Chapter 11 — Probability Weighting and Rank Dependent Utility

Cumulative Prospect Theory (CPT) is one very popular alternative to Expected Utility Theory. It was developed, primarily by Kahneman and Tversky to explain certain observed anomalies in behavior. The first difference is that the utility function, or value function as it is more commonly called in CPT, is not concave but S-shaped. The second is that the expectation is not computed with probabilities, but with decision weights. Using non-concave utility functions is not inconsistent with EUT, though, of course, it is inconsistent with risk aversion. Probability weighting is not generally consistent with EUT.

Neither of these ideas are original to CPT. Non-concave utility functions were proposed very soon after Von Neumann and Morgenstern’s original proof of EUT in 1947. Probably the first was the Friedman-Savage (1948) utility function which has a convex portion in the middle. Several different types of weighted utility were discussed briefly at the end of Chapter 1. One that is almost identical to that used in CPT is the weighting in Rank Dependent Utility (RDU). It is also called Rank Dependent Expected Utility and when first developed by Quiggin (1982) was called Anticipated Utility.¹²

Either CPT or RDU can be described as follows. Let \( x \) and \( \pi \) denote vectors of payoffs and the probabilities of those payoffs. Choices are made by maximizing

\[
E_\omega[u(\tilde{x})] = \sum \omega_n(\pi, x)u(x_n).
\]

(1)

This is essentially an expected utility of the payoff; however, the expectation is computed with decision weights, \( \omega_n \), rather than the actual probabilities. The decision weights are not subjective probabilities although they are referred to as such at times. They are used even when the gamble has clear objective probabilities. Kahneman and Tversky (1979, p. 280) describe decision weights as “measure[ing] the impact of events on the desirability of prospects not merely the perceived likelihood of these events.” What distinguishes both RDU and CPT from other weighted utilities is that each decision weight, \( \omega_n \), depends on the set of probabilities and payoffs and not just a particular probability or payoff, \( \pi_n \) and \( x_n \).

Probability weighting and S-shaped utility are not intrinsically linked and will be presented separately. This chapter looks at cumulative probability weighting which is the basis for RDU and is one feature in CPT. The next chapter examines CPT.

Probability Weighting and Dominance

There are many different utility theories that use probability weighting. In Weighted Expected Utility, Implicit Weighted Utility, Implicit Expected Utility, Disappointment Aversion, and the original version of Prospect Theory (among others), the weighting depends on the outcomes \( x \) or on the overall utility. In RDU and CPT, the weighting depends on the cumulative probabilities.

The idea of using distorted probabilities in computing “expected” utility is quite old. Edwards (1954, p 395) claimed that “subjects, when they bet, prefer some probabilities to others.” The first proposed probability weighting schemes evaluated gambles using

¹ A special case of RDU is Yaari’s (1987) Dual Theory. It is dual to EUT in requiring a linear utility function but uses the same weighting functions as RDU. As such, Yaari’s theory and EUT are dual special cases within RDU.
² The choice alternatives in RDU are probability distributions of outcomes. Choquet expected utility is essentially the same theory applied to the choice among acts with consequences. This is the same distinction between von Neumann Morgenstern expected utility and Savage subjective expected utility. The name comes from its use of a Choquet integral.
In particular low probabilities were typically over-weighted and high probabilities under weighted; that is, \( \omega \) is a decreasing function. This type of weighting is also part of Prospect Theory, the predecessor of CPT.

Such weighting methods can explain the Allais Paradox and the Common Ratio Effect. However, as shown by Machina (1983), they also necessarily make some selections that are first-degree stochastically dominated. To see this, compare the lottery that pays \( x \) for sure to another with \( N \) outcomes that pays \( x + \varepsilon_n \) each with equal probability \( \frac{N-1}{N} \) for increasing, bounded values of \( \varepsilon, 0 < \varepsilon_1 \leq \varepsilon_2 \leq \ldots \leq \varepsilon_N \). The safe lottery is evaluated at \( u(x) \). For a continuous value function and \( \varepsilon_N \) sufficiently small, the risky lottery is evaluated at

\[
\sum_{n=1}^{N} \omega(N^{-1})u(x + \varepsilon_n) = \omega(N^{-1})\sum_{n=1}^{N} u(x + \varepsilon_n) \sim \omega(N^{-1})Nu(x)
\]  

If there is some \( N \) for which \( \omega(N^{-1}) < N^{-1} \), the safe lottery is preferred even though it is first order stochastically dominated because all payments in the risky lottery are larger. If there is no \( N \) for which \( \omega(N^{-1}) < N^{-1} \) and \( \omega \) is not the identity function, then there must be some \( N \) for which \( \omega(N^{-1}) > N^{-1} \). In that case a lottery that pays \( x - \varepsilon_n \) with equal probabilities of \( \frac{N-1}{N} \) is evaluated above the safe lottery even though the risky lottery is now dominated.

Kahneman and Tversky eliminated this dominance problem by adding a preliminary stage of editing to the choice process. In this stage, any gamble that is first-degree stochastically dominated by another is eliminated. However, this approach, too, has problems. It is possible to construct examples in which \( A \) is preferred to \( B \) which is preferred to \( C \) under the probability weighting, but \( C \) first-degree stochastically dominates \( A \) as in the example above. So if only \( A \) and \( B \) are available, \( A \) will be chosen. But if all three are available, \( C \) will eliminate \( A \) in the first stage, and then \( B \) will be chosen over \( C \). So \( C \) is never chosen, but its presence affects the choice made. More complicated examples in which the order of elimination of the dominated lotteries leads to different choices are also possible. For example, there can be a fourth choice \( D \) which is not preferred to any of the others under probability weighting, but stochastically dominates \( C \). In this case, the order of culling determines which prospect is chosen. If \( A \) is culled first by comparison to \( C \), then \( B \) is chosen. But if \( C \) is first culled due to \( D \), then \( A \) is chosen over \( B \) and \( D \).

More generally the dominance problem can be removed by weighting the cumulative distribution rather than the individual probabilities, and then determining the individual decision weights as differences. Under cumulative weighting, the outcomes are first ordered from lowest to highest, \( x_1 < x_2 < \ldots < x_N \).

Any two outcomes that are the same are merged into a single outcome with the sum of the probabilities so these inequalities are strict. The cumulative probabilities, \( \Pi_n = \sum_{i=1}^{n} \pi_m \), are determined. A weighting function is applied to the cumulative probabilities, \( \Omega(\Pi_n) \). The individual weights are determined by differencing

\[
\omega_n = \Omega(\Pi_n) - \Omega(\Pi_{n-1}) \quad \text{with} \quad \Pi_0 \equiv 0.
\]  

Obviously, the weighting function in (4) must be strictly increasing to ensure that all weights from every possible gamble are strictly positive. If this were not true, then the weights and the probabilities would disagree about arbitrage. Similarly, \( \Omega(0) = 0 \) so that impossible events are given a weight of zero. Finally \( \Omega(1) = 1 \) so that certain events are assessed as certain.

---

3 If you are familiar with CPT, one detail is being skipped here. In CPT two weighting functions are applied separately to the cumulative probability distribution for losses and to the complementary cumulative probabilities for gains. This difference is discussed in the next chapter.
This also assures that the weights sum (or integrate) to 1. Any function satisfying these conditions is a valid cumulative probability weighting function. Usually, the weighting function is assumed continuous so that events with very small probabilities can never be assigned a large decision weight.

Substituting (4) into (1), any prospect is evaluated as

$$\mathbb{E}_\Omega[u(\bar{x})] = \sum_{n=1}^{N} \omega_n u(x_n) = \sum_{n=1}^{N} [\Omega(\Pi_n) - \Omega(\Pi_{n-1})] u(x_n)$$

(5)

again with the understanding that $\Pi_0 = 0$. Equation (5) is a direct application of (4) in (1), but other reformulations as also convenient. Recombining terms, the evaluation can also be expressed as

$$\mathbb{E}_\Omega[u(\bar{x})] = \Omega(\Pi_N) u(x_N) - \Omega(\Pi_{N-1}) [u(x_N) - u(x_{N-1})] - \cdots - \Omega(\Pi_1) [u(x_2) - u(x_1)]$$

$$= u(x_N) - \sum_{n=1}^{N-1} \Omega(\Pi_n) [u(x_{n+1}) - u(x_n)].$$

(6)

So the evaluation is the utility of the best payoff less the difference in the utilities of the best two payoffs multiplied by the decision weight that the best payoff is not realized, less the difference in the utilities of the second and third best payoffs multiplied by the decision weight that the best two payoffs is not realized, etc.\(^4\)

For a continuous cumulative distribution $\Pi(x)$ with density $\pi(x)$, the resulting weighting density is

$$\omega(x) = \frac{d\Omega(\Pi(x))}{dx} = \Omega'(\Pi(x)) \pi(x).$$

(7)

Obviously to apply (7), $\Omega$ must be differentiable; though in a discrete state space, differentiability is not even relevant. If $\Omega$ is not differentiable, then the obvious interpretation of (7) will create atoms in the weighting density when none existed in the original continuous distribution. Using (7) and integrating by parts, the valuation of any risky prospect with a continuous distribution is

$$\mathbb{E}_\Omega[u(\bar{x})] = \int_a^b \omega(x) u(x) dx = \int_a^b \Omega'(\Pi(x)) \pi(x) u(x) dx$$

$$= \Omega(\Pi(x)) v(x) \bigg|_a^b - \int_a^b \Omega'(\Pi(x)) u'(x) dx = v(b) - \int_a^b \Omega(\Pi(x)) u'(x) dx.$$

(8)

This is the obvious continuous distribution analog of (5). In this integral the limit $a$ is any value less than or equal to the minimum outcome of $\bar{x}$ and $b$ is any value greater than or equal to the maximum outcome for $\bar{x}$.

It is important to note that a rank-dependent assignment of weights does not admit to assigning decision weights to events without knowing the payoffs associated with all possible events. For example, suppose a share of stock will be worth one of $S_1 < S_N$ given in increasing order. The decision weight for state 1 when pricing the stock would be $\Omega(\pi_1)$. However, the decision weight for pricing a put option on the stock would be $\Omega(1-\pi_1)$ because state 1 would then have the largest outcome. This becomes even more complicated in portfolio analysis because different portfolios order outcomes differently. If there are $S$ states, there are $S!$ possible orderings to consider. Fortunately, some orderings can be disregarded as shown later.

\(^4\) RDU can be based on complementary probabilities. In this case using the notation here, (6) is re-expressed as

$$\mathbb{E}_\Omega[u(\bar{x})] = u(x_1) + \sum_{n=1}^{N} [1 - \Omega(\Pi_n)][u(x_{n+1}) - u(x_n)].$$

This is the utility of the worst payoff plus the increase in utility to the second worst payoff multiplied by the decision weight that the worst payoff is exceeded, etc.
As mentioned previously a probability weighting method that assigns weights based only on the marginals necessarily admits to first-order stochastic dominance. Assigning decision weights based on the cumulative distribution as in (4) eliminates this problem as Theorem 11.1 shows.

**Theorem 11.1: Cumulative Weighting and First-Order Stochastic Dominance.** First-order stochastic dominance is preserved under cumulative probability weighting

\[ F \succ_{1SD} G \iff F \succ_{\Omega} G \quad \text{and} \quad F \succ_{1SD} G \iff F \succ_{\Omega} G \tag{9} \]

i.e., \( F \) first-order stochastically dominates \( G \) if and only if \( \Omega(F) \) first-order stochastically dominates \( \Omega(G) \) for all strictly increasing \( \Omega(\cdot) \) that map \([0,1]\) to \([0,1]\).

**Proof:** The backward implication (\( \iff \)) is obvious because the identity function \( \Omega(F) = F \) is a strictly increasing mapping giving the natural probabilities. The forward implication follows because \( \Omega \) is an increasing function so \( \Omega(G(z)) \geq \Omega(F(z)) \) if and only if \( G(z) \geq F(z) \). Therefore, \( G(z) \geq F(z), \forall z \iff \Omega(G(z)) \geq \Omega(F(z)), \forall z \). The first relation is the definition of first-order stochastic dominance given in Chapter 2. If \( F \) and \( G \) are the cumulative distributions of \( \bar{x} \) and \( \bar{y} \), the second relation guarantees that

\[ \mathbb{E}[u(\bar{x}) - u(\bar{y})] = \int_{a}^{b} [\Omega(G(z)) - \Omega(F(z))]u'(z)dz \geq 0. \tag{10} \]

The term in brackets is nonnegative as just shown, and \( u' \) is nonnegative for all increasing utility functions. Therefore, from (8) \( F \) is preferred to \( G \) for all cumulative weighting functions. \[ \blacksquare \]

Theorem 11.1 proves that any probability weighting method must be based on cumulative probabilities if it is to eliminate first-order stochastic dominances. So cumulative probability weighting as used in RDU is a necessary condition for (1) to represent increasing preferences. Sufficient conditions for RDU must be weaker than those required for EUT as the latter is a special case. Not surprisingly, it is the Independence Axiom that is replaced with something weaker.

A number of different equivalent axioms have been proposed to replace the Independence Axiom. The simplest to use, though not most intuitive, is the Distorted Independence Axiom.

**Distorted Independence Axiom:** For any cumulative probability distribution \( F \), define the weight-distorted probability distribution \( \Omega(F) \) where \( \Omega(0) = 0, \Omega(1) = 1 \), and \( \Omega \) is strictly increasing and continuous with a continuous inverse.\(^5\) Choices satisfy the Distorted Independence Axiom if for every \( \pi \in [0,1] \) and cumulative distributions \( F_1, F_2, G \)

\[ \Omega(F_1) \succ \Omega(F_2) \Rightarrow \Omega(\pi F_1 + (1 - \pi)G) \succ \Omega(\pi F_2 + (1 - \pi)G). \tag{11} \]

In other words, the Distorted Independence Axiom is the standard Independence Axiom applied to the distorted distributions. With this axiom, the proof of RDU follows immediately by applying the EUT proof using the \( \Omega \)-distorted decision weights \( \omega_n \).

**Theorem 11.2: Rank Dependent Utility.** If there is a preference relation \( \succ \) over a set of simple lotteries \( \mathcal{P} \) that is complete, transitive, always prefers first-degree stochastically dominating outcomes, and satisfies the Archimedean and Distorted Independence Axioms, then \( X \equiv \{p, x\} \succ Y \equiv \{q, y\} \) if and only if there exist a strictly increasing utility function, \( u \), and a strictly increasing probability weighting \( \Omega \) with \( \Omega(0) = 0 \) and \( \Omega(1) = 1 \) such that \( V(X) \geq V(Y) \) where

\(^5\) It is easy to show that \( \Omega(F) \) satisfies all conditions of a cumulative distribution and so may be treated as one.
\[ V(X) \equiv \sum \omega_n u(x_n) \quad \omega_n \equiv \Omega(P_n) - \Omega(P_{n-1}) \quad \text{and} \quad P_n \equiv \sum_{j=1}^{n} p_j. \]  

**Proof:** The proof is exactly like the proof of EUT using the decision weights in place of the probabilities. 

While the theorem is straightforward, the Distorted Independence Axiom lacks an underlying intuition. It describes properties of the weighting function which is to be derived, and therefore should be a derived property. An alternate equivalent axiom with more intuition is the Comonotonic Independence Axiom.

**Comonotonic Independence Axiom:** Let \( Z \equiv \{\pi, z\} \) represent a lottery that has ordered payoffs \( z_n \) with probabilities \( \pi_n \). Comonotonic independence is the property that for any two lotteries \( X = \{p, x\} \) and \( Y = \{q, y\} \) with \( x_n = y_m \), the preference \( X \succ Y \) is repeated as \( X^z \succ Y^z \) for the lotteries \( X^z \equiv \{p(x_1, \ldots, x_{n-1}, z, x_{n+1}, \ldots, x_N)\} \) and \( Y^z \equiv \{q(y_1, \ldots, y_{m-1}, z, y_{m+1}, \ldots, y_M)\} \) for any \( z \) in both ranges, \( z \in (x_{n-1}, x_{n+1}) \) and \( z \in (y_{m-1}, y_{m+1}) \). The same is true for the strict and indifference rankings, \( \succ \) and \( \sim \).

The condition that \( z \) lie in both intervals \( (x_{n-1}, x_{n+1}) \) and \( (y_{m-1}, y_{m+1}) \) ensures that the \( X \) and \( Y \) outcome orderings are preserved in \( X^z \) and \( Y^z \). Therefore the decision weights for the new lotteries remain the same as in the original, \( \omega_{X^z} = \omega_X \) and \( \omega_{Y^z} = \omega_Y \). This axiom is weaker than the Independence Axiom which requires the same preference ordering even if the common replacement \( z \) lies outside one of the two ranges and so changes the ordering of the outcomes. Unlike the decision weights, the probabilities are not affected by altering the outcome ordering so the valuation is not affected.

A proof of RDU based on the Comonotonic Independence Axiom can be found in Chateauneuf (1997). As with EUT an axiom like the sure-thing principle is required to prove RDU valid for continuous distributions.

**Rank Dependent Utility, Risk, and Risk Aversion**

Under EUT, a certain payment equal to the expected value of a random payoff is always preferred to the payoff itself. This is known as weak risk aversion. More generally the less risky or two prospects with the same expectation is always preferred. This is known as strong risk aversion. In other words, weak risk aversion can be described as an aversion to risk while strong risk aversion is an aversion to an increase in risk.

As shown in Theorem 2.2, both of these comparisons depend only on the concavity of the utility function so they are not really separate concepts under EUT. However, concave utility is insufficient to guarantee even weak risk aversion under RDU. Furthermore, weak and strong risk aversion are distinct under RDU. With some probability weighting structures, preferences can be weakly risk averse without being strongly risk averse. The restrictions on the weighting functions required for weak and strong risk aversion are pessimism and decreasing pessimism, respectively.

**Definitions: Pessimism and Increasing Pessimism.** A probability weighting function is pessimistic if \( \Omega(P) \geq P, \forall P \). That is for any cumulative probability, the cumulative decision weight that the realized outcome is less (worse) is no smaller than the actual probability. A weighting function displays increasing pessimism if it is concave. The weight applied for the next outcome with probability \( p \) at a cumulative probability \( P \) is \( \omega = \Omega(p + P) - \Omega(P) \). This is decreasing in \( P \) if \( \Omega'(p + P) - \Omega'(P) \leq 0 \); that is, if \( \Omega \) is concave.\(^6\)

\(^6\) Some authors order the outcomes from the best to the worst which of course changes these definitions.
Theorem 11.3: Rank Dependent Utility and Weak Risk Aversion. A concave utility function and a pessimistic weighting function are individually necessary and jointly sufficient for RDU choices to display weak (Arrow-Pratt) risk aversion with \( \mathbb{E}_\Omega[u(\hat{x})] \leq u(\mathbb{E}_\Omega[\hat{x}]) \) \( \forall \hat{x} \).

Proof: (Necessity) Because the identity function is a valid weighting function and gives weights equal to probabilities, EUT is a special case of RDU. So a concave utility function is a necessary condition if risk aversion is to apply for all valid weighting functions. Similarly a linear utility function is a weakly concave function and evaluates choices under RDU as \( \mathbb{E}_\Omega[\hat{x}] \).

So for weak risk aversion to hold, we must have \( \mathbb{E}_\Omega[\hat{x}] \leq \mathbb{E}[\hat{x}] \) because the latter is the evaluation of a deterministic gamble with a single payment of \( \mathbb{E}[\hat{x}] \). Now consider a gamble, \( \hat{x} \), that pays 1 with probability \( P \) and 0 with probability \( 1 - P \). Then

\[
\mathbb{E}_\Omega[\hat{x}] = 0 \cdot \Omega(P) + 1 \cdot [1 - \Omega(P)] \leq \mathbb{E}[\hat{x}] = 1 - P \\
\Rightarrow \Omega(P) \geq P.
\]

(Sufficiency) Sufficiency can be demonstrated in two steps. First, from Fubini's Theorem, expectations can be computed as

\[
\mathbb{E}[\hat{x}] = \int_{-\infty}^{\infty} [1 - F(x)]dx - \int_{-\infty}^{0} F(x)dx. \tag{14}
\]

So

\[
\mathbb{E}[\hat{x}] - \mathbb{E}_\Omega[\hat{x}] = \int_{-\infty}^{\infty} [\Omega(F(x)) - F(x)]dx + \int_{-\infty}^{0} [\Omega(F(x)) - F(x)]dx \geq 0. \tag{15}
\]

The inequality follows because the integrands are nonnegative by assumption. Therefore, \( \mathbb{E}_\Omega[\hat{x}] \leq \mathbb{E}[\hat{x}] \). Second, \( \mathbb{E}_\Omega[u(\hat{x})] \leq \mathbb{E}_\Omega[\hat{x}] \) by the concavity of \( u \) and Jensen’s inequality. Together, \( \mathbb{E}_\Omega[u(\hat{x})] \leq \mathbb{E}_\Omega[\hat{x}] \leq \mathbb{E}[\hat{x}] \) and weak risk aversion is verified.

The theorem does not say that weak risk aversion is only possible with both concave utility and concave probability weighting. There could be other combinations that result in risk aversion for some gambles. What it does say is that utility must be concave and the weighting function must be pessimistic if all risks are to be valued below their expected value. The same is true in the next theorem.

Under EUT, weak (Arrow-Pratt) risk aversion and strong (Rothschild-Stiglitz) risk aversion are the same. Both arise solely from concavity of the utility function. Under RDU, these two types of risk aversion are distinct. In particular a pessimistic weighting function with \( \Omega(P) \geq P \), is not sufficient for strong risk aversion.

To verify that these two conditions are not jointly sufficient for strong risk aversion consider the following RDU structure, \( u(x) = x \), which is weakly concave, and

\[
\Omega(P) = \begin{cases} 
2P & 0 \leq P \leq \frac{1}{4} \\
\frac{1}{2} + \frac{1}{2}P & \frac{1}{4} \leq P \leq \frac{1}{2} \\
P & \frac{1}{2} \leq P \leq 1 
\end{cases}, \tag{16}
\]

which is pessimistic, \( \Omega(P) \geq P \). A gamble that pays 10 with probability 0.4 and 20 with probability 0.6 has an expected payoff of 16, but due to the pessimism in the weighting function its RDU evaluation is 0.575 \cdot 10 + 0.425 \cdot 20 = 14.25 even though utility is linear. The gamble \( \hat{y} = \hat{x} + \hat{\epsilon} \) where \( \hat{\epsilon} = \pm 5 \) with equal probability when \( \hat{x} = 20 \) and 0 otherwise is riskier in a Rothschild-Stiglitz sense and must be valued below \( \hat{x} \) if strong risk aversion holds. The \( \hat{y} \) gamble has three outcomes \{10, 15, 25\} with probabilities \{0.4, 0.3, 0.3\}. The decision weights for those probabilities are \{0.575, 0.15, 0.275\}. So the RDU evaluation of \( \hat{y} \) is 14.875, and the riskier gamble is preferred. Strong risk aversion does not hold because \( \Omega \) has a convex portion that induces a local optimism about \( \hat{\epsilon} \). A tighter restriction than pessimism is required for strong risk aversion.
**Theorem 11.4: Rank Dependent Utility and Strong Risk Aversion.** A concave value function and an increasingly pessimistic (i.e., concave) weighting function are individually necessary and jointly sufficient for RDU choices to display strong (Rothschild-Stiglitz) risk aversion of rejecting any gamble in favor of a less risky one with the same expectation.

**Proof:** (Necessity) Because strong risk aversion implies weak risk aversion, concavity of the utility function is still necessary. However, pessimism in the weighting function is insufficient as demonstrated in the preceding example. As before, a linear utility function is a weakly concave function so for strong risk aversion to hold, we must have $E_\Omega[\hat{x} + \tilde{\varepsilon}] \leq E_\Omega[\hat{x}]$ whenever $E[\varepsilon | x] = 0$ for all risks $\hat{x}$.\(^7\)

Now consider a random payment, $\hat{x}$, that is 0 with probability $p - \pi$, $\frac{1}{2}$ with probability $1 - p - 2\pi$, and 1 with probability $1 - p$. The mean preserving spread, $\tilde{\varepsilon}$, takes on the values $\pm \frac{1}{2}$ with equal probabilities when $\hat{x} = \frac{1}{2}$ and is otherwise zero. So $\hat{x} + \tilde{\varepsilon}$ is equal to 0 with probability $p$ and to 1 with probability $1 - p$. The gamble $\hat{x}$ is less risky than $\hat{x} + \tilde{\varepsilon}$ so it is preferred under EUT. The two gambles can be made equivalent if one or more of the payoffs of $\hat{x}$ are reduced. The adjustment made here is to reduce the payoff $\frac{1}{2}$ which is the only difference between the gambles.\(^8\) The two gambles as adjusted are

$$
\tilde{x} = \begin{cases}
\frac{1}{2} & 1 - p - \pi \\
\frac{1}{2} - \delta_\Omega & 2\pi \\
0 & p - \pi
\end{cases}
$$

$$
\tilde{x} + \tilde{\varepsilon} = \begin{cases}
1 & 1 - p \\
0 & p
\end{cases}.
$$

The two decision-weight expectations are

$$
E_\Omega[\tilde{x} + \tilde{\varepsilon}] = \Omega(p) \cdot 0 + [1 - \Omega(p)] \cdot 1 = 1 - \Omega(p).
$$

and

$$
E_\Omega[\tilde{x}] = \Omega(p - \pi) \cdot 0 + [\Omega(p + \pi) - \Omega(p - \pi)] \cdot \frac{1}{2} + [\Omega(1) - \Omega(p + \pi)] \cdot 1
$$

$$
= [\Omega(p + \pi) - \Omega(p - \pi)] \cdot \left(\frac{1}{2} - \delta_\Omega\right) + 1 - \Omega(p + \pi).
$$

The risk premium, $\delta_\Omega$, is the fixed reduction in the safe payment of $\frac{1}{2}$ required to equate the two expectations in (18) and (19)

$$
\delta_\Omega = \frac{\Omega(p) - \frac{1}{2} \left[\Omega(p - \pi) + \Omega(p + \pi)\right]}{\Omega(p + \pi) - \Omega(p - \pi)} = \frac{1}{2} \frac{\left[\Omega(p) - \Omega(p - \pi)\right] - [\Omega(p + \pi) - \Omega(p)]}{\Omega(p + \pi) - \Omega(p - \pi)}.
$$

The numerator is a second difference and the denominator is twice a first difference. So for small $\pi$, the risk premium is

$$
\delta_\Omega \approx -\frac{1}{4} \pi \frac{\Omega''(p)}{\Omega'(p)}.
$$

A necessary condition for the safer prospect to be preferred is that the risk premium is positive. As $\Omega$ is an increasing function, $\Omega''(p)$ must be negative meaning $\Omega$ is locally concave. As this must be true for all $p$, $\Omega$ must be globally concave.

(Sufficiency). Let $F$ and $G$ be two gambles with the same expectation and with $F$ less risky in a Rothschild-Stiglitz sense. Then $F$ second-degree stochastically dominates $G$ under the natural probabilities. Using the quantile description of second degree stochastic dominance as

\(^7\) Note that the increased riskiness is an objective statement and defined with the actual probabilities not the decision weights; i.e., it is not assumed that $E_\Omega[\varepsilon | x] = 0$.

\(^8\) Showing that any (or all) payoffs of $\tilde{x}$ must be reduced to equate the expectations proves that $\tilde{x}$ is preferred. It is convenient for later analysis to reduce only the payoff of $\frac{1}{2}$, the one that is altered between the gambles.
given in equation (25) of chapter 2,
\[ 0 \leq H(P^*) = \int_0^{P^*} \left[ F^{-1}(P) - G^{-1}(P) \right] dP \quad \forall P^*. \quad (22) \]

Under probability weighting, the relevant distributions are \( \Omega(F) \) and \( \Omega(G) \). So \( F \) will be preferred for all concave utilities if
\[ 0 \leq \int_0^{P^*} \left[ F^{-1}(\Omega^{-1}(P)) - G^{-1}(\Omega^{-1}(P)) \right] dP \quad \forall P^*. \quad (23) \]

Use the change in variable \( Q = \Omega^{-1}(P) \) so \( P = \Omega(Q) \) and \( dP = \Omega'(Q)dQ \). The integrating (23) by parts gives
\[ \int_0^{\Omega^{-1}(P^*)} \left[ F^{-1}(Q) - G^{-1}(Q) \right] \Omega'(Q)dQ = H(Q)\Omega'(Q)\bigg|_0^{\Omega^{-1}(P^*)} - \int_0^{\Omega^{-1}(P^*)} H(Q)\Omega''(Q)dQ. \quad (24) \]

\( H \) is nonnegative everywhere per (22) and \( H(0) = 0 \). \( \Omega' > 0 \) for all valid weighting functions so the first term on the right-hand side of (24) is nonnegative. In the second term \( H \geq 0 \) and \( \Omega'' \leq 0 \), so that integral is non-positive, and the second term is also nonnegative. Therefore, \( F \) stochastically dominates \( G \) under the probability weighting as well.

Theorem 11.4 proved that Rothschild-Stiglitz increases in risk are disliked under RDU, but they said nothing about the magnitude of the dislike. An Arrow-Pratt type of analysis can determine the risk premium for small risks. Equation (21) gave the risk premium for small probability risks under risk neutrality. Under risk aversion, the premium is larger.

To determine the risk premium for a small increase in risk when the investor is not risk neutral, small changes in both the probabilities and the payoffs are required. As in the Arrow-Pratt exercise, the risk premium is determined by equating the “expected” utilities of two gambles. Because weak and strong risk aversion are not the same, two risky gambles must be compared.

\[ \tilde{x} = \begin{cases} h & 1 - p - \pi \\ \hat{x} - \delta & 2\pi \\ \ell & p - \pi \end{cases} \quad \tilde{y} = \begin{cases} h & 1 - p - \pi \\ \hat{x} + \varepsilon & \pi \\ \hat{x} - \varepsilon & \pi \\ \ell & p - \pi \end{cases} \quad (25) \]

The riskier gamble, \( \tilde{y} \), has the intermediate outcome split increasing the risk in \( \tilde{x} \). The safer gamble has the intermediate outcome reduced by the risk premium.

The difference in the decision weighted utilities depends only on the intermediate outcomes near \( \hat{x} \). It is
\[ \mathbb{E}_\omega[\tilde{x} - \tilde{y}] = [\Omega(p + \pi) - \Omega(p - \pi)]u(\hat{x} - \delta_{\text{RDU}}) - [\Omega(p) - \Omega(p - \pi)]u(\hat{x} - \varepsilon) - [\Omega(p + \pi) - \Omega(p)]u(\hat{x} + \varepsilon). \quad (26) \]

The risk premium is the value of \( \delta \) that sets this difference to 0. Making the usual Taylor expansion in \( u \) and collecting term gives
\[ 0 \approx [\Omega(p + \pi) - \Omega(p - \pi)]\left[u(\hat{x}) - \delta_{\text{RDU}}u'(\hat{x})\right] - [\Omega(p) - \Omega(p - \pi)]\left[u(\hat{x} - \varepsilon)u'(\hat{x}) + \frac{1}{2}u''(\hat{x})\right]
\[ - [\Omega(p + \pi) - \Omega(p)]\left[u(\hat{x}) + \varepsilon u'(\hat{x}) + \frac{1}{2}\varepsilon^2 u''(\hat{x})\right]
\[ = -[\Omega(p + \pi) - \Omega(p)]\left[\delta_{\text{RDU}}u'(\hat{x}) - \frac{1}{2}\varepsilon^2 u''(\hat{x})\right] + [2\Omega(p) - \Omega(p - \pi) - \Omega(p + \pi)]\varepsilon u'(\hat{x}) \quad (27) \]

\[ \Rightarrow \delta_{\text{RDU}} = -\frac{1}{2} \frac{\varepsilon u''(\hat{x})}{u'(\hat{x})} + \varepsilon \frac{2\Omega(p) - \Omega(p - \pi) - \Omega(p + \pi)}{\Omega(p + \pi) - \Omega(p - \pi)} \]
The first term is the standard Arrow-Pratt risk premium for concave utility evaluated at the point where the risk was introduced. The second term is the risk-premium determined in (20), $\delta\Omega$, multiplied by twice the magnitude of the risk, $2\varepsilon$. So for a small risk with standard deviation $\sigma$ centered at $\hat{x}$, the RDU risk premium is

$$\delta_{\text{RDU}} = \delta_{\text{EUT}} + 2\sigma\delta\Omega.$$  \hspace{1cm} (28)

**Monotonic Risk**

In RDU, weak and strong risk aversion are distinct concepts and both of them require assumptions about the probability weighting function. A natural question is can a stronger definition of risk produce risk aversion with no additional assumptions on the probability weighting function? The answer to the question is yes; RDU investors are averse to increases in monotonic riskiness for any weighting function provided only that their utility is concave.

A random variable $\tilde{y}$ with cumulative distribution $G$ is monotonic riskier than $\tilde{x}$ with cumulative distribution $F$ if for any quantile range, $0 \leq Q < P \leq 1$, the spread in outcomes required is larger. That is

$$G^{-1}(Q) - G^{-1}(P) \geq F^{-1}(Q) - F^{-1}(P), \quad \forall \ 0 \leq Q < P \leq 1. \hspace{1cm} (29)$$

The empirical validity of monotonic risk is illustrated by the following choices.

<table>
<thead>
<tr>
<th>prob:</th>
<th>2%</th>
<th>1%</th>
<th>47%</th>
<th>47%</th>
<th>1%</th>
<th>2%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{x}$</td>
<td>-1000</td>
<td>0</td>
<td>-50</td>
<td>50</td>
<td>0</td>
<td>1000</td>
</tr>
<tr>
<td>$\tilde{y}$</td>
<td>-1000</td>
<td>-1000</td>
<td>0</td>
<td>0</td>
<td>1000</td>
<td>1000</td>
</tr>
</tbody>
</table>

When questioned, many people pick $\tilde{y}$ over $\tilde{x}$ even though it is Rothschild-Stiglitz riskier. The MPS that creates $\tilde{y}$ from $\tilde{x}$ takes the values $\mp 950$ and $\pm 1050$ with probabilities 0.495 and 0.505 when $\tilde{x} = \mp 50$. However, $\tilde{y}$ is not monotonic riskier than $\tilde{x}$ because

$$G^{-1}(95\%) - G^{-1}(5\%) = 0 - 0 = 0 < F^{-1}(95\%) - F^{-1}(5\%) = 50 - (-50) = 100.$$  \hspace{1cm} (30)

Apparently the concentration of the outcomes of $\tilde{y}$ nearer the mean value of 0 appears to make it less risky. From this example, the important difference between monotonic and Rothschild-Stiglitz riskiness is clear. Monotonic risk spreads all of the outcomes rather than just some of them.

When increased riskiness is defined as monotonic rather than Rothschild-Stiglitz, then aversion to an increase in risk requires only a concave utility function. No restrictions on the probability weighting function are required.

**Theorem 11.5: Rank Dependent Utility and Monotonic-Risk Aversion.** A concave value function is a necessary and sufficient condition for RDU choices to display an aversion to any increase in monotonic risk.

**Proof:** Necessity follows immediately because an increase in monotonic risk is an increase in Rothschild-Stiglitz risk and a concave utility function is necessary for the latter to be disliked. Sufficiency follows from the definition of monotonic risk in (29). Setting $Q = 0$, and rearranging terms gives

$$F^{-1}(0) - G^{-1}(0) \leq F^{-1}(P) - G^{-1}(P). \hspace{1cm} (31)$$

If $G$ is monotonically riskier than $F$, then the domain of $G$ must include the entire domain of $F$ so the left-hand side of (31) is nonnegative. This means that $F^{-1}(P) - G^{-1}(P) \geq 0$ for all $P$. Therefore, $\int_{-\infty}^{0}[F^{-1}(p) - G^{-1}(p)] dp \geq 0, \forall P$ because the integrand is nonnegative for every $p$. This is the
quantile condition for second-order stochastic dominance. So \( \tilde{x} \) is preferred by all concave utility functions.

**Inverse-S Probability Weighting**

Based on experiments, the weighting function is not concave, but has an inverse S shape, concave for low probabilities and convex for higher ones as shown in the figure. This means that its slope is highest near the endpoints of its domain so the worst and best outcomes are typically over-weighted relative to their actual probabilities unless they are fairly common. The particular weighting function shown is the one proposed by Tversky and Kahneman (1992) for CPT:

\[
\Omega(P) = \frac{P^\delta}{[P^\delta + (1-P)^\delta]^{1/\delta}} \quad 0 < \delta \leq 1. \tag{32}
\]

Several other weighting functions have also been proposed:

- Lattimore, Baker, & Witte (1992): \[\frac{\gamma P^\delta}{\gamma P^\delta + (1-P)^\delta} \quad 0 < \delta, \gamma > 0\]
- Wu & Gonzalez (1996): \[\Omega(P) = \frac{P^\delta}{[P^\delta + (1-P)^\delta]^{\gamma}} \quad 0 < \delta, \gamma \geq 1 \tag{33}\]
- Prelec (1998): \[\Omega(P) = \exp[-\beta(-\elln P)^\alpha] \quad \alpha, \beta > 0\]

The TK weighting function is a special case Wu and Gonzalez with \( \gamma = 1/\delta \). For \( \gamma = 1 \), the Wu and Gonzalez weighting function is the same as the Lattimore, Baker, and Witte weighting function. Karmarkar (1978) had previously proposed this special case.

![Probability Weighting Function](image)

**Figure 11.1: The Tversky-Kahneman Probability Weighting Function**

This figure illustrates the probability weighting function suggested by Tversky and Kahneman. It has an inverted S shape that emphasizes the probabilities for the extreme high and low outcomes.

Another particularly simple cumulative weighting function with an inverse S-shape is **neo-linear weighing**

---

9 The TK weighting function is not monotone for all parameter values, and so can assign negative decision weights. For example, for \( \delta = 0.25 \), \( \Omega(P) \) is decreasing over the range of cumulative probabilities \( 1.56% < P < 23.62% \). It is monotone for all values of \( \delta \) greater than the root of the equation \((1-\delta)^{2-\delta} = \delta^{1-\delta}\) with solution \( \delta \approx 0.279 \). See Ingersoll (2008).
Whenever there are three or more outcomes, each outcome other than the best and the worst is assigned a decision weight that is smaller than its probability by the same proportion, \( \beta \)

\[
\omega_n \equiv \Omega(P_n) - \Omega(P_{n-1}) = \beta (P_n - P_{n-1}) = \beta p_n \quad \forall 1 < n < N.
\] (35)

Either the worst outcome or the best outcome or both have decision weights in excess of their probabilities. The decision weight for the worst outcome exceeds its probability when \( p_1 < \alpha/(1 - \beta) \). The best outcome, \( N \), is over-weighted when \( p_N < (1 - \alpha - \beta)/(1 - \beta) \).

The evaluation of any gamble under neo-linear probability weighting is a weighted average of the unweighted expected utility and the utilities of the best and worst outcomes

\[
\mathbb{E}_{\Omega}[u(\tilde{x})] = \sum_{n=1}^{N} \omega_n u(x_n) = (\alpha + \beta \omega(x_1) + \beta \sum_{n=2}^{N-1} p_n u(x_n) + [1 - \alpha - \beta p_N] u(x_N)
\]

\[
= \sum_{n=1}^{N} (\alpha + \beta \omega(x_1) + (1 - \alpha - \beta) u(x_N) = \beta \mathbb{E}[u(\tilde{x})] + \alpha u(x_1) + (1 - \alpha - \beta) u(x_N).
\] (36)

### Rank Dependent Utility and the Paradoxes

Rank Dependent Utility can be used to explain a number of the paradoxes that have been observed. RDU uses a weakened independence axiom, and can explain the Allais Paradox and the Common Ratio Effect which both exhibit violations of that axiom. The Ellsberg Paradox is also not necessarily a paradox under RDU.

The Allais Paradox, as described in Chapter 2, is that most people violate the Independence Axiom by showing a preference for \( B \) over \( A \) and \( C \) over \( D \) in the following choices

\[
A: \begin{cases} 
33\% \text{ chance of } & 2500 \\
66\% \text{ chance of } & 2400 \\
1\% \text{ chance of } & 0
\end{cases} \quad B: 100\% \text{ chance of } 2400
\]

\[
C: \begin{cases} 
33\% \text{ chance of } & 2500 \\
67\% \text{ chance of } & 0
\end{cases} \quad D: \begin{cases} 
34\% \text{ chance of } & 2400 \\
66\% \text{ chance of } & 0
\end{cases}
\] (37)

Eliminating the common 66\% chance of winning 2400 from \( A \) and \( B \) leaves a 33 to 1 chance of winning 2500 in \( A \) as worse than getting 2400 for sure in \( B \). Similarly eliminating the common 66\% chance of 0 from \( C \) and \( D \) again leaves a 33 to 1 chance of winning 2500 in \( C \) and a sure 2400 in \( B \). So both the only possible EUT choices are \( A \) and \( C \) or \( B \) and \( D \). But those choices are possibly in RDU.

In more detail, assign \( u(2400) = 0 \) without loss of generality and denote \( u_H \equiv u(2500) > 0 \) and \( u_L \equiv u(0) < 0 \). Under RDU, these choices impose a the conditions

\[
B > A \iff u_M = 0 > [1 - \Omega(0.67)] u_H + 0 + \Omega(0.01) u_L
\]

\[
C > D \iff [1 - \Omega(0.67)] u_H + \Omega(0.67) u_L > 0 + \Omega(0.66) u_L
\]

Recalling that \( u_L < u_M = 0 \), this gives
\[ \Omega(0.67) - \Omega(0.66) < \frac{[1 - \Omega(0.67)]u_M}{-u_L} < \Omega(0.01). \] (39)

Under EUT, the weighting function is the identify function so both the left-hand side and right-hand side are 0.01 leading to a contradiction. But for other weighting functions, (39) can be satisfied. What is required is that an initial 1% chance has a higher decision weight than the one percentage point increase from 66% to 67%. For example, taking \( \Omega(P) = P^{0.5} \), the left- and right-hand sides of (39) are 0.006 and 0.1. For the Tversky-Kahneman weighting function in (32) with \( \delta = 0.6 \) the left- and right-hand sides of (39) are 0.006 and 0.058. So the two inequalities are definitely possible.

Bergen’s Paradox, or the Common Ratio Effect, described in [Chapter 1] is also possibly explained by RDU. Most participants in Common-Ratio experiments select \( A \) and \( D \) when offered the following types of choices

\[
A: \begin{cases} \text{prob. } p \text{ of } H \\ \text{prob. } 1-p \text{ of } L \end{cases} \quad \text{or} \quad B: \begin{cases} \text{prob. } q \text{ of } M \\ \text{prob. } 1-q \text{ of } L \end{cases}
\]

\[
C: \begin{cases} \text{prob. } \alpha p \text{ of } H \\ \text{prob. } 1-\alpha p \text{ of } L \end{cases} \quad \text{or} \quad D: \begin{cases} \text{prob. } \alpha q \text{ of } M \\ \text{prob. } 1-\alpha q \text{ of } L \end{cases}
\]

for \( L < M < H, 0 < p < q < 1, \) and \( 0 < \alpha < 1. \) With no loss of generality, we can set \( u_L = 0 \) so under EUT these choices require

\[ A > B \Rightarrow pu_H > qu_M \quad D > C \Rightarrow \alpha pu_H < \alpha qu_M, \] (41)

which are clearly contradictory under EUT.

With only two outcomes, the decision weight of the better outcome with probability \( \pi \) is \( \omega \pi = 1 - \Omega(1-\pi) \). The choices of \( A \) over \( B \) and \( D \) over \( C \) under RDU requires that

\[ A > B \Rightarrow \omega_p u_H > \omega_q u_M \quad \text{and} \quad D > C \Rightarrow \omega_{\alpha p} u_H < \omega_{\alpha q} u_M. \] (42)

These are both true when

\[ \frac{\omega_{\alpha p}}{\omega_{\alpha q}} < \frac{u_M}{u_H} < \frac{\omega_p}{\omega_q} \] (43)

which is possible. For example, for \( \Omega(\Pi) = \Pi^\gamma \), the decision weight for the better outcome is \( \omega \pi = 1 - (1-\pi)^\gamma \). When \( \gamma < 1 \), the right-hand side always exceeds the left-hand side so the utility ratio could fall between them. This is a pessimistic weighting functions as \( \Pi < \Pi^\gamma \), and the difference between the decision weight is less than the probability of winning \( \omega \pi - \pi = (1-\pi) - (1-\pi)^\gamma < 0 \).

The Ellsberg Paradox is a bit different. It does not involve a violation of the Independence Axiom but rather an inconsistent belief about probabilities. Because RDU uses decision weights rather than probabilities, it can also explain the Ellsberg Paradox. However, it requires a compound-lottery interpretation.

In the Ellsberg Paradox, there are two decisions. Each consists of a choice between two gambles with possible payouts of \( x \) and 0. In each decision, one gamble has a known probability of winning, and the other has an unknown probability of winning that lies in a range centered on the known probability that is presumably symmetric about its center. It is not crucial to the paradox that the symmetry is believed.

The paradox arises because the unknown win probabilities are related, and there is no possible belief about the random probability for which the usual choices are both optimal. In fact, one of the choices must violate first-order stochastic dominance for any subjective probabil-
probability belief.

Let \( \Pi(p) \) be the cumulative probability distribution of the unknown win probability, \( p^* \), with marginals \( \pi_m \). To simplify the analysis, set \( u(0) = 0 \). Under EUT, the evaluations of the two gambles are

\[
V^{\text{known}}_{\text{EUT}}(p) = pu(x) \quad V^{\text{unknown}}_{\text{EUT}}(\tilde{p}) = u(x)\mathbb{E}[\tilde{p}] = u(x)\sum_{m=1}^{M} p_m \pi_m. \tag{44}
\]

The gamble with the known winning probability is selected if and only if \( p \geq \Pi[p^*] \). As always, an EUT valuation is linear in the probabilities.

Under RDU, the gamble with a known win probability is evaluated as

\[
V^{\text{known}}_{\text{RDU}}(p) = u(0)\Omega(1-p) + u(x)[1-\Omega(1-p)] = \omega_p u(x) \tag{45}
\]

where \( \omega_p \equiv 1 - \Omega(1-p) \) is the decision weight of a win as before. Using the same weighted average evaluation, RDU assesses the gamble with an unknown win probability to be

\[
V^{\text{unknown}}_{\text{RDU}}(\tilde{p}) = \sum_{m=1}^{M} V^{\text{known}}_{\text{RDU}}(p_m)\pi_m = u(x)\sum_{m=1}^{M} \omega_{p_m} \pi_m. \tag{46}
\]

Again the utility of the amount won plays no role in the choice; both are proportional to \( u(x) \). All that matters is whether the known decision weight, \( \omega_p \), is greater or less than the expected value of the unknown decision weight, \( \mathbb{E}[\tilde{\omega}_p] = \sum \pi_m \omega_{p_m} \).

In the Ellsberg Paradox, it is reasonable to assume that the probability distribution for the unknown win probability is centered at the known probability of its alternative; i.e., \( p = \mathbb{E}[\tilde{p}] \equiv p^* \). The comparison of the known-probability gamble to the unknown-probability gamble is then

\[
V^{\text{unknown}}_{\text{RDU}}(\tilde{p}) \geq V^{\text{known}}_{\text{RDU}}(\tilde{p}) \iff u(x)\sum_{m=1}^{M} \omega_{p_m} \pi_m \geq u(x)\omega_p. \tag{47}
\]

By Jensen’s inequality the right-hand side is larger, and therefore the gamble with the known probability is preferred if \( \omega \) is a concave function of \( p \). As \( \omega \equiv 1 - \Omega(1-p) \), it is concave when \( \Omega \) is convex. Of course, that is only a sufficient condition to explain the paradox. It is not necessary.

While this analysis explains the Ellsberg Paradox, it seems to be internally inconsistent. Probability weighting was used to modify the probabilities of the distribution of the payoff, but it was not used to modify the distribution of the winning probabilities. As a higher win probability is clearly better, it seems natural under RDU order the possibilities of the win probability as \( p_1 < p_2 < \ldots < p_M \). Denoting the cumulative probability of the win probabilities by \( \Pi(p) \), the RDU evaluation of the gamble with the unknown win probability is

\[
V^{\text{unknown}}_{\text{RDU}}(\tilde{p}) = \sum_{m=1}^{M} V^{\text{known}}_{\text{RDU}}(p_m)[\Omega(\Pi_m) - \Omega(\Pi_{m-1})] = u(x)\sum_{m=1}^{M} \omega_{p_m} [\Omega(\Pi_m) - \Omega(\Pi_{m-1})] \tag{48}
\]

instead of (46). The probability weighting function for the \( \tilde{p} \) and \( \tilde{x} \) distributions could be the same or different.

As an illustration consider the Ellsberg Paradox presented in Chapter 1. There are 300 balls in an urn. One hundred of the balls are red, and the remaining 200 balls are white and blue in unknown proportions. As described in the table, Participants choose between \( A \), winning $100 if red is picked, and \( B \), winning $100 if white is picked. They are also choose between \( C \), winning $100 if either red

<table>
<thead>
<tr>
<th>Amount Won</th>
<th>red</th>
<th>white</th>
<th>blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A )</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( B )</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>( C )</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>( D )</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
or blue is picked, and $D$, winning $100 if white or blue is picked.

Suppose the participants believe there are either 50 white balls and 150 blue balls or vice versa with equal probability, and have a weighting function $\Omega(P) = P^{1/2}$. Then for choice $A$, $\omega_0$ is $1 - (1 - 1/3)^{1/2} = 0.184$. For $D$, $\omega_0 = 1 - (1 - 2/3)^{1/2} = 0.423$. For $B$ the two possible winning probability weights are $1 - (1 - 1/6)^{1/2} = 0.087$ and $1 - (1 - 1/2)^{1/2} = 0.293$. For $C$, they are $1 - (1 - 1/2)^{1/2} = 0.293$ and $1 - (1 - 5/6)^{1/2} = 0.592$. In each case the two probabilities are assumed equally likely so the smaller one has a decision weight of $\Omega(1/2) = 0.707$. The valuations of $B$ and $C$ are then

$$V^B = 0.707 \cdot 0.087 + 0.293 \cdot 0.293 = 0.147 < V^A = 0.184$$

$$V^C = 0.707 \cdot 0.293 + 0.293 \cdot 0.592 = 0.380 < V^D = 0.423.$$  

(49)

Note that this $\Omega$ is concave so $\omega$ is convex and weighting the decision weights with probabilities as in (47) rather than (48) would give the opposite ordering.

For this problem the ranking to be used on the $p$ outcomes seems obvious. Higher $p$ is better so, they should be ranked from smallest to largest. However, if there are more than two outcomes for $\hat{x}$, the application in (48) presents some problems. The probability distributions on $\tilde{p}$ need to be ordered from worst to best, and there is no obvious way to do this. Except for changes in the probability of the best or worst outcomes, it is not clear if an increase in the probability of any of the intermediate $\hat{x}$ outcome is good or bad.

**RDU Portfolios in a Complete Market**

The portfolio problem under RDU is the same as the standard problem, but uses decision weights instead of probabilities. However, there is one important difference. As the weighting function is applied to the cumulative probability, the weighting changes when portfolios with different outcome orderings are evaluated. We know that in a complete market, EUT maximizers with homogeneous beliefs all hold portfolios whose outcomes are ordered identically, but we do not know that to be true for RDU investors. And in an incomplete market, it certainly need not be true because it does not hold for EUT investors.

Once the decision weights have been determined, the optimization problem for a RDU investor in a complete market is

$$\max_{x_s} \sum \omega_s u(x_s) \quad \text{subject to} \quad \sum x_s q_s = w_0 \quad \text{and} \quad x_s \leq x_{s+1}. \quad (50)$$

The last set of constraints ensures that the ordering of outcomes is the same as the one used in computing the decision weights. The entire solution to the problem requires that all $S$ factorial permutations of the outcomes with their weights be considered. The optimal portfolio is the one that satisfies (50) and has the highest evaluation after all permutations are considered, though in practice, not all permutations usually need to be examined explicitly.

The Lagrangian for this problem is

$$\mathcal{L} = \sum_{s=1}^{S} \omega_s u(x_s) + \eta \left[ w_0 - \sum_{s=1}^{S} q_s x_s \right] + \sum_{s=1}^{S-1} \kappa_s (x_{s+1} - x_s). \quad (51)$$

The final $S-1$ constraints are inequality constraints and the Kuhn-Tucker conditions must be applied. The multipliers $\eta$ and $\kappa$ are nonnegative. The first-order conditions are

$$0 = \partial \mathcal{L} / \partial x_s = \omega_s u'(x_s) - \eta q_s - \kappa_s + \kappa_{s-1} \quad \text{for } s = 1, \ldots, S$$

$$0 \leq \partial \mathcal{L} / \partial \kappa_s = x_{s+1} - x_s \quad \text{← ordering constraint} \quad (52)$$

$$0 = \kappa_s (x_{s+1} - x_s) \quad \text{for } s = 0, \ldots, S-1 \quad \text{← } \kappa_s = 0 \text{ or } x_{s+1} = x_s$$
along with the budget constraint. When \( x_{s+1} > x_s \) so the ordering constraint is not binding, then \( \kappa_s = 0 \). When the constraint is binding, \( x_s = x_{s+1} \) and typically \( \kappa_s > 0 \) so that \( u'(x_s) = \eta q_s / \omega_s + \kappa_s > \eta q_s / \omega_s \). That is, the investor would like to increase \( x_s \) but “cannot” do so because of the constraint. The constraint is self-imposed; \( x_s \) can, of course, be increased, but this will change the outcome orderings which may change the weights creating a new problem to be solved.

For concave utility, marginal utility is invertible so

\[
x_s^* = u'^{-1}\left( (\eta q_s / \omega_s - \kappa_s) / \omega_s \right).
\]

If \( x_s^* < x_{s+1}^* \), then \( \kappa_s = \kappa_{s+1} = 0 \), and

\[
x_s^* = u'^{-1}(\eta q_s / \omega_s).
\]

This is just the standard result with the decision weight replacing the probability. If some consecutive states, \( s' \) through \( s'' \), have the same payoff, then \( \kappa_{s'} \) through \( \kappa_{s''-1} \) can be positive, and

\[
\begin{align*}
x_s^* &= \cdots = x_{s'}^* = u'^{-1} \left( (\eta q_s / \omega_s - \kappa_s) / \omega_s \right) \\
x_s^* &= \cdots = x_{s''}^* = u'^{-1} \left( (\eta q_s / \omega_s + \kappa_s) / \omega_s \right) \\
\end{align*}
\]

The inequalities follow because the multipliers are nonnegative and marginal utility is decreasing. The first-order conditions can therefore be expressed as

\[
u'^{-1}(\eta q_s / \omega_s) \leq x^* \leq u'^{-1}(\eta q_s / \omega_s)
\]

with strict equality holding when \( x_{s-1}^* < x_s^* < x_{s+1}^* \).

As an example, consider the following simple economy. States \( a, b, \) and \( c \) have probabilities of 20%, 30% and 50% and state prices of 0.3, 0.3 and 0.4, respectively. The SDF, or state price per unit probability, is highest in state \( a \) and lowest in state \( c \) so the optimal portfolio for a risk-averse expected utility maximizer has its outcomes ordered inversely, \( x_a < x_b < x_c \).

Consider an investor with utility \( u(x) = x^{0.6} \) using the TK probability weighting function given in (32) with \( \delta = 0.7 \). Assuming his portfolio returns are also ordered \( x_a < x_b < x_c \), then the decision weights are

\[
\begin{align*}
\omega_a &= \Omega(0.2) = 0.2560 \\
\omega_b &= \Omega(0.5) - \Omega(0.2) = 0.2013 \\
\omega_c &= 1 - 0.2560 - 0.2013 = 0.5426
\end{align*}
\]

Using these weights in (53) and ignoring the ordering constraint (i.e., setting \( \kappa_s = 0 \)), gives an optimal portfolio of \( x' = (0.575, 0.315, 1.837) \) as shown in columns four through six in the upper panel of Table 9.1. Unfortunately, these payoffs are ordered \( x_b < x_a < x_c \), not the assumed order.

State \( b \) has the smallest outcome so the weight for it should be \( \omega_b = \Omega(0.3) = 0.3281 \). The weight for state \( a \) should be \( \omega_a = \Omega(0.5) - \Omega(0.3) = 0.1293 \). The weight for state \( c \) remains the same. Columns seven and eight, under corrected order, recalculate the expected utility for the portfolio using these new decision weights.

But this is not the end of the problem. The portfolio was not optimized for these new decision weights. The second panel of the table repeats the optimization problem assuming \( x_b < x_a < x_c \). The optimized portfolio is now \( x' = (0.985, 0.096, 1.689) \). Once again the outcome ordering differs from the assumption; it insists on the original ordering. Consequently, the optimal portfolio under this probability weighting must have \( x_a = x_b^* \). 10

To solve the problem, states \( a \) and \( b \) are merged into a single state with probability 0.5 and state price 0.6. The decision weight is \( \omega_{ab} = \Omega(0.5) = 0.4574 \). Columns nine through eleven in both panels show the final optimized portfolio.

10 As the ordering constraint is binding, the multiplier, \( \kappa_{ab} \), associated with this constraint is nonzero to satisfy the first-order condition.
weighting function with function has an inverted S shape with only a single crossing of the 45° line. The effect must always be present in at least one of the tails if the cumulative weighting expected-utility maximizing portfolio for the extreme outcome. For example, this occurs in the lower tail for the TK EUT portfolios. For an inverted S-shaped probability weighting, typically outcomes in states \(s\) leads to a contradiction with the first-order conditions for the optimal portfolio indicating the opposite ordering. Therefore, the true optimal portfolio must have equal outcomes in states \(a\) and \(b\).

### Table 9.1: Optimal RDU Portfolio

This table determines the optimal portfolio for an investor with a utility function \(v(x) = x^{0.6}\) who uses a TK probability weighting function (equation (31)) with parameter \(\delta = 0.7\). The two sections of the table demonstrate that either assumed ordering of the state’s returns \((a, b, c)\) or \((b, a, c)\) leads to a contradiction with the first-order conditions for the optimal portfolio indicating the opposite ordering. Therefore, the true optimal portfolio must have equal outcomes in states \(a\) and \(b\).

<table>
<thead>
<tr>
<th>State</th>
<th>Assumed order (x_a &lt; x_b &lt; x_c)</th>
<th>Corrected order (x_b &lt; x_a &lt; x_c)</th>
<th>Constrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>(w(a,b,c)) (x^*) (w\cdot v(x))</td>
<td>(w(b,a,c)) (w\cdot v(x))</td>
<td>(w) (x^*) (w\cdot v(x))</td>
</tr>
<tr>
<td>20%</td>
<td>25.60% 0.575 0.184</td>
<td>12.93% 0.093</td>
<td>45.74% 0.437 0.278</td>
</tr>
<tr>
<td>(b)</td>
<td>20.13% 0.315 0.101</td>
<td>32.81% 0.164</td>
<td>54.26% 1.845 0.784</td>
</tr>
<tr>
<td>(c)</td>
<td>54.26% 1.832 0.780</td>
<td>54.26% 0.780</td>
<td>54.26% 1.845 0.784</td>
</tr>
<tr>
<td>50%</td>
<td>“(\mathbb{E}_w) (v(\cdot)) = 1.065”</td>
<td>“(\mathbb{E}_w) (v(\cdot)) = 1.037”</td>
<td>“(\mathbb{E}_w) (v(\cdot)) = 1.062”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>State</th>
<th>Assumed order (x_b &lt; x_a &lt; x_c)</th>
<th>Corrected order (x_a &lt; x_b &lt; x_c)</th>
<th>Constrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>(b)</td>
<td>(w(b,a,c)) (x^*) (w\cdot v(x))</td>
<td>(w(a,b,c)) (w\cdot v(x))</td>
<td>(w) (x^*) (w\cdot v(x))</td>
</tr>
<tr>
<td>30%</td>
<td>32.81% 0.985 0.325</td>
<td>20.13% 0.210</td>
<td>45.74% 0.437 0.278</td>
</tr>
<tr>
<td>(a)</td>
<td>12.93% 0.096 0.032</td>
<td>25.60% 0.066</td>
<td>54.26% 1.845 0.784</td>
</tr>
<tr>
<td>(c)</td>
<td>54.26% 1.689 0.743</td>
<td>54.26% 0.780</td>
<td>54.26% 1.845 0.784</td>
</tr>
<tr>
<td>50%</td>
<td>“(\mathbb{E}_w) (v(\cdot)) = 1.100”</td>
<td>“(\mathbb{E}_w) (v(\cdot)) = 1.006”</td>
<td>“(\mathbb{E}_w) (v(\cdot)) = 1.062”</td>
</tr>
</tbody>
</table>

While this example was created, it was not chosen specifically to achieve unusual results. While we tend to think of an SDF as a property of a market, it is actually a personal measure. When probability beliefs differ, each investor has his own SDF. The same is true here; For each RDU investor the subjective SDF is the state price divided by that investor’s decision weight, \(q\cdot w\), and the optimal portfolio’s return is decreasing in the subjective SDF. The state price is unchanged under probability weighting so states in which \(w_b > w_a\) have high returns compared to EUT portfolios. For an inverted S-shaped probability weighting, typically \(w_b > w_a\) in both tails of the distribution. This means the right tail of the returns distribution is longer and the left tail is shorter making the probability-weighted optimal portfolio’s distribution right skewed relative to the distribution of an EUT portfolio.

When \(w_b > w_a\) in the right tail, the increase in portfolio returns only leads to a skewing of the distribution. However, in the left tail the increased returns can potentially alter the outcome ordering which affects the probability weighting as in the example. Often the left tail will often

\[11\] It is possible to construct scenarios where the probability weight for either extreme outcome is less than the associated probability, and, therefore, the probability-weighted optimal portfolio has a smaller return than the expected-utility maximizing portfolio for the extreme outcome. For example, this occurs in the lower tail for the TK weighting function with \(\delta = 0.65\) if the worst state had a probability in excess of 35.87%. However, the best and worst states are typically very rare in most applications so right skewing of both tails of the optimal portfolio should be the typical result. The effect must always be present in at least one of the tails if the cumulative weighting function has an inverted S shape with only a single crossing of the 45° line.
be completely flattened so that the portfolio’s return is constant over a range of states that have different SDFs. The returns on the optimal portfolio in the example are only weakly decreasing in the SDF rather than strictly so as they would be under EUT.

At times it is possible to alter the ordering to the portfolio outcomes without two states having the same return. Suppose there are four states with probabilities \( \pi' = (0.2, 0.2, 0.4, 0.2) \) and state prices \( q' = (0.4, 0.3, 0.2, 0.1) \). The SDF is \( m' = (2, 1, 0.75, 0.5) \) so an EUT maximizer would hold a portfolio with \( x_a < x_b < x_c < x_d \). Consider a probability weighting function satisfying \( \Omega(0.2) = 0.3, \Omega(0.4) = 0.5, \Omega(0.6) = 0.56, \) and \( \Omega(0.8) = 0.75 \). These weights are consistent with an inverted S shape. If an RDU portfolio’s returns are ordered inversely to the SDF, then

\[
\omega_a = \Omega(0.2) = 0.3 \quad \omega_b = \Omega(0.4) - \Omega(0.2) = 0.2 \\
\omega_c = \Omega(0.6) - \Omega(0.4) = 0.25 \quad \omega_d = \Omega(1) - \Omega(0.8) = 0.25.
\]  

(57)

The \( \omega \)-SDF is \( m^\omega = q'/\omega \); its values are

\[
m^\omega_a = 1.33 \quad m^\omega_b = 1 \quad m^\omega_c = 1.2 \quad m^\omega_d = 0.4.
\]  

(58)

So the optimal (constrained ordering) portfolio would need to have \( x_a < x_c < x_b < x_d \), which is not in the order assumed. If the portfolio returns are ordered in this fashion, then

\[
\omega_a = \Omega(0.2) = 0.3 \quad \omega_c = \Omega(0.6) - \Omega(0.2) = 0.26 \\
\omega_b = \Omega(0.8) - \Omega(0.6) = 0.19 \quad \omega_d = \Omega(1) - \Omega(0.8) = 0.25.
\]  

(59)

The \( \omega \)-SDF values are

\[
m^\omega_a = 1.33 \quad m^\omega_c = 1.15 \quad m^\omega_b = 1.05 \quad m^\omega_d = 0.4.
\]  

(60)

The \( \omega \)-SDF ordering is now correct with the ordering of the returns on the optimal portfolio’s returns agreeing with the assumption. Therefore, a potential correct order for the optimal portfolio has been found though other permutations should be checked to assure that there is no other assumption that also gives a correct \( \omega \)-SDF ordering and has a higher expected utility.

This ordering does not, however, match that of the returns on an optimal portfolio of an EUT maximizer. The optimal portfolio returns are strictly decreasing in the state price per unit decision weight, but they are not even weakly decreasing in the state price per unit probability. For example, a log utility investor would hold the portfolio \( x' = (0.5, 1.0, 1.33, 2.0) \) using the probabilities, or the portfolio \( x' = (0.75, 0.95, 0.87, 2.51) \) using RDU decision weights.

The previous examples illustrate the complications of determining an optimal RDU portfolio. The ordering of returns on the optimal portfolio might not be possible ex ante meaning that many orderings will need to be checked. In any practical problem the difficulty is multiplied immensely as with \( n \) states, there are \( n! \) factorial standard portfolio problems to be solved and then compared — one for each of the possible orderings. A related complication is that the efficient set need not be convex when optimal portfolios have different orderings. If the efficient set is not convex, the market portfolio might not be efficient which means there would be no representative investor. These issues will be addressed in the next chapter.

**Probability Weighting and Portfolio Separation**

The most commonly analyzed market structure, other than a complete market, is one in which mutual-fund separation holds. Two-fund separation, in particular, is of interest because it yields strong predictions and sound intuition. Under two-fund money, the set of optimal portfolios is spanned by the risk-free asset and a single risky portfolio which must be the market...
portfolio (of risky assets). The efficient set is convex so there always is a representative investor who holds the market. Assets can therefore be priced in the usual fashion based on the first-order conditions of the representative investor holding the market portfolio. As discussed in Chapter 7, two-fund money separation holds when all investors’ utility functions are in the LRT class with the same cautiousness or when all asset distributions come from the separating distributions defined by Ross (1978).

Neither of these conditions is sufficient for two-fund separation with probability weighting. For example, two LRT-utility investors with the same cautiousness hold the risky assets in the same proportion under EUT when they have homogeneous beliefs. But probability weighting creates a form of heterogeneous beliefs so their risk-asset portfolios will only be the same if they have the same form of probability weighting as well even if they have homogeneous objective beliefs.

Probability weighting also may eliminate distributional-based separation. Ross’ (1978) necessary and sufficient condition for two-fund money is that returns are characterized by

\[ r_i = r_f + b_i y + \varepsilon_i \]

with \( \mathbb{E}[\varepsilon_i | y] = 0 \) \( \forall i \)

and \( \exists \alpha \) such that \( \mathbf{1}' \alpha = 1, \mathbf{a}' \varepsilon \equiv 0 \).

Because the identity function is (weakly) inverted S-shaped, this condition is clearly necessary for two-fund separation under RDU but it is not sufficient.

Suppose \( r_f = 1, y = \{ -1, 2 \} \) with equal probability, and an asset with \( b = 1 \) has \( \varepsilon = 0 \) when \( y = -1 \) and \( \varepsilon = \pm 1 \) with equal probability when \( y = 2 \). The residual risk \( \varepsilon \) is conditionally mean zero, as required, making the asset more risky than the portfolio \( \alpha \) with rate of return \( \tilde{y} \). Consider a risk-averse investor with a utility function \( u(w) = w \) for \( w > 0 \) and \( u(w) = 3w \) for \( w \leq 0 \) and a probability weighting function that assigns \( \Omega(0.5) = 0.5, \Omega(0.75) = 0.70 \). This investor will compute an expected payoff and utility (since all outcomes are positive) of 1.6 for the asset. The expected payoff and utility of the \( \alpha \) portfolio is 1.5. Levering this portfolio down will decrease its expected return and utility. Levering the portfolio up will increase its expected return, but also decrease its expected utility as the worst outcome is now in the high marginal utility region. So this investor’s optimal portfolio cannot be a levered position in portfolio \( \alpha \), and two-fund separation does not hold.

This example illustrates the difficulty of establishing two-fund separation with probability weighting. Two-fund separation requires showing that for any portfolio there exists a portfolio in a small class (those that have no idiosyncratic risk) that gives at least as high expected utility. For Ross’ distribution, any portfolio with a given \( b \) has the same expected return as and is more risky than portfolio \( \alpha \) levered to have the same \( b \). This specific levered portfolio stochastically dominates all assets with the same \( b \) value so we need not consider each utility function separately — two-fund separation holds trivially. However, with probability weighting, a levered position in portfolio \( \alpha \) no longer dominates all other portfolios with the same \( b \) because the convex portion of an inverted-S-shaped weighting function can increase the subjective mean after an objective-mean-preserving spread. Therefore, risk-neutral investors and those sufficiently close to risk-neutral will prefer the portfolio that is objectively dominated. This does not mean that two-fund separation fails, but it does mean that to verify separation it is no longer sufficient to compare portfolios with the same value of \( b \). We must potentially compare every portfolio with idiosyncratic risk to all levered index positions with the same or higher subjective mean.

Two related questions immediately arise. Which weighting functions do preserve two-fund separation for Ross’ distributions? And can the class of separating distributions be further restricted so that mutual fund separation does hold for inverse-S weighting functions? In fact, the first question is already mostly answered by the example. Any strictly convex portion of the
probability weighting function will increase the mean of some portfolios with residual risk removing the dominance of the levered $\alpha$ portfolio. Ross’ two-fund separation theorem can remain valid only for concave weighting functions. The next theorem shows that this is sufficient as well.

**Theorem 11.6: Two-fund Separation with Concave Probability Weighting.** If the returns on all assets are as described in (61), then optimal portfolios for all risk-averse investors are combinations of the risk-free asset and a single portfolio of risky assets if and only if all investors have concave probability weighting functions.

**Proof:** The necessity of the Ross distributions is obvious because EUT is a special case of RDU and they are necessary under EUT. The necessity of the concavity of the weighting function has already been discussed. As shown in Theorem 11.4, a concave weighting function is required for strong risk averse behavior and that is required for Ross’ separation result.

Sufficiency is demonstrated if the concavity of the weighting function preserves the second-order stochastic dominance inherent in the natural probability distribution of the Ross distribution.

As shown in Chapter 2, second-order stochastic dominance can be established with the quantiles, namely

$$
\mathbb{E}_p[u(x)] \geq \mathbb{E}_g[u(y)] \quad \forall u \text{ with } u' \geq 0, u'' \leq 0 \iff 0 \leq \int_0^p [F^{-1}(p) - G^{-1}(p)] dp \quad \forall P. \quad (62)
$$

So second-order stochastic dominance is preserved under weighting function $\Omega$ if

$$
\mathbb{E}_{\Omega(p)}[u(x)] \geq \mathbb{E}_{\Omega(y)}[u(y)] \iff 0 \leq \int_0^p [F^{-1}(\Omega^{-1}(p)) - G^{-1}(\Omega^{-1}(p))] dp \quad \forall P. \quad (63)
$$

Using the change in variable, $p = \Omega(q)$, the integral in (63) can be re-expressed as

$$
\int_0^{\Omega^{-1}(p)} [F^{-1}(q) - G^{-1}(q)] \Omega'(q) dq = H(q)\Omega'(q)\Bigg|_{0}^{\Omega^{-1}(p)} - \int_0^{\Omega^{-1}(p)} H(q)\Omega''(q) dq
$$

where $H(Q) = \int_0^Q [F^{-1}(q) - G^{-1}(q)] dq$.

The first term on the right-hand side of (64) is nonnegative because $H(0) = 0$ and $H$ and $\Omega'$ are nonnegative elsewhere. The remaining integral is nonpositive as $H$ is nonnegative and $\Omega''$ is nonpositive. Therefore the integral in (63) is nonnegative and stochastic dominance is preserved.

The intuition for this theorem is that decision weight density is the product of the natural probability density and the derivative of the cumulative weighting function, $\Omega'$. A concave weighting function has a decreasing derivative so it increases the decision weights of lower payoffs relative to those of higher payoffs; i.e., $\mathbb{E}_{\Omega}[\tilde{\epsilon} \mid y] < 0$. So any mean-preserving spread adds risk and reduces the expectation. This ensures that objectively dominated prospects remain dominated under probability weighting.

Two-fund separation is not preserved under inverse-S weighting functions because some increased risk may be liked if it induces an increased probability-weighted mean. To preserve two-fund separation, objective-mean preserving spreads that increase the probability weighted mean must be excluded. One way to accomplish this is to assume enough symmetry so that any mean-increasing alterations in one tail have offsetting mean-reducing alterations in the other tail. This requires symmetric distributions and “no better than symmetric” probability weighting adjustments.

As a first step, Theorem 11.7 shows that the Rothschild-Stiglitz notion of riskiness is preserved under symmetric probability weightings with $\Omega(1 - F) = 1 - \Omega(F)$.
Theorem 11.7: Symmetric Weighting Preserves Increasing Risk for Symmetric Random Variables. Consider two random variables \( \tilde{x} \) and \( \tilde{y} \) with the same mean and symmetric cumulative distributions \( F \) and \( G \) with \( \tilde{y} \) riskier than \( \tilde{x} \) in the sense of Rothschild and Stiglitz. Let \( \Omega(\cdot) \) be a symmetric inverted-S probability weighting function (i.e., \( \Omega(1 - F) = 1 - \Omega(F) \)) with \( \Omega \) increasing and concave below \( \frac{1}{2} \). Then \( \Omega(G(y)) \) is riskier than \( \Omega(F(x)) \).

**Proof:** The symmetric transformation \( \Omega \) preserves the symmetry of \( \tilde{x} \) and \( \tilde{y} \) so their expectations under probability weighting remain unchanged and equal. As in (63) and (64) above, \( \tilde{y} \) is riskier if and only if

\[
0 \leq \int_0^1 [F^{-1}(\Omega^{-1}(p)) - G^{-1}(\Omega^{-1}(p))] dp = \int_0^{1-\Omega^{-1}(P)} [F^{-1}(q) - G^{-1}(q)] \Omega'(q) dq.
\]  

(65)

For \( P \leq \frac{1}{2}, \Omega^{-1}(P) \leq \frac{1}{2} \) as well. In this region \( \Omega \) is concave so Theorem 11.6 applies and the integral is nonnegative. For \( P > \frac{1}{2} \)

\[
\int_0^{1-\Omega^{-1}(P)} [F^{-1}(q) - G^{-1}(q)] \Omega'(q) dq = \int_0^{1-\Omega^{-1}(P)} [F^{-1}(q) - G^{-1}(q)] \Omega'(q) dq + \int_{\frac{1}{2}}^{1-\Omega^{-1}(P)} [F^{-1}(q) - G^{-1}(q)] \Omega'(q) dq.
\]

(66)

By symmetry the last two integrals are equal in magnitude and opposite in sign so their sum is zero. The first integral is nonnegative as in (65) because \( 1 - \Omega^{-1}(P) \leq \frac{1}{2} \) when \( P \geq \frac{1}{2} \).

Theorem 11.8: Two-Fund Separation under Inverse-S Probability Weighting.

Sufficient conditions for two-fund separation under probability weighting are: (i) returns satisfy the Ross conditions in (61); (ii) the distributions of all rates of returns including \( \tilde{y} \) are symmetric; and (iii) the cumulative probability weighting function for the distribution \( F \) has the form \( \Omega(F) \equiv \Xi(\Xi(F)) \). \( \Xi(\cdot) \) is increasing and concave. \( \Xi(\cdot) \) is increasing and concave below \( F = \frac{1}{2} \) and complementary around \( \frac{1}{2} \) with \( \Xi(1 - F) = 1 - \Xi(F) \).

**Proof:** Let \( F \) and \( G \) be the cumulative distributions of \( r_f + \tilde{y} \) and \( r_f + b_i \tilde{y} + \tilde{e}_i \). Because \( F \) and \( G \) are the distributions of symmetric random variables, the transformations \( \Xi(F) \) and \( \Xi(G) \) preserve the riskiness ordering as shown in Theorem 11.7. Therefore, \( \Xi(G) \) remains riskier than \( \Xi(F) \) in a Rothschild-Stiglitz sense. Applying Theorem 11.6, the increasing concave transformation \( \tilde{\gamma} \) preserves the second-order stochastic dominance.

Unfortunately, the assumption in this theorem does not apply to the KT weighting function for any parameter value nor to many of the other probability weighting functions used. The second derivative of \( \Omega(F) \equiv \Xi(\Xi(F)) \) is \( \Omega''(F) = \Xi''(F) + \Xi'(F) \). \( \Xi''(\cdot) \) is assumed to be negative so the first term is negative; therefore, the inflection point in \( \Omega(F) \), if there is one, must occur when \( \Xi'' \) is positive. By assumption, this must be for \( F \geq \frac{1}{2} \). However, the inflection point for the KT probability weighting function occurs at a probability less than \( \frac{1}{2} \) for all values of \( \delta \). In fact, using a non-parametric approach to determine the “least favored” probability, Wu and Gonzalez (1966) have estimated that the inflection point is no higher than 40%. If this is correct, then the probability weighting function cannot be represented as required in this theorem.

---

12 Assumption (ii) does not require that the distributions of \( \tilde{e} \) be symmetric. If the distributions of \( \varepsilon \) conditional on \( \tilde{y} = \tilde{y} + a \) and that of \( -\tilde{e} \) conditional on \( \tilde{y} = \tilde{y} - a \) are identical, then the asset returns will be symmetric.

13 As noted previously, assuming \( \Xi(1 - F) = 1 - \Xi(F) \) is equivalent to using identical weighting functions for losses and gains, \( \Omega'(1 - F) = \Omega'(F) \) and requiring \( \Omega'(\frac{1}{2}) = \frac{1}{2} \). Of course, after the \( \tilde{\gamma} \) transformation, \( \Omega^+ \) and \( \Omega^- \) will no longer be identical.