be obtained outside EU, in the dual theory of choice under risk (DT; Yaari [36]) and, more generally and behaviorally more relevant, under rank-dependent utility (RDU; Quiggin [25]). To achieve this, we substitute or complement the primal second moment $m_2$ by its dual counterpart, and substitute or complement the derivatives of the utility function $U$ by the respective derivatives of the probability weighting function. This modification enables us to develop for these two canonical non-EU models a simple and intuitive local index of risk attitude that resembles the one in (1.1) under EU. Our results allow for quite arbitrary utility and probability weighting functions including inverse $s$-shaped functions such as the probability weighting functions in Prelec [24] and Wu and Gonzalez [33], which are descriptively relevant (Abdellaoui [1]). Thus, we allow for violations of strong risk aversion (Chew, Karni and Safra [7] and Roell [26]) in the sense of aversion to mean-preserving spreads à la Rothschild and Stiglitz [28] (see also Machina and Pratt [18]).

1 For ease of exposition, we assume $U$ to be twice continuously differentiable, with positive first derivative.

2 Dual moments are sometimes referred to as mean order statistics in the statistics literature; see Section 2 for further details.

3 The RDU model encompasses both EU and DT as special cases and is at the basis of (cumulative) prospect theory (Tversky and Kahneman [32]). According to experimental evidence collected by Harrison and Swarthout [16], RDU seems to emerge even as the most important non-EU preference model from a descriptive perspective.

4 In a very stimulating strand of research, Chew, Karni and Safra [7] and Roell [26] have developed the “global” counterparts of the results presented here; see also the more recent Chateauneuf, Cohen and Meilijson [4, 5] and Ryan [29]. Surprisingly, the “local” approach has received no attention under DT and RDU, except—to the best of our knowledge—for a relatively little used paper by Yaari [35]. Specifically, Yaari exploits a uniformly ordered local quotient of derivatives (his Definition 4) with the aim to establish global results, restricting attention to DT. Yaari does not analyze the local behavior of the risk premium nor does he make a reference to dual moments. For global measures of risk aversion under prospect theory, we refer to Schmidt and Zank [30].

5 The insightful Nau [21] proposes a significant generalization of the PA-measure of local risk aversion in
Our paper is organized as follows. In Section 2 we define the second dual moment and we use it in Section 3 to develop the local index of absolute risk aversion under DT. In Section 4 we extend our results to the RDU model. Section 5 discusses related literature and Section 6 illustrates our results in examples. In Section 7 we present an application to portfolio choice and we provide a conclusion in Section 8. Some supplementary material, including some technical details to supplement Sections 3 and 4, the proof of a result in Section 5, and two illustrations to supplement Section 6, suppressed in this version to save space, is contained in an online appendix.

2 The Second Dual Moment

The second dual moment about the mean of an arbitrary risk \( \tilde{\varepsilon} \), denoted by \( \bar{m}_2 \), is defined by

\[
\bar{m}_2 := \mathbb{E} \left[ \max \left( \tilde{\varepsilon}^{(1)}, \tilde{\varepsilon}^{(2)} \right) \right] - \mathbb{E} [\tilde{\varepsilon}],
\]

where \( \tilde{\varepsilon}^{(1)} \) and \( \tilde{\varepsilon}^{(2)} \) are two independent copies of \( \tilde{\varepsilon} \). The second dual moment can be interpreted as the expectation of the largest order statistic: it represents the expected best outcome among two independent draws of the risk.\(^6\)

Our analysis will reveal that for an RDU maximizer who evaluates a small zero-mean risk, the second dual moment stands on equal footing with the variance as a fundamental measure of risk. While the variance provides a measure of risk in the “payoff plane”\(^7\) the second dual moment can be thought of as a measure of risk in the “probability plane”. Indeed, for a risk \( \tilde{\varepsilon} \) with cumulative distribution function \( F \), so

\[
m := \mathbb{E} [\tilde{\varepsilon}] = \int x \, dF(x),
\]

we have that

\[
m_2 = \int (x - m)^2 \, dF(x), \quad \text{while} \quad \bar{m}_2 = \int (x - m) \, d(F(x))^2.
\]

another direction. He considers the case in which probabilities may be subjective, utilities may be state-dependent, and probabilities and utilities may be inseparable, by invoking Yaari’s \(^{34}\) elementary definition of risk aversion of “payoff convex” preferences, which agrees with the Rothschild and Stiglitz \(^{28}\) concept of aversion to mean-preserving spreads under EU.\(^8\)

The definition and interpretation of the 2-nd dual moment readily generalize to the \( n \)-th order, \( n \in \mathbb{N}_{>0} \), by considering \( n \) copies.\(^7\)

We refer to Meyer \(^{19}\) and Eichner and Wagener \(^{12}\) for insightful comparative statics results on the mean-variance trade-off and its compatibility with EU.\(^7\)

Formally, our integrals with respect to functions are Riemann-Stieltjes integrals. If the integrator is a cumulative distribution function of a discrete (or non-absolutely continuous) risk, or a function thereof, then the Riemann-Stieltjes integral does not in general admit an equivalent expression in the form of an ordinary Riemann integral.\(^8\)
For the sake of brevity and in view of (2.1), we shall term the second dual moment about the mean, \( \bar{m}^2 \), the maxiance by analogy to the variance. Our designation “small” in “small zero-mean risk” will quite naturally refer to a risk with small maxiance under DT and to a risk with both small variance and small maxiance under RDU.

One readily verifies that for a zero-mean risk \( \tilde{\varepsilon} \)
\[
\mathbb{E} \left[ \max \left( \tilde{\varepsilon}^{(1)}, \tilde{\varepsilon}^{(2)} \right) \right] = -\mathbb{E} \left[ \min \left( \tilde{\varepsilon}^{(1)}, \tilde{\varepsilon}^{(2)} \right) \right].
\]
The miniance—the expected worst outcome among two independent draws—is perhaps a more natural measure of “risk”, but agrees with the maxiance for zero-mean risks upon a sign change.

Just like the first and second primal moments occur under EU when the utility function is linear and quadratic, the first and second dual moments correspond to a linear and quadratic probability weighting function under DT. For further details on mean order statistics and their integral representations we refer to David [8]. In the stochastic dominance literature, these expectations of order statistics and their higher-order generalizations arise naturally in an interesting paper by Muliere and Scarsini [20], when defining a sequence of progressive \( n \)-th degree “inverse” stochastic dominances by analogy to the conventional stochastic dominance sequence (see Ekern [13] and Fishburn [14]).

3 Local Risk Aversion under the Dual Theory

Consider a DT decision-maker. His evaluation of a risk \( A \) with cumulative distribution function \( F \) is given by
\[
\int x \, dh(F(x)), \tag{3.1}
\]
\[\text{This is easily seen from the Riemann-Stieltjes representations of the miniance and maxiance. Indeed,}
\]
\[
-\mathbb{E} \left[ \min \left( \tilde{\varepsilon}^{(1)}, \tilde{\varepsilon}^{(2)} \right) \right] = \int x \, d \left( 1 - F(x)^2 \right) = -2 \int x \, dF(x) + \int x \, d \left( F(x)^2 \right) = \mathbb{E} \left[ \max \left( \tilde{\varepsilon}^{(1)}, \tilde{\varepsilon}^{(2)} \right) \right],
\]
where the last equality follows because \( \int x \, dF(x) = 0 \) when \( \tilde{\varepsilon} \) is a zero-mean risk.

In a related strand of the literature, Eeckhoudt and Schlesinger [9] (see also Eeckhoudt, Schlesinger and Tsetlin [10]) and Eeckhoudt, Laeven and Schlesinger [11] derive simple nested classes of lottery pairs to sign the \( n \)-th derivative of the utility function and probability weighting function, respectively. Their approach can be seen to control the primal moments for EU and the dual moments for DT.

Expressions similar (but not identical) to dual moments also occur naturally in decision analysis applications. For example, the expected value of information when the information will provide one of two signals is the maximum of the two posterior expected values (e.g., payoffs or utilities) minus the highest prior expected value. This generalizes to the case of \( n > 2 \) possible signals. See Smith and Winkler [31] for a related problem.
where the probability weighting (distortion) function \( h : [0, 1] \rightarrow [0, 1] \) satisfies the usual properties (\( h(0) = 0, h(1) = 1, h'(p) > 0 \))\(^{12}\)^\(^{13}\).

In order to develop the local index of absolute risk aversion under DT we start from a lottery \( A \) given by the following representation.\(^{14}\)^\(^{15}\)

\[ A \]
\[ \begin{array}{ccc}
  \rightarrow & -1 & \leftarrow \\
  p_0 & & \\
  \downarrow & & \\
  1 - p_0 & & 1
\end{array} \]

We transform lottery \( A \) into a lottery \( B \) given by: \(^{16}\)

\[ B \]
\[ \begin{array}{ccc}
  \rightarrow & -1 & \leftarrow \\
  \hspace{1cm} p_0 - \varepsilon_1 & & \\
  \hspace{1cm} 2\varepsilon_1 & & \\
  \downarrow & & \\
  1 - p_0 - \varepsilon_1 & & 1
\end{array} \]

To obtain \( B \) from \( A \) we subtract a probability \( \varepsilon_1 \) from the probabilities of both states of the world in \( A \) without changing the outcomes and we assign these two probabilities jointly, i.e., \( 2\varepsilon_1 \), to a new intermediate state to which we attach an outcome \( x \) with \(-1 < x < 1\). If \( x \equiv 0 \), then \( E[A] = E[B] \) and \( B \) is a mean-preserving contraction of \( A \).

The value of \( x \) will be chosen such that the decision-maker is indifferent between \( A \) and \( B \). Naturally the difference between 0 and \( x \), denoted by \( \rho = 0 - x \), represents the risk premium associated to the risk change from \( A \) to \( B \). As we will show in Section 5.2 this definition of the risk premium can be viewed as a natural generalization of the PA risk premium to the

\(^{12}\)For ease of exposition, we assume \( h \) to be twice continuously differentiable.

\(^{13}\)Rather than distorting “decumulative” probabilities (as in Yaari), we adopt the convention to distort cumulative probabilities. Our convention ensures that aversion to mean-preserving spreads corresponds to \( h'' < 0 \) (i.e., concavity) under DT, just like it corresponds to \( U'' < 0 \) under EU, which facilitates the comparison. Should we adopt the convention to distort decumulative probabilities, the respective probability weighting function \( \tilde{h}(p) := 1 - \tilde{h}(1 - p) \) would be convex when \( h \) is concave.

\(^{14}\)In all figures, values along (at the end of) the arrows represent probabilities (outcomes).

\(^{15}\)Of course, we assume \( 0 < p_0 < 1 \).

\(^{16}\)We assume \( 0 < \varepsilon_1 < \min\{p_0, 1 - p_0\} \).
case of risk changes with probability mass less than unity. Depending on the shape of $h$ the risk premium $\rho$ may be positive or negative. If (and only if) $h''(p) < 0$, the corresponding DT maximizer is averse to mean-preserving spreads, and would universally prefer $B$ over $A$ when $x$ were 0. Thus, to establish indifference between $A$ and $B$ for such a decision-maker, $x$ has to be smaller than 0, in which case $\rho$ is positive.

In general, for $x \equiv 0 - \rho$ in $B$, indifference between $A$ and $B$ implies:

$$h(p_0)(w_0 - 1) + (1 - h(p_0))(w_0 + 1) = h(p_0 - \varepsilon_1)(w_0 - 1) + (h(p_0 + \varepsilon_1) - h(p_0 - \varepsilon_1))(w_0 - \rho) + (1 - h(p_0 + \varepsilon_1))(w_0 + 1),$$

where $w_0$ is the decision-maker’s initial wealth level. From (3.2) we obtain the explicit solution

$$\rho = \frac{h(p_0) - h(p_0 - \varepsilon_1) - (h(p_0 + \varepsilon_1) - h(p_0))}{h(p_0 + \varepsilon_1) - h(p_0 - \varepsilon_1)}.$$

(3.3)

By approximating $h(p_0 \pm \varepsilon_1)$ in (3.3) using second-order Taylor series expansions around $h(p_0)$, we obtain the following approximation for the DT risk premium:

$$\rho \approx \frac{\bar{m}_2}{2\Pr} \left( -\frac{h''(p_0)}{h'(p_0)} \right).$$

(3.4)

Here, $\bar{m}_2$ is the unconditional maxiance of the risk $\tilde{\varepsilon}_1$ that describes the mean-preserving spread from $B$ with $x \equiv 0$ to $A$. Unconditionally, $\tilde{\varepsilon}_1$ takes the values $\pm 1$ each with probability $\varepsilon_1$. Furthermore, $\Pr$ is the total unconditional probability mass associated to $\tilde{\varepsilon}_1$; see Figure 3.

Figure 3: Mean-Preserving Spread from $B$ with $x \equiv 0$ to $A$.

Observe that lottery $A$ is obtained from lottery $B$ (with $x \equiv 0$) by attaching the risk $\tilde{\varepsilon}_1$ to the intermediate branch of $B$. That is, the risk $\tilde{\varepsilon}_1$ is effective conditionally upon realization of the intermediate state of lottery $B$, which occurs with probability $2\varepsilon_1$. One readily verifies that, for this risk $\tilde{\varepsilon}_1$, we have that, unconditionally, $\bar{m}_2 = 2\varepsilon_1^2$ and $\Pr = 2\varepsilon_1$. We consider the unconditional maxiance of the zero-mean risk $\tilde{\varepsilon}_1$ to be “small” and compute the Taylor

---

17See the references in footnote 4 for global results on risk aversion under DT and RDU.
expansions up to the order $\varepsilon_1^2$. Henceforth, maxiances and variances are always understood to be unconditional.

It is important to compare the result in (3.4) to that obtained by PA presented in (1.1). In PA the local approximation of the risk premium is proportional to the variance, while under DT it is proportional to the maxiance.

We note that the local approximation of the risk premium in (3.4) remains valid in general, for non-binary zero-mean risks $\hat{\varepsilon}_1$ with small maxiance, just like, as is well-known, (1.1) is valid for non-binary zero-mean risks with small variance.\(^{18}\)

4 Local Risk Aversion under Rank-Dependent Utility

Under DT the local index arises from a risk change with small maxiance. To deal with the RDU model, which encompasses both EU and DT as special cases, we naturally have to consider changes in risk that are small in both variance and maxiance. To achieve this, we start from a lottery $C$ given by:\(^{19}\)

Figure 4: Lottery $C$

Similar to under DT, we transform lottery $C$ into a lottery $D$ by reducing the probabilities of both states in $C$ by a probability $\varepsilon_1$ and assigning the released probability $2\varepsilon_1$ to an intermediate state with outcome $y$, where $-\varepsilon_2 < y < \varepsilon_2$. This yields a lottery $D$ given by:

Figure 5: Lottery $D$

\(^{18}\)Detailed derivations are suppressed to save space. They are contained in an online appendix (available from the authors' webpages; see \url{http://www.rogerlaeven.com}).

\(^{19}\)We assume $\varepsilon_2 > 0$.  

7
Of course, when \( y \equiv 0 \), \( D \) is a mean-preserving contraction of \( C \). All RDU decision-makers that are averse to mean-preserving spreads therefore prefer \( D \) over \( C \) in that case. Formally, an RDU decision-maker evaluates a lottery \( A \) with cumulative distribution function \( F \) by computing
\[
\int U(x) \, dh(F(x)),
\]
and is averse to mean-preserving spreads if and only if \( U'' < 0 \) and \( h'' < 0 \).

In general, we can search for \( y \) such that indifference between \( C \) and \( D \) occurs. The discrepancy between the resulting \( y \) and 0 is the RDU risk premium associated to the risk change from \( C \) to \( D \) and its value, denoted by \( \lambda = 0 - y \), is the solution to
\[
h(p_0) U(w_0 - \varepsilon_2) + (1 - h(p_0)) U(w_0 + \varepsilon_2) = h(p_0 - \varepsilon_1) U(w_0 - \varepsilon_2) + (h(p_0 + \varepsilon_1) - h(p_0 - \varepsilon_1)) U(w_0 - \lambda) + (1 - h(p_0 + \varepsilon_1)) U(w_0 + \varepsilon_2) .
\]
It will be positive or negative depending on the shapes of \( U \) and \( h \).

Approximating the solution to (4.2) by Taylor series expansions, up to the first order in \( \lambda \) around \( U(w_0) \) and up to the second orders in \( \varepsilon_1 \) and \( \varepsilon_2 \) around \( U(w_0) \) and \( h(p_0) \), we obtain the following approximation for the RDU risk premium:
\[
\lambda \simeq \frac{m_2}{2\Pr} \left( - \frac{U''(w_0)}{U'(w_0)} \right) + \frac{\bar{m}_2}{2\Pr} \left( - \frac{h''(p_0)}{h'(p_0)} \right).
\]
Here, \( m_2 \) and \( \bar{m}_2 \) are the unconditional variance and maxiance of the risk \( \tilde{\varepsilon}_{12} \) that dictates the mean-preserving spread from \( D \) with \( y \equiv 0 \) to \( C \). Unconditionally, \( \tilde{\varepsilon}_{12} \) takes the values \( \pm \varepsilon_2 \) each with probability \( \varepsilon_1 \). Furthermore, \( \Pr \) is the total unconditional probability mass associated to \( \tilde{\varepsilon}_{12} \).

Comparing (4.3) to (1.1) and (3.4) reveals that the local approximation of the RDU risk premium aggregates the (suitably scaled) EU and DT counterparts, with the variance and maxiance featuring equally prominently.

As shown in online supplementary material, the local approximation of the RDU risk premium in (4.3) also generalizes naturally to non-binary risks \( \tilde{\varepsilon}_{12} \).

5 Related Literature

5.1 Global Results: Comparative Risk Aversion under RDU

Not only the local properties of the previous sections are valid under DT and RDU but also the corresponding global properties, just like in the PA-approach under the EU model (see, in

\footnote{See the references in footnote 4.} \footnote{It is straightforward to verify that for \( \tilde{\varepsilon}_{12} \) we have that, unconditionally, \( m_2 = 2\varepsilon_1\varepsilon_2^2, \bar{m}_2 = 2\varepsilon_1^2\varepsilon_2 \), and \( \Pr = 2\varepsilon_1 \).}
particular, Theorem 1 in Pratt \[22\]). In this section, we restrict attention to the RDU model. (The DT model occurs as a special case.) We first note that the definition of the RDU risk premium in (4.2) applies also when $\varepsilon_1$ and $\varepsilon_2$ are “large”, as long as $0 < \varepsilon_1 \leq \{p_0, 1 - p_0\} < 1$ and $\varepsilon_2 > 0$ are satisfied. We then state the following result:

**Proposition 5.1** Let $U_i, h_i, \lambda_i(p_0, w_0, \varepsilon_1, \varepsilon_2)$ be the utility function, the probability weighting function, and the risk premium solving (4.2) for RDU decision-maker $i = 1, 2$. Then the following conditions are equivalent:

(i) $-\frac{U''_2(w)}{U'_2(w)} \geq -\frac{U''_1(w)}{U'_1(w)}$ and $-\frac{h''_2(p)}{h'_2(p)} \geq -\frac{h''_1(p)}{h'_1(p)}$ for all $w$ and all $p \in (0, 1)$.

(ii) $\lambda_2(p_0, w_0, \varepsilon_1, \varepsilon_2) \geq \lambda_1(p_0, w_0, \varepsilon_1, \varepsilon_2)$ for all $0 < \varepsilon_1 \leq \{p_0, 1 - p_0\} < 1$, all $w_0$, and all $\varepsilon_2 > 0$.

Because the binary symmetric zero-mean risk $\tilde{\varepsilon}_{12}$ in Section 3 induces a risk change that is a special case of a mean-preserving spread, the implication $(i) \Rightarrow (ii)$ in Proposition 5.1 in principle follows from the classical results on comparative risk aversion under RDU (Yaari \[35\], Chew, Kani and Safra \[7\], and Roell \[26\]). The reverse implication $(ii) \Rightarrow (i)$ formalizes the connection between our local risk aversion approach and global risk aversion under RDU.

Due to the simultaneous involvement of both the utility function and the probability weighting function, the proof of the equivalences between (i) and (ii) under RDU is more complicated than that of the analogous properties under EU (and DT). Our proof of Proposition 5.1 (which is contained in online supplementary material available from the authors’ webpages) is based on the total differential of the RDU evaluation, and the sensitivity of the risk premium with respect to changes in payoffs.

### 5.2 Relation to the Pratt-Arrow Definition of the Risk Premium

Our definition of the risk premium under RDU in (4.2) can be viewed as a natural generalization of the risk premium of Pratt \[22\] and Arrow \[2, 3\]. To see this, first note that the PA-definition, under which a risk is compared to a sure loss equal to the risk premium, occurs when $p_0 = \varepsilon_1 = \frac{1}{2}$. Then, lottery $D$ becomes a sure loss of $\lambda$ the value of which is such that the decision-maker is indifferent to the risk of lottery $C$.

When $\varepsilon_1 < \frac{1}{2}$, our definition of the RDU risk premium provides a natural generalization of the PA-definition. This becomes readily apparent if we omit the common components of lotteries $D$ and $C$ with the same incremental RDU evaluation and isolate the risk change, which yields

\[22\] Recall that the probability $\varepsilon_1$ and payoff $\pm \varepsilon_2$ in (4.2) can be “large” as long as $0 < \varepsilon_1 \leq \{p_0, 1 - p_0\} < 1$ and $\varepsilon_2 > 0$. 

9
Figure 6: Lottery $D$ after Omitting the Components in Common with Lottery $C$.

$$D \setminus C \cap D \xrightarrow{2\varepsilon_1} -\lambda$$

and

Figure 7: Lottery $C$ after Omitting the Components in Common with Lottery $D$.

The value of $\lambda$ thus represents the risk premium for the risk change induced by a risk that, unconditionally, takes the values $\pm \varepsilon_2$ each with probability $\varepsilon_1$.

When $\varepsilon_1 < \frac{1}{2}$, the original comparison between $C$ and $D$ is a comparison between two risky situations as in Ross [27], Machina and Neilson [17], and Pratt [23]. The removal of common components, however, reveals that we essentially face a PA-comparison between a single loss and a risk with the same total probability mass, which is now allowed to be smaller than unity.

5.3 Related Measures of Risk

Dual moments can be related to the Gini coefficient named after statistician Corrado Gini and used by economists to measure the dispersion of the income distribution of a population, summarizing its income inequality. In risk theory, the Gini coefficient $G$ of a risk $A$ is usually defined by

$$G = \frac{\mathbb{E} \left[ |A^{(1)} - A^{(2)}| \right]}{2\mathbb{E} |A|},$$

(5.4)

which represents half the relative (i.e., normalized) expected absolute difference between two independent draws of the risk $A$. One can verify that, equivalently but less well-known,

$$G = \frac{\overline{m}_2}{m}.$$

(5.5)

This alternative expression features the ratio of the maxiance and the first moment.

Furthermore, $n$-th degree expectations of first order statistics also appear in Cherny and Madan [6] as performance measures in the context of portfolio evaluation. In this setting, the expected maximal financial loss occurring in $n$ independent draws of a risk is used as a measure to define an acceptability index linked to a tolerance level of stress.
6 Examples

Owing to its local nature, our approximation is valid and can insightfully be applied when the probability weighting function is not globally concave, as is suggested by ample experimental evidence. Consider the canonical probability weighting function of Prelec \[24\] given by\[23\]

\[ h(p) = 1 - \exp\{-(-\log(1-p))^\alpha\}, \quad 0 < \alpha < 1. \quad (6.1) \]

It captures the following properties which are observed empirically: it is regressive (first, \( h(p) > p \), next \( h(p) < p \)), is inverse s-shaped (first concave, next convex), and is asymmetric (intersecting the identity probability weighting function \( h(p) = p \) at \( p^* = 1 - 1/\exp(1) \), the inflection point).\[24\] The upper panel of Figure 8 plots this probability weighting function for \( \alpha \in \{0.1, 0.3, \ldots, 0.9\} \). (Wu and Gonzalez \[33\] report estimated values of \( \alpha \) between 0.03 and 0.95.)

Its local index \(-\frac{h''(p)}{h'(p)}\) takes the form

\[ -\frac{h''(p)}{h'(p)} = -\frac{1 - \alpha (1 - (-\log(1-p))^\alpha) + \log(1-p)}{(1-p) \log(1-p)}. \quad (6.2) \]

Figure 8, lower panel, plots this local index for \( \alpha \in \{0.1, 0.3, \ldots, 0.9\} \).

\[23\] Recall our convention to distort cumulative probabilities rather than decumulative probabilities; see footnote 12. Prelec’s original function is given by \( w(p) = 1 - h(1 - p) \).

\[24\] Prelec’s function is characterized axiomatically as the probability weighting function of a sign- and rank-dependent preference representation that exhibits subproportionality, diagonal concavity, and so-called compound invariance.
Figure 8: Prelec’s Probability Weighting Function (upper panel) and its Local Index (lower panel). We consider $\alpha \in \{0.1, 0.3, \ldots, 0.9\}$.

The inverse $s$-shape of the probability weighting function (first concave, next convex) implies that its local index changes sign at the inflection point. More specifically, the local index associated with Prelec’s probability weighting function is decreasing (first positive, next negative) in $p$ for any $0 < \alpha < 1$. This property is naturally consistent with the inverse $s$-shape property of the probability weighting function: the inverse $s$-shape property is meant to represent a psychological phenomenon known as *diminishing sensitivity* in the probability domain (rather than the payoff domain), under which the decision-maker is less sensitive to changes in the objective probabilities when they move away from the reference points 0 and 1, and becomes more sensitive when the objective probabilities move towards these reference points.

A decreasing local index implies in particular that $h''' > 0$. (By Pratt [22], the sign of the derivative of the local index is the same as the sign of $(h''(p))^2 - h'(p)h'''(p)$.) Inverse $s$-shaped probability weighting functions, including Prelec’s canonical example, usually exhibit positive signs for the odd derivatives and alternating signs (first negative, then positive) for the even derivatives. For a probability weighting function that is inverse $s$-shaped (first concave, then convex) and has second derivative equal to zero at the inflection point, a positive sign of the third derivative means that the function becomes increasingly concave when we move to the
left of the inflection point and becomes increasingly convex when we move to the right of the inflection point.

In Figure 12 in the online appendix we also plot the local index $-\frac{h''(p)}{h'(p)}$ of the probability weighting function proposed by Tversky and Kahneman [32] (see also Wu and Gonzalez [33]) given by

$$h(p) = 1 - \frac{(1 - p)^{\beta}}{(1 - p)^{\beta} + p^{\beta}}^{1/\beta}, \quad 0 < \beta < 1,$$

(6.3)

for values of the parameter $\beta \in \{0.55, 0.65, \ldots, 0.95\}$ as found in experiments (Wu and Gonzalez [33] report estimated values of $\beta$ between 0.57 and 0.94). Observe the similarity between the shapes in Figure 8 and Figure 12.

The analysis in this paper reveals that for a small risk the sign and size of the maxiance’s contribution to the RDU risk premium, given by the second term on the right-hand side of (4.3), i.e.,

$$\frac{\bar{m}_2}{2\Pr} \left( -\frac{h''(p_0)}{h'(p_0)} \right),$$

varies with the probability level $p_0$, from strongly positive to strongly negative, in tandem with the local index $-\frac{h''(p)}{h'(p)}$ to which it is proportional.

We finally plot in Figure 9 our approximation to the RDU risk premium (4.3) of a risk with small variance and maxiance normalized to satisfy $\bar{m}_2 = \bar{x}_{\Pr} = 1$, as a function of both the initial wealth level $w_0$ and the probability level $p_0$. We suppose the utility function is given by the conventional power utility (note that we consider a pure rank-dependent model)

$$U(x) = x^\gamma,$$

with $\gamma = 0.5$ (consistent with the gain domain in Tversky and Kahneman [32] and with experimental evidence in Wu and Gonzalez [33]), and the probability weighting function is that of Prelec with parameter $\alpha = 0.65$. 

13
Figure 9: Surface of the RDU Risk Premium Approximation. We consider a risk with small variance and maxiance normalized to satisfy $\frac{\bar{m}}{2\sigma} = \frac{\bar{\beta}}{2\rho} = 1$ under power utility (with $\gamma = 0.5$) and Prelec’s probability weighting function (with $\alpha = 0.65$).

Figure 9 illustrates the interplay between the variance’s and the maxiance’s contributions to the RDU risk premium (4.3), depending on the local indices \(-U''(w)\) and \(-h''(p)\) evaluated in the wealth and probability levels $w_0$ and $p_0$, respectively. The orange surface represents our approximation (4.3) to the RDU risk premium $\lambda$, while the blue surface is the $\lambda = 0$-plane. To illustrate the effect of a change in variance or maxiance, we also plot in Figure 13 in the online appendix the surface of the RDU risk premium approximation (4.3) for a small risk with ratio between the variance and maxiance equal to 3 (upper panel) and $1/3$ (lower) panel, instead of a ratio of 1 as in Figure 9.

7 A Portfolio Application

In order to illustrate how the concepts we have developed can be used we consider a simple portfolio problem with a safe asset, the return of which is zero, and a binary risky asset with returns expressed by the following representation:\footnote{We assume $0 < R_0 < R_1$.}
Taking \( \frac{R_1}{R_0 + R_1} > p_0 \) makes the expected return strictly positive.

If an RDU investor has initial wealth \( w_0 \) his portfolio optimization problem is given by

\[
\arg \max_\alpha \left\{ h(p_0) U(w_0 - \alpha R_0) + (1 - h(p_0)) U(w_0 + \alpha R_1) \right\},
\]

with first-order condition (FOC) given by

\[
-R_0 h(p_0) U'(w_0 - \alpha R_0) + R_1 (1 - h(p_0)) U'(w_0 + \alpha R_1) \equiv 0.
\]

It is straightforward to show that the second-order condition for a maximum is satisfied provided \( U'' < 0 \).

Let us now pay attention to the RDU investor for whom it is optimal to choose not to invest in the risky asset, i.e., to select \( \alpha \equiv 0 \). Plugging \( \alpha \equiv 0 \) into the FOC we obtain the condition

\[
h(p_0) \equiv \frac{R_1}{R_0 + R_1}.
\]

Without surprise, \( h(p_0) > p_0 \). This value of \( h(p_0) \) expresses the intensity of risk aversion that induces the choice of \( \alpha \equiv 0 \).

Now consider a mean-preserving contraction of the return of the risky asset given by:

\[
\begin{align*}
&\begin{array}{c}
\text{Figure 11: Mean-Preserving Contraction of the Risky Asset}
\end{array}
\end{align*}
\]

\[
\begin{align*}
&\begin{array}{c}
\text{Figure 11: Mean-Preserving Contraction of the Risky Asset}
\end{array}
\end{align*}
\]

One may verify that such a mean-preserving contraction for a decision-maker who had decided not to participate in the risky asset may induce him to select a strictly positive \( \alpha \).
Hence, we raise the following question: By how much should we reduce the intermediate return \( \frac{R_1 - R_0}{2} \) to induce the decision-maker to stick to the optimal \( \alpha \) equal to zero? The answer to this question is denoted by \( \varsigma \).

Because we are concentrating on the situation where \( \alpha \equiv 0 \) is optimal, the analysis is related only to the shape of the probability weighting function. Indeed, the shape of \( U \) that appears in the FOC through different values of \( U' \) becomes irrelevant at \( \alpha \equiv 0 \). The reason to concentrate on \( \alpha \equiv 0 \) where only the probability weighting function matters under RDU pertains to the well-known fact that under EU a mean-preserving contraction of the risky return has an ambiguous effect on the optimal \( \alpha \) (Gollier [15]).

It turns out that, upon invoking Taylor series expansions and after several basic manipulations, the reduction \( \varsigma \) that answers our question raised above is given by

\[
\varsigma \approx \frac{\bar{m}_2}{2\text{Pr}} \left( -\frac{h''(p_0)}{h'(p_0)} \right),
\]  

(7.3)

where \( \bar{m}_2 \) is the maxiance of the risk that, unconditionally, takes the values \( \pm \frac{R_0 + R_1}{2} \) each with probability \( \varepsilon_1 \), and where \( \text{Pr} \) is the total probability mass of this risk. Again the second dual moment (instead of the primal one) appears, jointly with the intensity of risk aversion induced by the probability weighting function. In particular, the mean-preserving contraction is an improvement and has made the risky asset attractive if and only of \( \varsigma \) is positive.

8 Conclusion

Under EU, the risk premium is approximated by an expression that multiplies half the variance of the risk (i.e., its primal second central moment) by the local index of absolute risk aversion. This expression dissects the complex interplay between the risk’s probability distribution, the decision-maker’s preferences and his initial wealth that the risk premium in general depends on. Surprisingly, a similar expression almost never appears in non-EU models.

In this paper, we have shown that when one refers to the second dual moment—instead of, or on par with, the primal one—one obtains quite naturally an approximation of the risk premium in canonical non-EU models that mimics the one obtained within the EU model.

The PA-approximation of the risk premium under EU has induced thousands of applications and results in many fields such as operations research, insurance, finance, and environmental economics. So far, comparable developments have been witnessed to a much lesser extent outside the EU model. Hopefully, the new and simple expression of the approximated risk premium may contribute to a widespread analysis and use of risk premia for non-EU.

References

[1] Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting
functions. *Management Science* 46, 1497-1512.

[2] Arrow, K.J. (1965). *Aspects of the Theory of Risk-Bearing*. Yrjö Jahnsson Foundation, Helsinki.

[3] Arrow, K.J. (1971). *Essays in the Theory of Risk-Bearing*. North-Holland, Amsterdam.

[4] Chateauneuf, A., M. Cohen and I. Meilijson (2004). Four notions of mean preserving increase in risk, risk attitudes and applications to the Rank-Dependent Expected Utility model. *Journal of Mathematical Economics* 40, 547-571.

[5] Chateauneuf, A., M. Cohen and I. Meilijson (2005). More pessimism than greediness: A characterization of monotone risk aversion in the rank-dependent expected utility model. *Economic Theory* 25, 649-667.

[6] Cherny, A. and D. Madan (2009). New measures for portfolio evaluation. *Review of Financial Studies* 22, 2571-2606.

[7] Chew, S.H., E. Karni and Z. Safra (1987). Risk aversion in the theory of expected utility with rank dependent probabilities. *Journal of Economic Theory* 42, 370-381.

[8] David, H.A. (1981). *Order Statistics*. 2nd Ed., Wiley, New York.

[9] Eeckhoudt, L.R. and H. Schlesinger (2006). Putting risk in its proper place. *American Economic Review* 96, 280-289.

[10] Eeckhoudt L.R., H. Schlesinger and I. Tsetlin (2009). Apportioning of risks via stochastic dominance. *Journal of Economic Theory* 144, 994-1003.

[11] Eeckhoudt, L.R., R.J.A. Laeven and H. Schlesinger (2018). Risk apportionment: The dual story. Mimeo, IESEG, University of Amsterdam and University of Alabama.

[12] Eichner, T. and A. Wagener (2009). Multiple risks and mean-variance preferences. *Operations Research* 57, 1142-1154.

[13] Ekern, S. (1980). Increasing nth degree risk. *Economics Letters* 6, 329-333.

[14] Fishburn, P.C. (1980). Stochastic dominance and moments of distributions. *Mathematics of Operations Research* 5, 94-100.

[15] Gollier, C. (1995). The comparative statics of changes in risk revisited. *Journal of Economic Theory* 66, 522-536.

[16] Harrison, G.W. and J.T. Swarthout (2016). Cumulative prospect theory in the laboratory: A reconsideration, Mimeo, CEAR.

[17] Machina, M.J. and W.S. Neilson (1987). The Ross characterization of risk aversion: Strengthening and extension. *Econometrica* 55, 1139-1149.

[18] Machina, M.J. and J.W. Pratt (1997). Increasing risk: Some direct constructions. *Journal of Risk and Uncertainty* 14, 103-127.

[19] Meyer, J. (1987). Two-moment decision models and expected utility maximization. *American Economic Review* 77, 421-430.

[20] Muliere, P. and M. Scarsini (1989). A note on stochastic dominance and inequality measures. *Journal of Economic Theory* 49, 314-323.

[21] Nau, R.F. (2003). A generalization of Pratt-Arrow measure to nonexpected-utility preferences and inseparable probability and utility. *Management Science* 49, 1089-1104.
[22] Pratt, J.W. (1964). Risk aversion in the small and in the large. *Econometrica* 32, 122-136.

[23] Pratt, J.W. (1990). The logic of partial-risk aversion: Paradox lost. *Journal of Risk and Uncertainty* 3, 105-113.

[24] Prelec, D. (1998). The probability weighting function. *Econometrica* 66, 497-527.

[25] Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization* 3, 323-343.

[26] Roëll, A. (1987). Risk aversion in Quiggin and Yaari’s rank-order model of choice under uncertainty. *The Economic Journal* 97, 143-159.

[27] Ross, S.A. (1981). Some stronger measures of risk aversion in the small and the large with applications. *Econometrica* 49, 621-638.

[28] Rothschild, M. and J.E. Stiglitz (1970). Increasing risk: I. A definition. *Journal of Economic Theory* 2, 225-243.

[29] Ryan, M.J. (2006). Risk aversion in RDEU. *Journal of Mathematical Economics* 42, 675-697.

[30] Schmidt, U. and H. Zank (2008). Risk aversion in cumulative prospect theory. *Management Science* 54, 208-216.

[31] Smith, J.E. and R.L. Winkler (2006). The optimizer’s curse: Skepticism and postdecision surprise in decision analysis. *Management Science* 52, 311-322.

[32] Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297-323.

[33] Wu, G. and R. Gonzalez (1996). Curvature of the probability weighting function. *Management Science* 42, 1676-1690.

[34] Yaari, M.E. (1969). Some remarks on measures of risk aversion and on their uses. *Journal of Economic Theory* 1, 315-329.

[35] Yaari, M.E. (1986). Univariate and multivariate comparisons of risk aversion: a new approach. In: Heller, W.P., R.M. Starr and D.A. Starrett (Eds.). *Uncertainty, Information, and Communication*. Essays in honor of Kenneth J. Arrow, Volume III, pp. 173-188, 1st Ed., Cambridge University Press, Cambridge.

[36] Yaari, M.E. (1987). The dual theory of choice under risk. *Econometrica* 55, 95-115.
SUPPLEMENTARY MATERIAL
(FOR ONLINE PUBLICATION)
A Generalization to Non-Binary Risks

In this supplementary material, we first show that the local approximation for the DT risk premium in (3.4) remains valid for non-binary risks with small maxiance. Next, we prove that the RDU risk premium approximation in (4.3) also remains valid for non-binary risks with both small variance and small maxiance. Throughout this supplement, we consider \( n \)-state risks with probabilities \( p_i \) associated to outcomes \( x_i, i = 1, \ldots, n \), with \( n \in \mathbb{N}_{>0} \). We order states from the lowest outcome state (designated by state number 1) to the highest outcome state (designated by state number \( n \)), which means that \( x_1 \leq \cdots \leq x_n \).

We analyze the DT risk premium for a risk with \( n \geq 2 \) effective states that have equal unconditional probability of occurrence given by \( \frac{2 \varepsilon_1}{n}, 0 < \varepsilon_1 \leq \frac{1}{2} \). The outcomes are, however, allowed to be the same among adjacent states; this would correspond to a risk with non-equal state probabilities. Note the generality provided by this construction. We suppose that the risk has mean equal to zero, so \( \sum_{i=1}^{n} x_i = 0 \). One may verify that the unconditional maxiance of this \( n \)-state risk is given by

\[
\bar{m}_2 = \frac{4 \varepsilon_1^2}{n^2} \sum_{i=1}^{n} (2i - 1) x_i, \quad (A.1)
\]

and that the total probability mass \( \mathbb{P} x = 2 \varepsilon_1 \). Observe that the maxiance is of the order \( \varepsilon_1^2 \), i.e., \( \bar{m}_2 = O (\varepsilon_1^2) \).

Similar to the main text, this zero-mean risk is attached to the intermediate branch of lottery \( B \) (with \( x \equiv 0 \)) to induce a mean-preserving spread. (We normalize the outcomes of the zero-mean risk by restricting them to the interval \([-1, 1]\). This ensures that the initial ordering of outcomes in lottery \( B \) is not affected and can easily be generalized.) The DT risk premium \( \rho \) then occurs as the solution to

\[
(h (p_0 + \varepsilon_1) - h (p_0 - \varepsilon_1)) (w_0 - \rho) = \sum_{i=1}^{n} \left( h \left( p_0 - \varepsilon_1 + \frac{i \varepsilon_1}{n} \right) - h \left( p_0 - \varepsilon_1 + (i - 1) \frac{2 \varepsilon_1}{n} \right) \right) (w_0 + x_i). \quad (A.2)
\]

From (A.2) we obtain the explicit solution

\[
\rho = -\sum_{i=1}^{n} \frac{h \left( p_0 - \varepsilon_1 + i \frac{2 \varepsilon_1}{n} \right) - h \left( p_0 - \varepsilon_1 + (i - 1) \frac{2 \varepsilon_1}{n} \right)}{h (p_0 + \varepsilon_1) - h (p_0 - \varepsilon_1)} x_i. \quad (A.3)
\]

By invoking Taylor series expansions around \( h (p_0) \) up to the second order in \( \varepsilon_1 \) we obtain
from (A.3) the following approximation for the DT risk premium:

\[
\rho \simeq - \sum_{i=1}^{n} \frac{1}{2} \left(2i - 1\right) \frac{\epsilon_1^2}{2} \frac{h''(p_0)}{h'(p_0)} x_i
\]

\[
= \frac{\bar{m}_2}{2 \Pr} \left( - \frac{h''(p_0)}{h'(p_0)} \right),
\]

where the last equality follows directly from (A.1).

Finally, turning to the risk premium under RDU, we consider, as under DT, an \( n \)-state zero-mean risk with unconditional state probabilities \( \frac{2 \epsilon_1}{n} \), so \( \sum_{i=1}^{n} x_i = 0 \) and \( \Pr = 2 \epsilon_1 \), now assumed to satisfy additionally that \( \bar{m}_2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 = O \left( \epsilon_2^2 \right) \) for some \( \epsilon_2 > 0 \). Upon attaching this zero-mean risk to the intermediate branch of lottery \( D \) (with \( y \equiv 0 \) and assuming without losing generality that \( |x_i| < \epsilon_2 \)), the RDU risk premium \( \lambda \) occurs as the solution to

\[
(h(p_0 + \epsilon_1) - h(p_0 - \epsilon_1)) U(w_0 - \lambda)
= \sum_{i=1}^{n} \left( h\left(p_0 - \epsilon_1 + i \frac{2 \epsilon_1}{n}\right) - h\left(p_0 - \epsilon_1 + (i - 1) \frac{2 \epsilon_1}{n}\right) \right) U(w_0 + x_i).
\]

(A.4)

Invoking Taylor series expansions up to the first order in \( \lambda \) around \( U(w_0) \) and up to the second order in \( x_i \) and \( \epsilon_1 \) around \( U(w_0) \) and \( h(p_0) \), we obtain from (A.4), at the leading orders, the desired approximation for the RDU risk premium:

\[
\lambda \simeq - \sum_{i=1}^{n} \frac{1}{2} \frac{2 \epsilon_1 U''(w_0)}{2 \epsilon_1 U'(w_0)} x_i^2 - \sum_{i=1}^{n} \frac{1}{2} \left(2i - 1\right) \frac{4 \epsilon_1^2}{n^2} \frac{h''(p_0)}{h'(p_0)} x_i
\]

\[
= \frac{\bar{m}_2}{2 \Pr} \left( - \frac{U''(w_0)}{U'(w_0)} \right) + \frac{\bar{m}_2}{2 \Pr} \left( - \frac{h''(p_0)}{h'(p_0)} \right).
\]
B Proof of Proposition 5.1

First, note that (i) is equivalent to

(i) \( U_2(U_1^{-1}(t)) \) and \( h_2(h_1^{-1}(u)) \) are concave functions of \( t \) and \( u \) for all \( t \) and all \( u \in (0, 1) \).

(v) \( \frac{U_2(y) - U_2(x)}{U_2(w) - U_2(v)} \leq \frac{U_1(y) - U_1(x)}{U_1(w) - U_1(v)} \) and \( \frac{h_2(s) - h_2(r)}{h_2(q) - h_2(p)} \leq \frac{h_1(s) - h_1(r)}{h_1(q) - h_1(p)} \) for all \( v < w \leq x < y \) and all \( 0 < p < q \leq r < s < 1 \).

The equivalence of (i), (iv) and (v) follows trivially from the equivalence of (a), (d) and (e) in Theorem 1 of Pratt [22] and the corresponding DT counterpart equivalences.

Second, we will prove that (the equivalent) (i), (iv) and (v) imply (ii). Reconsider (4.2). Fix (a feasible) \( \epsilon_1 > 0 \) (satisfying \( 0 < \epsilon_1 \leq \{p_0, 1 - p_0\} < 1 \)). Note that if we let \( \epsilon_2 \to 0 \) in (4.2), then \( \lambda_i \to 0 \). Define

\[
V_i(\lambda_i, \epsilon_2) = (h_i(p_0 + \epsilon_1) - h_i(p_0 - \epsilon_1)) U_i(w_0 - \lambda_i) - ((h_i(p_0) - h_i(p_0 - \epsilon_1)) U_i(w_0 - \epsilon_2) + (h_i(p_0 + \epsilon_1) - h_i(p_0)) U_i(w_0 + \epsilon_2)).
\]

We compute the total differential \( dV_i = \frac{\partial V_i}{\partial \lambda_i} \, d\lambda_i + \frac{\partial V_i}{\partial \epsilon_2} \, d\epsilon_2 \). It is given by

\[
- (h_i(p_0 + \epsilon_1) - h_i(p_0 - \epsilon_1)) U'_i(w_0 - \lambda_i) \, d\lambda_i + ((h_i(p_0) - h_i(p_0 - \epsilon_1)) U'_i(w_0 - \epsilon_2) - (h_i(p_0 + \epsilon_1) - h_i(p_0)) U'_i(w_0 + \epsilon_2)) \, d\epsilon_2.
\]

Equating the total differential to zero yields

\[
\frac{d\lambda_i}{d\epsilon_2} = \frac{h_i(p_0) - h_i(p_0 - \epsilon_1)}{h_i(p_0 + \epsilon_1) - h_i(p_0 - \epsilon_1)} \frac{U'_i(w_0 - \epsilon_2)}{U'_i(w_0 - \lambda_i)} - \frac{h_i(p_0 + \epsilon_1) - h_i(p_0)}{h_i(p_0 + \epsilon_1) - h_i(p_0 - \epsilon_1)} \frac{U'_i(w_0 + \epsilon_2)}{U'_i(w_0 - \lambda_i)}. \tag{B.5}
\]

From (i), as in Pratt [22], Eqn. (20),

\[
\frac{U'_2(x)}{U'_2(w)} \leq \frac{U'_1(x)}{U'_1(w)}, \quad \text{for } w < x, \quad \text{and} \quad \frac{U'_2(x)}{U'_2(y)} \geq \frac{U'_1(x)}{U'_1(y)}, \quad \text{for } x < y.
\]

Furthermore, from (v),

\[
\frac{U_2(y) - U_2(x)}{U_2(w) - U_2(v)} + \frac{U_2(w) - U_2(v)}{U_2(w) - U_2(v)} \leq \frac{U_1(y) - U_1(x)}{U_1(w) - U_1(v)} + \frac{U_1(w) - U_1(v)}{U_1(w) - U_1(v)}, \quad \text{for } v < w \leq x < y.
\]

Taking \( w = x \) yields

\[
\frac{U_2(y) - U_2(v)}{U_2(w) - U_2(v)} \leq \frac{U_1(y) - U_1(v)}{U_1(w) - U_1(v)}, \quad \text{for } v < w < y.
\]
hence
\[
\frac{U_2(w) - U_2(v)}{U_2(y) - U_2(v)} \geq \frac{U_1(w) - U_1(v)}{U_1(y) - U_1(v)},
\]
and also
\[
\frac{U_2(y) - U_2(w)}{U_2(y) - U_2(v)} \leq \frac{U_1(y) - U_1(w)}{U_1(y) - U_1(v)},
\]
for \( v < w < y \). In all inequalities in this paragraph, \( U_i \) may be replaced by \( h_i \), with \( v, w, x \) and \( y \) restricted to \((0, 1)\).

Thus, from (B.5) and the inequalities above,
\[
\frac{d\lambda_2}{d\varepsilon_2} \geq \frac{d\lambda_1}{d\varepsilon_2},
\]
hence (ii).

We have now proved that (ii) is implied by (the equivalent) (i), (iv) and (v). We finally show that (ii) implies (i), or rather that not (i) implies not (ii). This goes by realizing that, by the arbitrariness of \( w_0, p_0, \varepsilon_1 \) with \( 0 < \varepsilon_1 \leq \{ p_0, 1 - p_0 \} < 1 \) and \( \varepsilon_2 > 0 \), if (i) does not hold on some interval (of \( w \) or \( p \)), one can always find feasible \( w_0, p_0, \varepsilon_1 \) and \( \varepsilon_2 \), such that (B.6), hence (ii), hold on some interval but with the inequality signs strict and flipped. \( \square \)
C Figures

Figure 12: Tversky-Kahneman Probability Weighting Function (upper panel) and its Local Index (lower panel). We consider $\beta \in \{0.55, 0.65, \ldots, 0.95\}$. 
Figure 13: Surface of the RDU Risk Premium Approximation. We consider a risk with small variance and maxiance, normalized to satisfy $\frac{\bar{m}^2}{\bar{m}^2_2} = 1$ with $\frac{m_2}{\bar{m}_2} = 3$ (upper panel) and $\frac{m_2}{\bar{m}_2} = 1$ with $\frac{m_2}{\bar{m}_2} = 1/3$ (lower panel), under power utility (with $\gamma = 0.5$) and Prelec’s probability weighting function (with $\alpha = 0.65$).