1. There are three dimensions of an investor that matter in asset trading: the investor’s initial endowments, her preferences, and her beliefs (or subjective probabilities) about the uncertain future. In a two-period economy where an investor trades assets at date 0 and receives returns at date 1, the investor’s endowments are usually represented by her date-0 initial wealth $W_0 > 0$, and her preferences represented by a von Neumann-Morgenstern utility function $u(\cdot)$ for her date-1 random wealth $\tilde{W}$. It has been a standard assumption in finance theory that an investor in making her date-0 investment decisions seeks to maximize $E[u(\tilde{W})]$, where the expectation $E[\cdot]$ is taken using the investor’s subjective probability distribution for $\tilde{W}$. In this elementary course, we shall only deal with the case where at date 0 all investors agree that there is an objective probability law according to which the date-1 uncertainty will be resolved. The EU theory developed under the assumption that there exists an objective probability law for the future uncertain events will be referred to as the objective EU theory.

2. The first part of this note defines rational investor, and poses three behavior axioms that imply that in a two-period investment environment with objective probabilities, rational investors are EU-maximizers. The second part of this note then defines fair gamble, risk aversion, dollar return, rate of return, excess return, and risk premium, and examines a risk-averse investor’s portfolio choice problem when there are only one risky asset and one riskless asset available for trading. The third part of this note compares risk aversion of different investors, and defines the notion of one investor being more risk averse than another investor. The fourth part of this note compares riskiness of investment projects,
and defines the notion of one investment project being riskier than another project. The fifth part of this note consists of applications of the preceding results to asset pricing, efficient risk-sharing, capital structure and the pricing of insurance policy. Finally, in the last part of this note, we briefly go over the basic ideas of prospect theory.

1 Three Axioms and the EU Theory

1. A decision maker facing a set $A$ of feasible alternatives is said to be rational if she is endowed with a weak preference relation $\succeq$ on $A$.\footnote{A binary relation $\succeq$ on $A$ is complete if for all $a, b \in A$, either $a \succeq b$ or $b \succeq a$ (or both), and it is transitive if for all $a, b, c \in A$, $a \succeq c$ whenever $a \succeq b$ and $b \succeq c$. A binary relation $\succeq$ on $A$ is a weak preference on $A$ if it is both complete and transitive. From a weak preference $\succeq$ on $A$, we can derive the strict preference $\succ$ on $A$ as follows. For all $a, b \in A$, $a \succ b$ if it is not true that $b \succeq a$. Similarly, we can derive the indifference relation $\sim$ on $A$ from the weak preference $\succeq$ on $A$: for all $a, b \in A$, we have $a \sim b$ if both $a \succeq b$ and $b \succeq a$.} We say that a (real-valued) utility function $u : A \rightarrow \mathbb{R}$ represents $\succeq$ if and only if for all $a, b \in A$,

$$u(a) > u(b) \iff a \succ b.$$ 

For example, suppose that $A = \{a, b, c\}$ and $a \succ b \sim c$. Among the following three real-valued functions, $u(\cdot)$ and $v(\cdot)$ both represent $\succeq$, but $w(\cdot)$ does not:

$$u(a) = 3, \quad u(b) = 0, \quad u(c) = 0;$$

$$v(a) = 5, \quad v(b) = 3, \quad v(c) = 3;$$

$$w(a) = 1, \quad w(b) = 1, \quad w(c) = 0.$$ 

2. Consider a date-0 investment environment where all investment projects (or lotteries) generate cash flows only at date 1, and there are $N$ possible levels of date-1 cash flows, denoted by $z_1 < z_2 < \cdots < z_N$. Define $Z \equiv \{z_1 < z_2 < \cdots < z_N\}$. In this case, an investment project (or a lottery) $p$ is a probability distribution over the set $Z$, which generates consumption level $z \in Z$ with probability $p(z)$. Let $P$ be the set of all feasible lotteries; that is, $P$ contains all probability distributions over
the $\mathbf{Z}$. We call $\mathbf{P}$ the *investment opportunity set*. Consistent with our definition in the previous section, here an investor facing $\mathbf{P}$ is called *rational* if she is endowed with a weak preference $\succeq$ on $\mathbf{P}$.

3. Our first main result is that, a rational investor’s weak preference on $\mathbf{P}$ satisfies the following three behavioral assumptions if and only if she is an EU-maximizer.

**Axiom 1 (Axiom of Reduction)** For all $a \in [0, 1]$ and for all $p, r \in \mathbf{P}$, the investor feels indifferent about the *simple lottery* $ap + (1 - a)r$ (which yields the consumption level $z \in \mathbf{Z}$ with probability $ap(z) + (1 - a)r(z)$) and the *compound lottery* that with probability $a$ he gets to take the lottery $p$ and with probability $1 - a$ he gets to take the lottery $r$.\(^2\)

**Axiom 2 (Independence Axiom)** For all $p, q, r \in \mathbf{P}$ and $a \in (0, 1)$, $p \succ q \Rightarrow ap + (1 - a)r \succ aq + (1 - a)r$.

**Axiom 3 (Continuity Axiom)** For all $p, q, r \in \mathbf{P}$, $p \succ q \succ r \Rightarrow ap + (1 - a)r \succ q \succ bp + (1 - b)r$ for some $a, b \in (0, 1)$.

\(^2\)While a simple lottery is a probability distribution on $\mathbf{Z}$, a (two-stage) compound lottery is a probability distribution on $\mathbf{P}$. In other words, the outcome of a simple lottery is a cash flow $z \in \mathbf{Z}$, but the outcome of a compound lottery may be an opportunity to play another lottery $p \in \mathbf{P}$. Note also that the above assumed date-0 investment environment has ruled out investment projects involving cash outflows at date 0. Moreover, if the lottery $p$ represents one unit of a traded security, holding 0.5 units of that security may not be a feasible investment project at date 0. To see this, suppose that $p$ generates $z_N$ with probability one. Then 0.5 units of $p$ will generate 0.5$z_N$ for sure, but it may happen that $z_j \neq 0.5z_N$ for all $j = 1, 2, \cdots, N$.

\(^3\)Suppose that $N = 2$ and $z_j = j$, for all $j = 1, 2$. Suppose that $p, q, r \in \mathbf{P}$ are such that

$$p = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}, \quad q = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \quad r = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}.$$  

Axiom 1 says that a compound lottery that allows the investor to take $p$ and $q$ both with probability $\frac{1}{2}$ is regarded as equivalent to $r$ by the investor.
Here comes our first main result.

**Theorem 1** A weak preference $\succeq$ on $P$ has an expected utility function representation if and only if it satisfies Axioms 1-3. That is, there exists a utility function $u : Z \to \mathbb{R}$ such that

$$
\forall p, q \in P, \ p \succ q \iff H(p) \equiv \sum_{z \in Z} p(z)u(z) \equiv E_p[u] > E_q[u] \equiv \sum_{z \in Z} q(z)u(z) \equiv H(q),
$$

if and only if Axioms 1,2, and 3 hold.

Moreover, if there exist $p,q \in P$ such that $p \succ q$, then the utility function $u$ is unique up to a positive affine transformation in the sense that if $v : Z \to \mathbb{R}$ is such that

$$
\forall p, q \in P, \ \sum_{z \in Z} p(z)v(z) > \sum_{z \in Z} q(z)v(z) \iff p \succ q,
$$

then there exist some $a \in \mathbb{R}$ and $b \in \mathbb{R}_{++}$ such that for all $z \in Z$,

$$
u(z) = a + bv(z).
$$

4. To prove Theorem 1 we first establish 6 Lemmas from the above Axioms.

**Lemma 1** $p \succ q$ and $0 \leq a < b \leq 1$ imply that $bp + (1 - b)q \succ ap + (1 - a)q$.

**Proof.** The case $b = 1$ is easy; the assertion can be proved by simply replacing $r$ by $p$ in Axiom 2. Thus assume that $b < 1$. Let $w = \frac{b-a}{1-a} \in (0,1)$, Axiom 2 implies that $r \equiv wp + (1 - w)q \succ q$. Now we have

$$
bp + (1 - b)q = ap + (1 - a)r \succ ap + (1 - a)q,
$$

by Axiom 1 and then Axiom 2.

**Lemma 2** $p \succeq q \succeq r$ and $p \succ r$ imply the unique existence of $a \in [0,1]$ such that $q \sim ap + (1 - a)r$. 

4
Proof. If \( p \sim q \) then \( a = 1 \) will do, and by Lemma 1, for all \( b \in [0, 1) \), \( q \succ bp + (1 - b)r \), showing the uniqueness. The case of \( p \succ q \sim r \) is similar: \( a = 0 \) will do, and by Lemma 1, for all \( b \in (0, 1] \), we have \( bp + (1 - b)r \succ 0 \cdot p + (1 - 0) \cdot r = r \sim q \), showing the uniqueness. It remains to consider the case of \( p \succ q \succ r \). Define

\[
U = \{a \in [0, 1] : ap + (1 - a)r \succ q\}, \\
I = \{a \in [0, 1] : ap + (1 - a)r \sim q\}, \\
L = \{a \in [0, 1] : q \succ ap + (1 - a)r\}.
\]

By transitivity of \( \succeq \), \( U, I, \) and \( L \) form a partition of \([0, 1] \). Apparently \( 1 \in U \) and \( 0 \in L \). By Axiom 3, \( U \) and \( L \) respectively have elements different from 0 and 1. Moreover, \( U \) and \( L \) both have the continuum property: If \( a_1 < a_2 \) and \( a_1, a_2 \in U \), then by Lemma 1 \( a \in U \) for all \( a \in (a_1, a_2) \). The same is true for \( L \). It follows that \( U \) and \( L \) are both intervals. Observe also that by Lemma 1 \( I \) is either empty or a singleton. Suppose that \( I \) were empty and we would have a contradiction. Since \([0, 1]\) is connected and \( U \) and \( L \) form a partition for \([0, 1]\), one between \( U \) and \( L \) must be closed: either for some \( a, L = [0, a] \) and \( U = (a, 1] \) or \( L = [0, a) \) and \( U = [a, 1] \). Suppose that the former were true. Then \( u \in U \) if and only if \( u > a \), and we have

\[
up + (1 - u)r \succ q \succ ap + (1 - a)r,
\]

by the fact that \( u \in U \) and \( a \in L \). Axiom 3 then implies that there exists \( b \in (0, 1) \) such that

\[
q \succ b[up + (1 - u)r] + (1 - b)[ap + (1 - a)r] \in L,
\]

but since \( bu + (1 - b)a > a \), we must have

\[
b[up + (1 - u)r] + (1 - b)[ap + (1 - a)r] \in U,
\]

a contradiction! The case where \( L = [0, a) \) and \( U = [a, 1] \) can be ruled out analogously. It follows that \( I \) is nonempty, and hence a singleton. \( \parallel \)

Lemma 3 \( p \succ q \) and \( r \succ s \) and \( a \in [0, 1] \) imply that \( ap + (1 - a)r \succ aq + (1 - a)s \).

5
Proof. Repeatedly applying Axiom 2, we have
\[ ap + (1 - a)r \succ aq + (1 - a)r \succ aq + (1 - a)s. \]

Lemma 4  \( p \sim q \) and \( a \in [0, 1] \) imply that \( p \sim ap + (1 - a)q \).

Proof. Suppose instead that \( p \succ ap + (1 - a)q \). Then \( q \succ ap + (1 - a)q \) also. By Lemma 3, we have
\[ ap + (1 - a)q \succ a[ap + (1 - a)q] + (1 - a)[ap + (1 - a)q] = ap + (1 - a)q, \]
a contradiction. The case where \( ap + (1 - a)q \succ p \) can be ruled out analogously.

Lemma 5  \( p \sim q \) and \( a \in [0, 1] \) imply that \( ap + (1 - a)r \sim aq + (1 - a)r \) for all \( r \in P \).

Proof. The cases \( a = 0 \) and \( a = 1 \) are obvious, and so we assume that \( 0 < a < 1 \). Fix \( p \sim q \). For \( r \in P \) such that \( p \sim q \sim r \), the assertion follows easily from Lemma 4. Thus we consider the case \( p \sim q \succ r \) (the remaining case is similar). Suppose that \( ap + (1 - a)r \succ aq + (1 - a)r \), and we shall demonstrate a contradiction. By Lemma 1, we have
\[ ap + (1 - a)r \succ aq + (1 - a)r \succ r. \]
By Axiom 3, there must exist \( b \in (0, 1) \) such that
\[ b[ap + (1 - a)r] + (1 - b)[r] \succ aq + (1 - a)r, \]
or equivalently,
\[ abp + (1 - ab)r \succ aq + (1 - a)r. \]
However, by Lemma 1, we also have
\[ q \sim p \succ bp + (1 - b)r, \]
so that by Axiom 2, we have
\[ aq + (1 - a)r \succ a[bp + (1 - b)r] + (1 - a)r = abp + (1 - ab)r, \]
a contradiction.
Lemma 6  There exist $z^*, z_* \in \mathbb{Z}$ such that $P_{z^*} \succeq p \succeq P_{z_*}$ for all $p \in P$, where $P_z$ denotes the lottery that generates consumption $z$ with probability one.

Proof. Since $\mathbb{Z}$ is finite, we must have by Axiom 1

$$P_{z^*} \succeq P_z \succeq P_{z_*}, \forall z \in \mathbb{Z}.$$  

(Step 1): Let $P_n \subset P$ be the set of lotteries in $P$ assigning strictly positive probabilities to exactly $n$ elements of $\mathbb{Z}$. Then it is clear that $(P_1, P_2, \ldots, P_N)$ form a partition of $P$. We first claim that for all $m \in \{2, 3, \ldots, N\}$ and for all $p \in P_m$, there exist $q, r \in P_{m-1}$ and $a \in (0, 1)$ such that $p = aq + (1-a)r$, where $aq + (1-a)r$ is the compound lottery yielding lotteries $q$ and $r$ with respectively probabilities $a$ and $1-a$. The claim is obviously true for $m = 2$, and for $m \geq 3$, assume without loss of generality that $p$ assigns positive probabilities only to $z_1, z_2, \ldots, z_m$. Suppose that $p$, $q$ and $r$ are such that

$$p(z_i) > 0, \forall i = 1, 2, \ldots, m; \quad p(z_{m+1}) = p(z_{m+2}) = \cdots = p(z_N) = 0,$$

$$p(z_j) = q(z_j) = r(z_j) \geq 0, \forall j = 4, 5, \ldots, N,$$

and

$$q(z_1) = r(z_3) = 0.$$

We need to show that the following system of equations has a solution:

$$q(z_2) + q(z_3) = r(z_1) + r(z_2) = p(z_1) + p(z_2) + p(z_3),$$

$$a \begin{bmatrix} 0 & q(z_2) & q(z_3) \\ q(z_2) & r(z_2) & 0 \\ q(z_3) & 0 & 0 \end{bmatrix} + (1-a) \begin{bmatrix} r(z_1) \\ r(z_2) \\ 0 \end{bmatrix} = \begin{bmatrix} p(z_1) \\ p(z_2) \\ p(z_3) \end{bmatrix},$$

with the five unknowns satisfying

$$0 < a, q(z_2), q(z_3), r(z_1), r(z_2) < 1.$$

It is easy to verify that any $a, q(z_2), q(z_3), r(z_1), r(z_2) \in (0, 1)$ such that

$$\frac{p(z_2) + p(z_3)}{p(z_1) + p(z_2) + p(z_3)} > a > \frac{p(z_3)}{p(z_1) + p(z_2) + p(z_3)}$$
can serve the purpose.

(Step 2): Now we show that there exist sure lotteries $P_{z^*}$ and $P_{z^*}$ such that for all $q \in P$, $P_{z^*} \succeq q \succeq P_{z^*}$. Define $Q_m = \bigcup_{j=1}^{m} P_j$. That there exist sure lotteries $P_{z^*}$ and $P_{z^*}$ such that for all $q \in Q_1$, $P_{z^*} \succeq q \succeq P_{z^*}$ follows from the fact that $\succeq$ satisfies completeness and $Z$ is finite. For $m \geq 2$, suppose that it has already been proved that there exist sure lotteries $P_{z^*}$ and $P_{z^*}$ such that for all $q \in Q_{m-1}$, $P_{z^*} \succeq q \succeq P_{z^*}$. By Step 1 above, every $p \in Q_m$ can be represented as a compound lottery yielding lotteries $q$ and $r$ with respectively probabilities $a$ and $1 - a$, where $q, r \in Q_{m-1}$. Since there exist $P_{z^*}$ and $P_{z^*}$ such that $P_{z^*} \succeq q \succeq P_{z^*}$ and $P_{z^*} \succeq r \succeq P_{z^*}$, by Lemmas 1 and 4 we must also have

$$P_{z^*} \succeq aq + (1 - a)r \sim p \sim aq + (1 - a)r \succeq P_{z^*}.$$

Thus by mathematical induction the proof is complete. ||

5. Now we are ready to prove Theorem 1.

**Proof of Theorem 1.** We first prove sufficiency by distinguishing two cases.

**Case 1:** $P_{z^*} \sim P_{z^*}$.

By Lemma 6, the investor must feel indifferent about all $p, q \in P$ in this case. Let us pick any $k \in \mathbb{R}$ and define $u(z) = k$ for all $z \in Z$. We have immediately $E_p[u(z)] = \sum_{z \in Z} p(z)u(z) = k$ for all $p \in P$, so that the constant function $u(\cdot)$ does represent $\succeq$ on $P$. Note that any (positive or negative) affine transform of $u(\cdot)$ results in another constant function on $Z$, which also represents $\succeq$. Apparently, a non-constant function on $Z$ cannot represent $\succeq$.

**Case 2:** $P_{z^*} \succ P_{z^*}$.
Define $H(p)$ as such that

$$H(p)P_{z^*} + (1 - H(p))P_{z_*} \sim p,$$

where by Lemma 2, $H(p)$ is uniquely defined for all $p \in P$. In fact, Lemma 1 and Lemma 4 show that $H(p)$ is a utility function representing $\succeq$: $H(p) > H(q)$ if and only if $p > q$ for all $p, q \in P$.\(^4\) By Lemma 5 and the definition of $H(p)$, we have

$$ap + (1 - a)q$$

$$\sim a[H(p)P_{z^*} + (1 - H(p))P_{z_*}] + (1 - a)[H(q)P_{z^*} + (1 - H(q))P_{z_*}]$$

$$\sim [aH(p) + (1 - a)H(q)]P_{z^*} + [1 - aH(p) + (1 - a)H(q)]P_{z_*},$$

so that

$$H(ap + (1 - a)q) = aH(p) + (1 - a)H(q),$$

showing that $H(\cdot)$ is linear. Now define $u(z) = H(P_z)$. We have

$$H(p) = H\left(\sum_{z \in Z} p(z)P_z\right) = \sum_{z \in Z} p(z)H(P_z) = \sum_{z \in Z} p(z)u(z).$$

Thus $\succeq$ does have an expected utility function representation. This finishes our proof for the sufficiency.

The necessity is much easier. Suppose that $\succeq$ has an expected utility function $H(\cdot)$ representation, where $H$ is real-valued. Since $H(\cdot)$ represents $\succeq$ on $P$, $\succeq$ must be a weak preference on $P$.\(^5\) We must show that Axioms 1, 2, and 3 are all satisfied by $\succeq$. Axiom 1 follows exactly

\(^4\)Here note that $H(\cdot)$ is an ordinal utility function in the sense that any monotonic transform of $H(\cdot)$ also represents $\succeq$.

\(^5\)More precisely, note that for all $p, q \in P$, $H(p)$ and $H(q)$ are just two real numbers, and so we have either $H(p) \geq H(q)$ or $H(q) \geq H(p)$. Since $H(\cdot)$ represents $\succeq$ in the sense that $p \succeq q$ if and only if $H(p) \geq H(q)$, we conclude that for all $p, q \in P$, we have either $p \succeq q$ or $q \succeq p$. This proves that $\succeq$ is a complete binary relation on $P$.

Similarly, note that for all $p, q, r \in P$, $H(p) \geq H(r)$ whenever $H(p) \geq H(q)$ and $H(q) \geq H(r)$. This says that, again by the fact that $H(\cdot)$ represents $\succeq$, for all $p, q, r \in P$, we have $p \succeq r$ whenever $p \succeq q$ and $q \succeq r$. This proves that $\succeq$ is a transitive binary relation on $P$. At this point, we have confirmed that $\succeq$ must be a weak preference on $P$ if $\succeq$ can be represented by a real-valued expected utility function $H(\cdot)$.
because $H(\cdot)$ is linear.\textsuperscript{6} The linearity of $H(p)$ also implies that Axiom 2 holds.\textsuperscript{7} Finally, the continuity of $H(\cdot)$ on $\mathcal{R}$ proves that Axiom 3 must also hold.\textsuperscript{8} 9

Finally, we show that that $u(\cdot)$ is determined up to a positive affine transform. Suppose that $G(p) = \sum_{z \in \mathcal{Z}} p(z)u(z)$ is another expected utility function representing $\succeq$. By the fact that $p \sim H(p)P_{z^*} + [1 - H(p)]P_{z_0}$, we have

$$G(p) = G(H(p)P_{z^*} + [1 - H(p)]P_{z_0}) = H(p)G(P_{z^*}) + [1 - H(p)]G(P_{z_0})$$

$$= v(z_*) + [v(z^*) - v(z_*)]H(p),$$

and hence $G(p)$ must be an affine transform of $H(p)$. To see that this transform is positive, note that $P_{z^*} \succ P_{z_0}$ and hence $v(z^*) - v(z_*) > 0$. $\|$ 

\textsuperscript{6}More precisely, for any $a \in [0, 1]$ and for any $p, q \in \mathcal{P}$, the expected utility generated by the simple lottery $ap + (1 - a)q$ is $H(ap + (1 - a)q)$, while the expected utility generated by the compound lottery that with probability $a$ the investor gets to take lottery $p$ and with probability $1 - a$ she gets to take lottery $q$ is $aH(p) + (1 - a)H(q)$. Note that

$$H(ap + (1 - a)q) = \sum_{z \in \mathcal{Z}} [ap(z) + (1 - a)q(z)]u(z) = a\sum_{z \in \mathcal{Z}} p(z)u(z) + (1 - a)\sum_{z \in \mathcal{Z}} q(z)u(z) = aH(p) + (1 - a)H(q),$$

and hence the two expected utilities are equal! Since $H(\cdot)$ represents $\succeq$, the investor must feel indifferent about the simple lottery $ap + (1 - a)q$ and the above compound lottery. This proves that Axiom of Reduction holds.

\textsuperscript{7}Pick any $p, q, r \in \mathcal{P}$ such that $p \succ q$, and pick any $a \in (0, 1]$. Since $H(\cdot)$ represents $\succeq$, we must have $H(p) > H(q)$. Now

$$H(ap + (1 - a)r) = aH(p) + (1 - a)H(r) > aH(q) + (1 - a)H(r) = H(aq + (1 - a)r),$$

and since $H(\cdot)$ represents $\succeq$, we must have $ap + (1 - a)r \succ eq + (1 - a)r$. This proves that the Independence Axiom holds.

\textsuperscript{8}Pick any $p, q, r \in \mathcal{P}$ with $H(p) > H(q) > H(r)$. Since $h(z) = zH(p) + (1 - z)H(r)$ is a continuous function on $\mathcal{R}$, and since $h(1) > H(q) > h(0)$, there must exist $a, b \in (0, 1)$ such that $a$ is sufficiently close to 1 and $b$ is sufficiently close to 0, with $H(aq + (1 - a)r) = h(a) > H(q) > h(b) = H(bq + (1 - b)r)$. By the fact that $H(\cdot)$ represents $\succeq$, this shows that $aq + (1 - a)r > q > bq + (1 - b)r$; that is, the Continuity Axiom holds.

\textsuperscript{9}A weak preference represented by an expected utility function $H(p)$ may violate Axiom 3 if $H$ can take $+\infty$ or $-\infty$ as its functional value. For example, suppose that $H(p) = \sum_{z \in \mathcal{Z}} p(z)\log(z)$, where $0, 1, 2 \in \mathcal{Z}$. It is clear that $P_2 \succ P_1 \succ P_0$, but for all $a \in (0, 1)$, we have $P_1 \succ ap_2 + (1 - a)P_0$. 

10
6. The above $u(\cdot)$ is referred to as a von Neumann-Morgenstern (VNM) utility function (cf. von Neumann and Morgenstern (1953)), which is defined on $\mathbb{Z}$. Note that $u$ is a cardinal utility function. We usually take $\mathbb{Z}$ to be $\mathbb{R}_+$, representing the set of (non-negative) consumption or wealth levels.\(^\text{10}\) In a dynamic setting where an investor faces a stream of lotteries, we usually assume that the investor maximizes a discounted sum of temporal VNM utility functions. (We say that the investor’s utility function is time-additive or time-separable.) In continuous-time models, for example, we usually let $T = [0, T]$ denote the time span, and the investor’s objective function is represented as $E_0[\int_{t \in T} e^{-\rho t} u(\hat{c}_t) dt]$, where $E_0[\cdot]$ is the expectation operator conditional on the time-0 information, $\hat{c}_t$ is the investor’s random time-$t$ consumption, and $\rho$ and $e^{-\rho t}$ are referred to as respectively the discount rate and the discount factor. In discrete-time models, we usually let $T = \{0, 1, 2, \ldots, t, t + 1, \ldots, T\}$, and we assume that the investor seeks to maximize $E_0[\sum_{t = 0}^{T} \delta^t u(\hat{c}_t)]$, where $\delta \in (0, 1)$ is the discrete-time discount factor.\(^\text{11}\)

2 The Two-Asset Portfolio Problem

1. In the remainder of this note, investors are assumed to maximize expected utility when making investment decisions. An investor is represented by a pair $(W_0, u)$, where $W_0$ is the investor’s (non-random) initial wealth, which consists of cash, and $u(\cdot)$ is the investor’s VNM utility function for terminal wealth. We shall consider the investment problem facing an investor endowed with VNM utility function $u : \mathbb{R}_+ \rightarrow \mathbb{R}$. A fair gamble is a lottery that generates zero expected profits. An investor with VNM utility function $u$ is risk neutral (respectively, risk averse and risk seeking) if given any initial wealth $W_0$ and facing any fair gamble $\tilde{z}$, she would feel indifferent about (respectively, become

\(^{10}\) However, for the case where $\mathbb{Z} = \mathbb{R}_+$, the three axioms specified above are no longer sufficient for the preference to be representable by an expected utility function; a fourth axiom called “the sure thing principle” is now needed. See Kreps (1988).

\(^{11}\) This time-additive utility function gives rise to the so-called equity premium puzzle in finance literature, which we shall explain in a subsequent lecture.
worse off, and become better off) taking the fair gamble $\tilde{z}$:

$$E[u(W_0 + \tilde{z})] = (\text{respectively, } \leq \text{ and } \geq) u(W_0).$$

Lemma 7 (Jensen Inequality) If $f : \mathbb{R} \to \mathbb{R}$ is concave (respectively, convex), and $\tilde{x}$ is a random variable with finite $E[\tilde{x}]$, then

$$E[f(\tilde{x})] \leq (\text{respectively, } \geq) f(E[\tilde{x}]).$$

Proof. We shall prove the lemma only for the convex case. We proceed in a few steps.

Step (i). Fix any $b, c \in \mathbb{R}$. Then the function $g(z) \equiv \frac{f(z) - f(c)}{z - c}$ is non-decreasing on $(c, +\infty)$, and the function $h(z) \equiv \frac{f(z) - f(b)}{z - b}$ is non-decreasing on $(-\infty, b)$.

To see this, suppose that $c < x < b$, $b, c, x \in \mathbb{R}$. Note that by definition of convexity, we have

$$f\left(\frac{b - x}{b - c} \cdot c + \frac{x - c}{b - c} \cdot b\right) = f(x) \leq \frac{b - x}{b - c} f(c) + \frac{x - c}{b - c} f(b),$$

where the inequality follows from the definition of convex function. Hence we have

$$(b - c)[f(x) - f(c)] \leq (x - c)[f(b) - f(c)] \Rightarrow \frac{f(x) - f(c)}{x - c} \leq \frac{f(b) - f(c)}{b - c},$$

and

$$(b - c)[f(b) - f(x)] \geq (b - x)[f(b) - f(c)] \Rightarrow \frac{f(b) - f(x)}{b - x} \geq \frac{f(b) - f(c)}{b - c}.$$}

We have shown that for all $c < x < b$, if $f(.)$ is convex, then

$$\frac{f(b) - f(x)}{b - x} \geq \frac{f(b) - f(c)}{b - c} \geq \frac{f(x) - f(c)}{x - c}.$$

Readers interested in other properties of concave functions can look up Tiel (1984).
This implies that, with \( c \) being given, \( g(b) \geq g(x) \) whenever \( b > x > c \); and with \( b \) being given, \( h(x) \geq h(c) \) whenever \( b > x > c \). Thus \( g : (c, +\infty) \to \mathbb{R} \) and \( h : (-\infty, b) \to \mathbb{R} \) are both weakly increasing (or, non-decreasing).

**Step (ii).** The right-hand derivative \( f'_+(c) \equiv \lim_{z \downarrow c} g(z) \) and the left-hand derivative \( f'_-(b) \equiv \lim_{z \uparrow b} h(z) \) exist for all \( b, c \in \mathbb{R} \).

This follows from the axiom below.

**Axiom of Continuity.**

An infinite sequence of real numbers \( \{a_n; n \in \mathbb{Z}_+\} \) has a limit if either it is increasing and has an upper bound (that is, if there exists a real number \( u \) such that for all \( n \in \mathbb{Z}_+ \), \( a_n \leq a_{n+1} \leq u \)) or it is decreasing and has a lower bound (that is, if there exists a real number \( l \) such that for all \( n \in \mathbb{Z}_+ \), \( a_n \geq a_{n+1} \geq l \)).

Let \( \{a_n\} \) be any decreasing sequence converging to \( c \), and \( \{b_n\} \) be any increasing sequence converging to \( b \), so that \( \{g(a_n)\} \) and \( \{h(b_n)\} \) are respectively decreasing and increasing sequences. Note that \( \{g(a_n)\} \) and \( \{h(b_n)\} \) have respectively a lower bound and an upper bound: given any \( d < c \) and \( e > b \), we have, for all positive integers \( n \),

\[
  g(a_n) \geq \frac{f(c) - f(d)}{c - d}, \quad h(b_n) \leq \frac{f(e) - f(b)}{e - b}.
\]

Now by the axiom of continuity, the two infinite sequences \( \{g(a_n)\} \) and \( \{h(b_n)\} \) both have limits, which are exactly \( f'_+(c) \) and \( f'_-(b) \). (Why are these limits independent of the choices of \( \{a_n\} \) and \( \{b_n\}\)?)

**Step (iii).** For all \( c, d \in \mathbb{R} \) with \( c < d \),

\[
  f'_-(c) \leq f'_+(c) \leq f'_-(d).
\]
To see this, note that for any $a < c < d$ and sufficiently small $h > 0$, we have

\[
\frac{f(c) - f(c - h)}{h} \leq \frac{f(c + h) - f(c)}{h} \leq \frac{f(d) - f(d - h)}{h},
\]

which implies $f'_-(c) \leq f'_+(c) \leq f'_-(d)$ once we let $h \downarrow 0$.

**Step (iv).** For all $x, c \in \mathbb{R}$,

\[
f(x) \geq f(c) + f'_-(c)(x - c).
\]

To see this, note that if $x > c$, we have

\[
\frac{f(x) - f(c)}{x - c} \geq f'_+(c) \geq f'_-(c) \Rightarrow f(x) \geq f(c) + f'_+(c)(x - c) \geq f(c) + f'_-(c)(x - c),
\]

and if $x < c$, we have

\[
\frac{f(x) - f(c)}{x - c} \leq f'_-(c) \leq f'_+(c) \Rightarrow f(x) \geq f(c) + f'_-(c)(x - c) \geq f(c) + f'_+(c)(x - c).
\]

**Step (v).** Jensen inequality holds.

To see this, note that we have shown in the previous step that for all $x, c \in \mathbb{R}$,

\[
f(x) \geq f(c) + f'_-(c)(x - c).
\]

Replacing $c = E[\tilde{x}] \in \mathbb{R}$ into the last inequality and taking expectations on both sides of the inequality, we obtain

\[
E[f(\tilde{x})] \geq f(E[\tilde{x}]) + f'_-(E[\tilde{x}])E[\tilde{x} - E[\tilde{x}]] = f(E[\tilde{x}]).
\]

The proof is complete.\

14
Theorem 2  The investor with VNM utility function $u(\cdot)$ is risk averse if and only if $u(\cdot)$ is concave.

Proof. Consider necessity. For any $x, y \in \mathbb{R}_+$ and any $\lambda \in [0, 1]$, we want to show

$$u(\lambda x + (1 - \lambda)y) \geq \lambda u(x) + (1 - \lambda)u(y).$$

Define $W_0 = \lambda x + (1 - \lambda)y$ and consider a fair gamble

$$\tilde{z} = \begin{cases} 
(1 - \lambda)(x - y), & \text{with probability } \lambda; \\
\lambda(y - x), & \text{with probability } (1 - \lambda).
\end{cases}$$

Since the investor is risk averse, we have

$$u(\lambda x + (1 - \lambda)y) = u(W_0) \geq E[u(W_0 + \tilde{z})] = \lambda u(x) + (1 - \lambda)u(y),$$

and since $x, y,$ and $\lambda$ are chosen arbitrarily, this proves that $u(\cdot)$ is concave.

On the other hand, suppose that $u(\cdot)$ is concave, and hence its right-hand and left-hand derivatives both exist everywhere in the domain of definition.\footnote{See the proof of the preceding lemma. The following concave function has $u'_-(0) = 2$ and $u'_+(0) = 1$, so that it is not differentiable at $z = 0$:}

$$u(z) = \begin{cases} 
z, & z \geq 0; \\
2z, & z < 0.
\end{cases}$$

It can be shown that a concave function $u : \mathbb{R} \to \mathbb{R}$ must be continuous, and it is differentiable except at a set of points which has zero Lebesgue measure on $\mathbb{R}$.

13 We have shown in the previous lemma that for all $x, y \in \mathbb{R}$,

$$u(x) \leq u(y) + u'_+(y)(x - y).$$

For any initial wealth $W_0$ and fair gamble $\tilde{z}$, putting $x = W_0 + \tilde{z}$ and $y = W_0 + E[\tilde{z}]$, we have

$$u(W_0 + \tilde{z}) \leq u(W_0 + E[\tilde{z}]) + u'_+(W_0 + E[\tilde{z}]) (\tilde{z} - E[\tilde{z}]).$$

Upon taking expectations on both sides of the above inequality and recognizing that $E[\tilde{z}] = 0$, we obtain

$$E[u(W_0 + \tilde{z})] \leq u(W_0),$$

and hence the investor is risk averse. ||
2. Given a VNM utility function \( u : \mathbb{R}_+ \rightarrow \mathbb{R} \) with \( u'(x) > 0 > u''(x) \) for all \( x > 0 \), define two new functions \( R^u_A(x) \) and \( R^u_R(x) \) as follows.

\[
R^u_A(x) = -\frac{u''(x)}{u'(x)}, \quad R^u_R(x) = -\frac{xu''(x)}{u'(x)}, \quad \forall x > 0.
\]

The two functional values \( R^u_A(x) \) and \( R^u_R(x) \) are referred to respectively as the Arrow-Pratt measures for absolute and for relative risk aversion at the wealth level \( x \); see Arrow (1970) and Pratt (1964). Observe that if \( R^u_A(x) = \rho > 0 \) is a constant function, then \( u(x) = -e^{-\rho x} \), and in this case, we refer to \( u(x) \) as a CARA (constant absolute risk aversion) utility function with \( \rho \) being the associated coefficient of absolute risk aversion. (If \( \rho = 0 \), then apparently \( u(x) = z \), so that the investor is risk neutral.) Similarly, \( u(x) \) is a CRRA (constant relative risk aversion) utility function if \( R^u_R(x) \) is a constant function. In this case either \( u(x) = \log(x) \) or \( u(x) = x^p \) for some \( p \in (0, 1) \) (up to a positive affine transform). If \( R^u_A(x) \) is an increasing (decreasing) function, then \( u(x) \) is referred to as an IARA (DARA) utility function. Some evidence has suggested that most investors have DARA preferences, and in this case if \( u'''(x) \) exists, then one can show that \( u''' \geq 0 \).

3. In a static setting, investors trade assets at date 0 and the assets generate cash flows at date 1. Let \( p \) and \( \tilde{x} \) be the date-0 price of an asset and

\[
\rho = R^u_A(x) = -\frac{u''(x)}{u'(x)} = -\frac{d\log(u'(x))}{dx} \Rightarrow -\rho x + k = \log(u'(x))
\]

\[
\Rightarrow e^{-\rho x + k} = u'(x) \Rightarrow u(x) = a - be^{-\rho x}
\]

for some \( k, a, b \in \mathbb{R}, b \in \mathbb{R}_+^+ \). Since a VNM utility function is unique up to a positive affine transform, we can, for example, pick \( a = 0, b = 1 \).

Note that if \( R^u_R(x) = -\frac{du''(x)}{dx} = \rho > 0 \) for some constant \( \rho \), then we have

\[
\rho = R^u_R(x) = -\frac{u''(x)}{u'(x)} = -\frac{d\log(u'(x))}{dx} \Rightarrow -\rho \log(x) + k = \log(u'(x)),
\]

and we have \( u'(x) = e^{kx^{-\rho}} \), so that if \( \rho = 1 \), \( u(x) = a + b \log(x) \); and if \( \rho \neq 1 \), then \( u(x) = a + bx^{1-\rho} \), for some \( k, a, b \in \mathbb{R}, b \in \mathbb{R}_+^+ \). Again we can pick \( a = 0 \) and \( b = 1 \). To make sure that \( u' > 0 > u'' \), we assume that \( \rho \in (0, 1) \).
the date-1 (random) cash flow generated by the asset. We refer to \( \frac{x_p}{p} - 1 \) and \( E\left[\frac{x_p}{p} - 1\right] \) as respectively the dollar return, the rate of return, and the expected rate of return on the asset. The asset is referred to as riskless if \( \frac{x_p}{p} \) is non-random, and in this case we denote \( E\left[\frac{x_p}{p} - 1\right] \) by \( r_f \). In the presence of a riskless asset, we refer to \( \frac{x_p}{p} - (1 + r_f) \) and \( E\left[\frac{x_p}{p} \right] - (1 + r_f) \) as respectively the excess rate of return and the risk premium on the asset.

4. Suppose that an investor with initial wealth \( W_0 \) can trade a riskless asset and a risky asset at date 0, and let \( \tilde{r} \) be the rate of return on the risky asset. Let \( a \) be the amount of her initial wealth to be invested in the risky asset. Define \( f(a) = E\left[u((W_0 - a)(1 + r_f) + a(1 + \tilde{r}))\right] \). Note that \( f(a) \) is the expected utility that the investor can obtain from the date-1 wealth

\[
\tilde{W} \equiv (W_0 - a)(1 + r_f) + a(1 + \tilde{r}),
\]

which is partially determined by the choice of \( a \). Assume \( u' > 0 > u'' \). Under some mild conditions, we have

\[
f'(a) = E\left[\frac{\partial}{\partial a} u((W_0 - a)(1 + r_f) + a(1 + \tilde{r}))\right] = E[u'(W_0(1 + r_f) + a(\tilde{r} - r_f))(\tilde{r} - r_f)],
\]

\[16\]In a pure exchange economy where consumers are endowed with consumption goods but they cannot produce any of those goods (an unrealistic assumption; I know), \( x \) is an exogenous variable, but \( p \), which is determined in the equilibrium of financial markets, is an endogenous variable. In a production economy where people can produce new goods (or greater amounts of existing goods) using existing goods, \( x \) must be determined endogenously also. However, it is standard in financial engineering that we take the prices of some traded assets (called underlying assets) as exogenous and then determine endogenously the prices of other assets (called derivative assets) accordingly. This procedure is subject to criticism because economic equilibrium should impose some restrictions on the prices of underlying assets, and arbitrarily assuming the behavior of those prices may imply inconsistency with investors’ rationality, or with the markets clearing condition. See He and Leland (1993).

\[17\]If \( a < 0 \), then the investor is selling the risky asset short; and if \( a > W_0 \), then the investor is selling the riskless asset short, or is borrowing.

\[18\]In general, a uniform integrability condition must hold so that switching the order of differentiation and the expectation operators is all right; see Section 4.5 of Chung (1974).
and
\[
f''(a) = E\left[ \frac{\partial^2}{\partial a^2} u((W_0 - a)(1 + r_f) + a(1 + \tilde{r})) \right] \\
= E[u''(W_0(1 + r_f) + a(\tilde{r} - r_f))(\tilde{r} - r_f)^2],
\]
where note that \(u''(W_0(1+r_f)+a(\tilde{r}-r_f))\) is a negative random variable and \((\tilde{r} - r_f)^2\) is a positive random variable, so that \(f''(a) < 0\). This shows that \(f(a)\) is strictly concave in \(a\). The investor seeks to

\[
\max_{a \in \mathbb{R}} f(a);
\]
that is, the investor would like to find the optimal amount of money \(a^*\) that should be spent on the risky asset, taking the asset returns \((\tilde{r}, r_f)\) as given, where optimality means that the investor’s expected utility is maximized.\(^{19}\) Since \(f(\cdot)\) is concave, our mathematical review shows that the optimal amount \(a^*\) to be invested in the risky asset can be found by solving the first-order condition \(f'(a^*) = 0\) if \(f'(+\infty) \equiv \lim_{x \to +\infty} f'(x) < 0 < f'(-\infty) \equiv \lim_{x \to -\infty} f'(x)\); one necessary condition for the existence of a solution to \(f'(a^*) = 0\) is that the probabilities of the two events \(\{\tilde{r} > r_f\}\) and \(\{\tilde{r} < r_f\}\) are both positive, which we shall assume to hold from now on.

**Lemma 8** If \(u(\cdot)\) is twice-differentiable with \(u' > 0 > u''\), then \(a^* > 0\) as long as \(E[\tilde{r}] > r_f\). That is, no matter how risk-averse the investor may be, she will take a long position in the risky asset.

**Proof.** To see this, note that

\[
a^* \leq 0 \iff f'(0) \leq f'(a^*) = 0,
\]
and that

\[
u'(W_0(1 + r_f))(E[\tilde{r}] - r_f) = E[u'(W_0(1 + r_f))(\tilde{r} - r_f)] = f'(0).\]

\(^{19}\)Hence we have assumed that markets are perfect. In particular, the investor is assumed to be a price-taker. (Where?)
Theorem 3 Assume that $E[\tilde{r}] > r_f$. Other things being equal, an increase in $W_0$ leads to an increase (respectively, a decrease) in $a^*$ if $R_A^v$ is a decreasing (respectively, increasing) function. Other things being equal, an increase in $W_0$ leads to an increase (respectively, a decrease) in $\frac{a}{W_0}$ (which is the portfolio weight for the risky asset) if $R_R^v$ is a decreasing (respectively, increasing) function.

Proof. Recall that

$$f(a, W_0) = E[u((W_0 - a)(1 + r_f) + a(1 + \tilde{r})].$$

The optimal amount $a^*$ to be invested in the risky asset is such that

$$L(a^*, W_0) \equiv \frac{\partial f}{\partial a}(a^*, W_0) = 0.$$

The implicit function theorem\(^\text{20}\) says that

$$\frac{da^*}{dW_0} = -\frac{\frac{\partial L}{\partial W_0}}{\frac{\partial L}{\partial a^*}},$$

\(^\text{20}\)Recall that given $e > 0$ and $x = (x_1, x_2) \in \mathbb{R}^2$, the following subset of $\mathbb{R}^2$

$$B(x, e) \equiv \{y \in \mathbb{R}^2 : \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2} < e\}$$

is called the open ball with center $x$ and radius $e$. A subset $A$ of $\mathbb{R}^2$ is an open set if for each $x \in A$, there exists a small $e > 0$ such that $A$ contains the open ball $B(x, e)$. Now, the implicit function theorem says the following. Suppose that $A \subset \mathbb{R}^2$ is an open set and that $F: A \rightarrow \mathbb{R}$ is continuously differentiable with $F_i(\cdot, \cdot)$ being its partial derivative with respect to its $i$-th argument. If at some $z = (z_1, z_2) \in A$,

$$F(z_1, z_2) = 0 \neq F_2(z_1, z_2),$$

then there exist $e > 0$, $\delta > 0$, such that for each $x_1 \in I \equiv (z_1 - \epsilon, z_1 + \epsilon)$, there correspondingly exists a unique $x_2 \in J \equiv (z_2 - \delta, z_2 + \delta)$ such that $F(x_1, x_2) = 0$. In other words, an injective function $f: I \rightarrow J$ can be defined by $F(x_1, f(x_1)) = 0$, or simply $x_2 = f(x_1)$, and moreover, this $f$ is continuously differentiable with

$$f'(x_1) = -\frac{F_1(x_1, x_2)}{F_2(x_1, x_2)},$$

19
as long as \( L \) is continuously differentiable in \( a^* \) and \( W_0 \) and the denominator on the right-hand side is non-zero. Note that except for degenerate cases when \( u'' < 0 \),

\[
\frac{\partial^2 f}{(\partial a)^2}(a, W_0) = E[u''(\tilde{W})(\tilde{r} - r_f)^2] < 0,
\]

where \( \tilde{W} = W_0(1 + r_f) + a(\tilde{r} - r_f) \), and that the aforementioned denominator is exactly \( \frac{\partial^2 f}{(\partial a)^2}(a^*, W_0) \). Thus the sign of \( \frac{da^*}{dW_0} \) is the same as the sign of

\[
\frac{\partial L}{\partial W_0} = E[u''(\tilde{W})(\tilde{r} - r_f)(1 + r_f)],
\]

or equivalently given that \( (1 + r_f) > 0 \), the sign of

\[
E[u''(\tilde{W})(\tilde{r} - r_f)].
\]

Suppose that \( R_u^a \) is a decreasing function so that when \( \tilde{W} \geq W_0(1+r_f) \), we have

\[
R_u^a(\tilde{W}) \leq R_u^a(W_0(1 + r_f)),
\]

and in the event that \( \tilde{W} < W_0(1 + r_f) \),

\[
R_u^a(\tilde{W}) \geq R_u^a(W_0(1 + r_f)).
\]

Now multiplying both sides of the above last two inequalities by \(-u'(\tilde{W})(\tilde{r} - r_f)\), we have

\[
u''(\tilde{W})(\tilde{r} - r_f) \geq -R_u^a(W_0(1 + r_f))u'(\tilde{W})(\tilde{r} - r_f).
\]

Taking expectations on both sides of the last inequality, and using the first-order condition, we conclude that

\[
E[u''(\tilde{W})(\tilde{r} - r_f)] \geq 0.
\]

The case where \( R_u^a(\cdot) \) is an increasing function is similar. This proves the first assertion.

Now, for the second assertion, note that

\[
\frac{da^*}{dW_0} = -\frac{E[u''(\tilde{W})(\tilde{r} - r_f)(1 + r_f)]}{E[u''(\tilde{W})(\tilde{r} - r_f)^2]},
\]

20
so that the elasticity of $a^*$ with respect to changes in $W_0$,

$$\eta \equiv \frac{da^*}{dW_0} \frac{W_0 - a}{a} = 1 + \frac{E[u''(\tilde{W})\tilde{W}(\tilde{r} - r_f)]}{-aE[u''(W)(\tilde{r} - r_f)^2]}.$$ 

Thus $\eta > 1$ (respectively, $\eta < 1$) if and only if $E[u''(\tilde{W})\tilde{W}(\tilde{r} - r_f)]$ is positive (respectively, negative).\(^{21}\)

Suppose that $R_R^u(\cdot)$ is an increasing function. In this case, we have

$$\tilde{r} \geq r_f \Rightarrow R_R^u(\tilde{W}) \geq R_R^u(W_0(1 + r_f))$$

and

$$\tilde{r} < r_f \Rightarrow R_R^u(\tilde{W}) \leq R_R^u(W_0(1 + r_f)),$$

implying that

$$\tilde{r} \geq r_f \Rightarrow [-u'(\tilde{W})(\tilde{r} - r_f)]R_R^u(\tilde{W}) \leq [-u'(\tilde{W})(\tilde{r} - r_f)]R_R^u(W_0(1 + r_f))$$

and

$$\tilde{r} < r_f \Rightarrow [-u'(\tilde{W})(\tilde{r} - r_f)]R_R^u(\tilde{W}) \leq [-u'(\tilde{W})(\tilde{r} - r_f)]R_R^u(W_0(1 + r_f)),$$

so that, whether or not $\tilde{r} \geq r_f$, we always have

$$u''(\tilde{W})\tilde{W}(\tilde{r} - r_f) \leq -R_R^u(W_0(1 + r_f))u'(\tilde{W})(\tilde{r} - r_f)$$

$$\Rightarrow E[u''(\tilde{W})\tilde{W}(\tilde{r} - r_f)] \leq -R_R^u(W_0(1 + r_f))E[u'(\tilde{W})(\tilde{r} - r_f)] = -R_R^u(W_0(1 + r_f))f'(a^*) = 0.$$ 

Thus $\eta \leq 1$ if $R_R^u(\cdot)$ is an increasing function. The case where $R_R^u(\cdot)$ is a decreasing function can be analogously proved. Note that if $u(\cdot)$ is a CRRA utility function, then $\eta = 1$, implying that the portfolio weight for the risky asset, $\frac{a^*}{W_0}$, is independent of the initial wealth $W_0$. \(\parallel\)

\(^{21}\)Observe that

$$\eta > 1 \Leftrightarrow \frac{a^* + da^*}{W_0 + dW_0} > \frac{a^*}{W_0};$$

that is, $\eta > 1$ if and only if increasing the initial wealth $W_0$ locally raises the portfolio weight for the risky asset.
Example 1 Suppose that Mrs. A’s VNM utility function is $u(x) = \sqrt{x}$. Suppose that her initial wealth is $W_0 = 100,000$, and that the riskless lending and borrowing rate is $r_f = 0$. The rate of return on the risky asset, $\tilde{r}$, is equally likely to be $-0.1$ and $0.5$. How much should she borrow or lend? (Show that she should borrow 700,000.)

Example 2 Suppose that Mrs. B’s VNM utility function is $u(x) = -e^{-x}$, and the riskless rate is $r_f = 0.1$. The rate of return on the risky asset, $\tilde{r}$, is a normal random variable with mean $0.3$ and variance $0.04$. Mrs. B has initial wealth $W_0 = 1,000,000$. How much should she spend
on the risky asset?  

Example 3  

Consider an investor endowed with the VNM utility function 

\[ u(W) = \frac{W^{1-\rho}}{1-\rho}, \quad 0 < \rho < 1, \]  

where \( W \) is the investor’s date-1 wealth. The investor is faced with two traded assets at date 0, one with risky rate of return \( \tilde{r} \), and the other with sure rate of return \( r_f \) (which is a lending and

\[ M_z(t) \equiv E[e^{t\tilde{z}}], \quad \forall t > 0. \]

If \( \tilde{z} \sim N(\mu, \sigma^2) \), then

\[ M_z(t) = e^{\mu t + \frac{1}{2}t^2\sigma^2}. \]

Now, consider an investor with von Neumann-Morgenstern utility function 

\[ u(x) = -e^{-Ax}, \quad A > 0 \]  

where \( A > 0 \) is the investor’s measure of absolute risk aversion. The investor seeks to

\[ \max E[u(\tilde{W})] = -E[e^{-A\tilde{W}}], \]

where the terminal wealth \( \tilde{W} \) is normally distributed

\[ \tilde{W} \sim N(\mu_W, \sigma^2_W). \]

Note that

\[ -A\tilde{W} \sim N(-A\mu_W, A^2\sigma^2_W), \]

and hence the investor’s objective function can be re-written as

\[ E[u(\tilde{W})] = -M_{-AW}(1) = -e^{-A\mu_W + \frac{1}{2}A^2\sigma^2_W} \]

\[ = -e^{A[\mu_W - \frac{1}{2}A\sigma^2_W]} = u(\mu_W - \frac{1}{2}A\sigma^2_W), \]

where \( \mu_W - \frac{1}{2}A\sigma^2_W \) is referred to as the certainty equivalent of \( \tilde{W} \) for the investor. Since \( u'(\cdot) > 0 \), the solution that solves

\[ \max E[u(\tilde{W})] = u(\mu_W - \frac{1}{2}A\sigma^2_W) \]

must be identical to the solution that solves

\[ \max \mu_W - \frac{1}{2}A\sigma^2_W. \]

For this reason, the investor essentially has a mean-variance utility function.
borrowing opportunity). Returns are generated at date 1. Let \( W_0 > 0 \) and \( a^* \) be the investor’s initial wealth and the amount of money that she chooses to spend on the risky asset.

(i) Show that if \( \rho = \frac{1}{2} \), \( r_f = 0 \), and

\[
\tilde{r} = \begin{cases} 
2, & \text{with probability } \frac{1}{2}, \\
-1, & \text{with probability } \frac{1}{2},
\end{cases}
\]

then the following portfolio is optimal for the investor:

\[
\left[ \begin{array}{c}
a^* W_0 \\
1 - a^* W_0
\end{array} \right] = \left[ \begin{array}{c} \frac{3}{2} \\
-\frac{1}{2}
\end{array} \right].
\]

(ii) Continue with part (i). If \( W_0 \) is equal to 100,000, how much money does the investor have to lend or borrow at date 0 in order to implement the above optimal portfolio strategy?

Solution. Note that for all \( W > 0 \),

\[
u'(W) = W^{-\rho} > 0 > -\rho W^{-\rho-1} = u''(W),
\]

implying that

\[
f(a) \equiv E[u(W_0(1 + r_f) + a(\tilde{r} - r_f))]
\]

is concave in \( a \), as long as \( W_0(1 + r_f) + a(\tilde{r} - r_f) > 0 \) with probability one. (This is true for \( a \geq 0 \).) Now, for part (i), it suffices to verify that \( a^* = \frac{3W_0}{2} \) satisfies the first-order condition. Note that

\[
f'(a^*) = W_0^{-\rho} \left[ \frac{1}{2} \times (1 + 2 \frac{a^*}{W_0})^{-\rho} \times 2 + \frac{1}{2} \times (1 - \frac{a^*}{W_0})^{-\rho} \times (-\frac{1}{2}) \right]
\]

\[
= W_0^{-\rho} \left[ \frac{2}{\sqrt{1 + 2 \frac{a^*}{W_0}}} - \frac{1}{2 \sqrt{1 - \frac{a^*}{W_0}}} \right]
\]

\[
= W_0^{-\rho} \times \frac{1}{2} \times 0 = 0.
\]

Since the terminal wealth is strictly positive with \( a^* = \frac{3W_0}{2} \), \( a^* = \frac{3W_0}{2} \) is indeed the optimal solution.

Now, for part (ii), it is easy to see that the investor has to borrow \( \frac{1}{2} \times 100,000 = 50,000 \) in order to implement the optimal portfolio.
5. The most widely adopted VNM utility functions include the negative exponential (or, CARA) function \( u(x) = -e^{-\rho x} \); the power function \( u(x) = x^p \), \( 0 < p < 1 \); the logarithmic function \( u(x) = \log(x) \); the linear function \( u(x) = x \); and the quadratic function \( u(x) = x - \frac{b}{2}x^2 \). As we have mentioned, the power and the log functions are CRRA functions. The linear function exhibits risk neutrality. The quadratic function implies that the investor only cares about the first two moments (the expectation and variance) of \( \tilde{W} \), which will be consistent with our analysis in Lecture 4, where an investor that cares only about the expectation and variance of \( \tilde{W} \) is solving her optimal portfolio problem when faced with \( N \) risky assets and one (why only one?) riskless asset available for trading. The quadratic function has the obvious drawback that up to a certain point the marginal utility of wealth becomes negative! Verify that it is also an IARA utility function.

As shown in the above theorem, the demand for the risky asset exhibits no income effect if the investor is endowed with a CARA utility function. That theorem also shows that the percentage of \( W_0 \) allocated to the risky asset will not vary when \( W_0 \) changes if the investor is endowed with a CRRA utility function. In other words, an investor’s optimal portfolio (to be defined more formally shortly) \( \left( \frac{a^*}{W_0}, 1 - \frac{a^*}{W_0} \right) \) is invariant to changes in her initial wealth if she is endowed with a CRRA utility function.

3 Comparing Risk Aversion for Multiple Investors

1. Given two investors, can we say that one investor is more risk averse than the other? Consider two investors U and V endowed with VNM utility functions \( u \) and \( v \) respectively (where \( u', v' > 0 \geq u'', v'' \)) and the same random wealth \( W + \tilde{z} \), where \( \tilde{z} \) is a fair gamble (or pure risk). Which one of them is willing to pay more for an insurance policy that removes from their wealth the pure risk \( \tilde{z} \)?

The maximum amount that investor U is willing to pay for such an insurance policy, denoted by \( \rho^u(W, \tilde{z}) \), or \( \rho^u \) for simplicity, is the solution
to the following equation where $\rho$ is the unknown:

$$u(W - \rho) = E[u(W + \hat{z})].$$

We shall refer to $\rho^u(W, \hat{z})$ as investor U’s risk premium for the pure risk $\hat{z}$. The preceding question can be rephrased as: which between $\rho^u$ and $\rho^v$ is greater? We shall say that investor $U$ is more risk averse than investor $V$ if and only if $\rho^u \geq \rho^v$ for all $W$ and all pure risks $\hat{z}$.

**Theorem 4** The following three statements are equivalent.

(i) For all (non-random) $W$ and all fair gambles $\hat{z}$, $\rho^u(W, \hat{z}) \geq \rho^v(W, \hat{z})$.

(ii) There exists a function $f : v(\mathbb{R}) \to \mathbb{R}$ with $f' > 0 \geq f''$ such that $u(\cdot) = f(v(\cdot))$.\(^{23}\)

(iii) For all $x \in \mathbb{R}$, $R_A^u(x) \geq R_A^v(x)$.

**Proof.** We shall assume that $u''(\cdot)$ and $v''(\cdot)$ are both continuous functions. We will show that (ii)$\Rightarrow$(i), (iii)$\Rightarrow$(ii), and (i)$\Rightarrow$(iii).

(ii)$\Rightarrow$(i). By definition, we have\(^{24}\)

$$\rho^u(W, \hat{z}) = W - u^{-1}(E[u(W + \hat{z})])$$

and

$$\rho^v(W, \hat{z}) = W - v^{-1}(E[v(W + \hat{z})]),$$

so that

$$\rho^u(W, \hat{z}) - \rho^v(W, \hat{z}) = v^{-1}(E[v(W + \hat{z})]) - u^{-1}(E[u(W + \hat{z})])$$

$$= v^{-1}(E[v(W + \hat{z})]) - u^{-1}(E[u(v(W + \hat{z}))])$$

$$= v^{-1}(E[v(W + \hat{z})]) - u^{-1}(E[f(v(W + \hat{z}))]).$$

---

\(^{23}\)Recall that the image of $h : A \to B$, denoted by $h(A)$, is the set $\{h(a) : a \in A\}$. The pre-image of $C \subset B$ under function $h$, denoted by $h^{-1}(C)$, is the set $\{a : h(a) \in C, a \in A\}$. For example, if $h : \mathbb{R} \to \mathbb{R}$ is defined by $h(a) = 1 - e^{-a}$, then $h(\mathbb{R}) = (-\infty, 1)$, and $h^{-1}((0, 1)) = (1 - e, 0)$. For another example, if $h : \mathbb{R} \to \mathbb{R}$ is defined by $h(z) = 2 + z^2$, then $h(\mathbb{R}) = [2, +\infty)$, $h^{-1}((0, 2)) = \{0\}$, and $h^{-1}((-1, 3)) = (-1, 1)$. In the theorem, $v(\mathbb{R})$ denotes the image of the function $v : \mathbb{R} \to \mathbb{R}$.

\(^{24}\)The inverse functions $u^{-1}$ and $v^{-1}$ are defined on respectively $u(\mathbb{R})$ and $v(\mathbb{R})$. 

26
where we have used the fact that (ii) holds, and hence
\[ u(\cdot) = f(v(\cdot)) \Rightarrow u(v^{-1}(\cdot)) = f(v^{-1}(\cdot)) = f(\cdot), \]
which is a concave function. Note that \( u' > 0 \) and hence \( v^{-1}(E[v(W + \tilde{z})]) - u^{-1}(E[f(v(W + \tilde{z}))]) \) is non-negative if and only if
\[ u(v^{-1}(E[v(W + \tilde{z})])) - u^{-1}(E[f(v(W + \tilde{z}))])) \geq 0, \]
or equivalently,
\[ f(E[v(W + \tilde{z})]) \geq E[f(v(W + \tilde{z}))], \]
but by Jensen’s inequality this last inequality is always true, because of concavity of \( f \)! Thus we have shown that if (ii) holds, then (i) holds.

(iii)⇒(ii). Define \( f(\cdot) = u((v^{-1}(\cdot)) \). We must show that \( f' > 0 \geq f'' \).
At first, since
\[ u'(x) = f'(v(x)) \cdot v'(x), \]
and since \( u', v' > 0 \), we indeed have \( f' > 0 \). Next, by differentiating
\[ u'(x) = f'(v(x)) \cdot v'(x), \]
we have
\[ u''(x) = f''(v(x))[v'(x)]^2 + f'(v(x))v''(x) \]
\[ \Rightarrow \frac{u''(x)}{u'(x)} = \frac{f''(v(x))[v'(x)]^2 + f'(v(x))v''(x)}{f'(v(x)) \cdot v'(x)} \]
\[ \Rightarrow R_A''(x) = R_A'(v(x))v'(x) + R_A'(x), \]
which, because \( v' > 0 \) and (iii) holds, implies that \( R_A'(v(x)) \geq 0 \). Because of \( f' > 0 \), we conclude that \( f''(v(x)) \leq 0 \). Thus we have shown that if (iii) holds then (ii) holds.

(i)⇒(iii). Suppose instead that (i) holds but for some \( x \in \mathbb{R} \), \( R_A'(x) < \frac{R_A''(x)}{R_A'(x)} \) so that (iii) does not hold. We shall demonstrate a contradiction. Since \( u''(\cdot) \) and \( v''(\cdot) \) are both continuous functions, \( R_A'(\cdot) \) and \( R_A''(\cdot) \) are
continuous functions also. If at \( x \in \mathbb{R} \), \( R^u_A(x) < R^v_A(x) \), then for each point \( y \) contained in a small interval \((x - e, x + e)\), where \( e > 0 \) is very small, it remains true that \( R^u_A(y) < R^v_A(y) \). (This is one well-known property of a real-valued continuous function.) Now, let \( W = x \) and let \( \tilde{z} \) be a fair gamble such that every realization of \( x + \tilde{z} \) is contained in the small interval \((x - e, x + e)\). Since on this interval, \( R^u_A(\cdot) < R^v_A(\cdot) \), the first two steps of our proof together show that, for this particular pair \((W, \tilde{z})\), we have \( \rho^u(W, \tilde{z}) < \rho^v(W, \tilde{z}) \), which contradicts the assumption that (i) holds. We thus conclude that if (i) holds then (iii) must hold also. \( \| \)

The next proposition shows that, when allowed to trade one risky asset and one riskless asset, if one investor \( U \) is more risk-averse than the other investor \( V \) according to the preceding definition, then the less risk-averse investor \( V \) will spend more money on the risky asset.

**Proposition 1** Suppose that \( E[\tilde{r}] > r_f \). Investor \( V \) will spend more than \( U \) on the risky asset if \( U \) is more risk averse than \( V \); that is, if for all (non-random) \( W_0 \) and all fair gambles \( \tilde{z} \), \( \rho^u(W_0, \tilde{z}) \geq \rho^v(W_0, \tilde{z}) \).

**Proof.** Let \( a \) be the optimal amount of money that \( V \) spends on the risky asset; that is,

\[
E[u'(W_0(1 + r_f) + a(\tilde{r} - r_f))(\tilde{r} - r_f)] = 0.
\]

We know that \( a > 0 \). It suffices to show that

\[
E[u'(W_0(1 + r_f) + a(\tilde{r} - r_f))(\tilde{r} - r_f)] \leq 0.
\]

By the preceding theorem, we have, for all \( z \in \mathbb{R} \),

\[
u'(z) = f'(v(z))v'(z)
\]
for some \( f : v(\mathbb{R}) \to \mathbb{R} \) with \( f' > 0 \geq f'' \), so that
\[
E[u'(W_0(1 + r_f) + a(\tilde{r} - r_f))(\tilde{r} - r_f)]
\]
\[
= E[f'(v(W_0(1 + r_f) + a(\tilde{r} - r_f))v'(W_0(1 + r_f) + a(\tilde{r} - r_f))(\tilde{r} - r_f)].
\]
Now observe that, because \( f'(-) \) and \( v(-) \) are respectively decreasing and increasing functions, we have\(^{25}\)
\[
f'(v(W_0(1 + r_f) + a(\tilde{r} - r_f)))(\tilde{r} - r_f) \leq f'(v(W_0(1 + r_f)))(\tilde{r} - r_f).
\]
Since \( v' > 0 \), upon multiplying the above inequality by \( v'(W_0(1 + r_f) + a(\tilde{r} - r_f)) \) and then taking expectations on both sides, we have
\[
E[f'(v(W_0(1 + r_f) + a(\tilde{r} - r_f))v'(W_0(1 + r_f) + a(\tilde{r} - r_f))(\tilde{r} - r_f)] \leq f'(v(W_0(1 + r_f)))E[v'(W_0(1 + r_f) + a(\tilde{r} - r_f))(\tilde{r} - r_f)] = 0.
\]
This shows that investor \( U \) will spend less money on the risky asset. \( \| \)

The preceding proposition cannot be extended to the more general case where riskless assets do not exist. For example, suppose both traded assets are risky, and denote their rates of return by respectively \( \tilde{r}_A \) and \( \tilde{r}_B \), where
\[
E[\tilde{r}_A - \tilde{r}_B | \tilde{r}_B] \geq 0,
\]
\(^{25}\)Recall that \( a > 0 \) because \( E[\tilde{r}] > r_f \). If the realized \( \tilde{r} \geq r_f \), then since \( a > 0 \) and \( f'' \leq 0 \), we have
\[
v(W_0(1 + r_f) + a(\tilde{r} - r_f)) > v(W_0(1 + r_f))
\]
\[
\Rightarrow f'(v(W_0(1 + r_f) + a(\tilde{r} - r_f))) \geq f'(v(W_0(1 + r_f)))
\]
\[
\Rightarrow f'(v(W_0(1 + r_f) + a(\tilde{r} - r_f))(\tilde{r} - r_f)) \leq f'(v(W_0(1 + r_f)))(\tilde{r} - r_f);
\]
on the other hand, if the realized \( \tilde{r} < r_f \), then since \( a > 0 \) and \( f'' \leq 0 \), we have
\[
v(W_0(1 + r_f) + a(\tilde{r} - r_f)) < v(W_0(1 + r_f))
\]
\[
\Rightarrow f'(v(W_0(1 + r_f) + a(\tilde{r} - r_f))) \geq f'(v(W_0(1 + r_f)))
\]
\[
\Rightarrow f'(v(W_0(1 + r_f) + a(\tilde{r} - r_f))(\tilde{r} - r_f)) \leq f'(v(W_0(1 + r_f)))(\tilde{r} - r_f).
for every realization of $\tilde{r}_B$. Here, asset B is less risky than asset A, but asset A also has a higher expected rate of return. More precisely, let $\tilde{r}_A = \tilde{r}_B + \tilde{z}$, where $\tilde{r}_B$ and $\tilde{z}$ are independent random variables, with

$$\tilde{z} = \begin{cases} 2, & \text{with probability } \frac{1}{2}; \\
-1, & \text{with probability } \frac{1}{2}; 
\end{cases}$$

$$\tilde{r}_B = \begin{cases} 1, & \text{with probability } \frac{1}{2}; \\
0, & \text{with probability } \frac{1}{2}. 
\end{cases}$$

Now let $v(\cdot)$ be such that $v' > 0 > v''$ and such that

$$v'(\frac{5}{2}) = 0, \quad v'(\frac{7}{4}) = 2, \quad v'(\frac{3}{2}) = 3, \quad v'(\frac{3}{4}) = 4.$$ 

Let $u(\cdot) = f(v(\cdot))$ with $f' \geq 0 \geq f''$, and with

$$f'(v(\frac{5}{2})) = 0, \quad f'(v(\frac{7}{4})) = 0, \quad f'(v(\frac{3}{2})) = 10, \quad f'(v(\frac{3}{4})) = 10.$$ 

Assume that both U and V have initial wealth $W_0 = 1$. Then it can be verified that V should optimally spend $\frac{1}{4}$ dollars on asset A:

$$E[v'(1 + \tilde{r}_B + \frac{\tilde{z}}{4})\tilde{z}] = 0.$$ 

Although $R^u_A(\cdot)$ lies above $R^v_A(\cdot)$, in this example, U wants to spend more than $\frac{1}{4}$ dollars on asset A! In fact, it can be easily verified that

$$E[u'(1 + \tilde{r}_B + \frac{\tilde{z}}{4})\tilde{z}] > 0.$$ 

The above example demonstrates the deficiency of Arrow-Pratt measure for absolute risk aversion. In fact, when the initial wealth is a random variable, the Arrow-Pratt measure also fails to arrive at the desired conclusion that more risk-averse people should be willing to pay more for a given insurance policy than their less risk-averse counterparts. We know that the conclusion is true if people are endowed with the common initial wealth is $W_0 + \tilde{z}$, where $W_0$ is non-random and $\tilde{z}$ is the insurable (pure) risk. However, suppose that U and V are both endowed with random initial wealth $\tilde{W}_0 + \tilde{z}$, where $\tilde{W}_0$ is independent.

30
of $\tilde{z}$. In this case, even if $u(\cdot) = f(v(\cdot))$ for some $f' > 0 > f''$, there exist $(\tilde{W}_0, \tilde{z})$ such that $V$ is willing to pay more than $U$ for an insurance policy that removes the pure risk $\tilde{z}$. For these reasons, Stephen Ross provides the following stronger definition of absolute risk aversion.\textsuperscript{26}

Two investors $U$ and $V$ with VNM utility functions $u$ and $v$ are such that $U$ is \textit{strongly more risk averse} than $V$ if $u', v' > 0 > u'', v''$, and if

$$\inf_{z \in \mathbb{R}} \frac{u''(z)}{v''(z)} \geq \sup_{x \in \mathbb{R}} \frac{u'(x)}{v'(x)},$$

where $\inf$ and $\sup$ denote the \textit{greatest lower bound} and \textit{least upper bound} operators.\textsuperscript{27}

\textbf{Lemma 9} \textit{If $U$ is strongly more risk-averse than $V$ in the sense of Ross, then $U$ is more risk-averse than $V$ in the sense of Arrow and Pratt.}

\textit{Proof.} Suppose that $U$ is strongly more risk-averse than $V$, so that

$$\inf_{z \in \mathbb{R}} \frac{u''(z)}{v''(z)} \geq \sup_{x \in \mathbb{R}} \frac{u'(x)}{v'(x)}.$$ 

This implies that, for all $W \in \mathbb{R}$, we have

$$\frac{u''(W)}{v''(W)} \geq \inf_{z \in \mathbb{R}} \frac{u''(z)}{v''(z)} \geq \sup_{x \in \mathbb{R}} \frac{u'(x)}{v'(x)} \geq \frac{u'(W)}{v'(W)},$$

$$\Rightarrow \frac{u''(W)}{u'(W)} \leq \frac{v''(W)}{v'(W)} \Rightarrow \frac{u''(W)}{u'(W)} \geq \frac{v''(W)}{v'(W)},$$

which, by the preceding theorem, implies that $U$ is more risk-averse than $V$ in the sense of Arrow and Pratt. ||


\textsuperscript{27}Given a subset $A$ of $\mathbb{R}$, a real number $u$ is an upper bound of $A$, if $u \geq a$, $\forall a \in A$. Let $U$ be the set of upper bounds of $A$. If $U$ has a smallest element (denoted by $\min U$), then we define that element of $U$ as $\sup A$; that is, $\sup A$ is the smallest among all the upper bounds of $A$. Similarly, a real number $b$ is a lower bound of $A$ if $a \geq b$, $\forall a \in A$. Let $B$ be the set of lower bounds of $A$. If $B$ has a largest element (denoted max $B$), then we define that element of $B$ by $\inf A$; that is, $\inf A$ is the largest among all the lower bounds of $A$. If $A$ has a largest element, then $\sup A = \max A$. If $A$ has a smallest element, then $\inf A = \min A$. 

31
Theorem 5  The following three statements are equivalent.

(i) $U$ is strongly more risk averse than $V$.

(ii) There are a constant $\lambda \in \mathbb{R}_{++}$ and a function $G : \mathbb{R} \to \mathbb{R}$ with $0 \geq G', G''$ such that

$$u(z) = \lambda v(z) + G(z), \quad \forall z \in \mathbb{R}. $$

(iii) For all pair $(\tilde{W}, \tilde{e})$ of random variables with $E[\tilde{e}|\tilde{W}] = 0$ for each and every realization of $\tilde{W}$, if we define the constants $\rho^u$ and $\rho^v$ as such that

$$E[u(\tilde{W} - \rho^u)] = E[u(\tilde{W} + \tilde{e})], \quad E[v(\tilde{W} - \rho^v)] = E[v(\tilde{W} + \tilde{e})],$$

then $\rho^u \geq \rho^v$.

Proof. We shall prove the equivalence of (i) and (ii), and refer the reader to Ross (1981) for the proof for the remaining assertions.

(i) $\Rightarrow$ (ii): Since $\inf_{z \in \mathbb{R}} \frac{u''(z)}{v''(z)} \geq \sup_{x \in \mathbb{R}} \frac{u'(x)}{v'(x)}$, there exists $\lambda \in \mathbb{R}_{++}$ such that

$$\inf_{z \in \mathbb{R}} \frac{u''(z)}{v''(z)} \geq \lambda \geq \sup_{x \in \mathbb{R}} \frac{u'(x)}{v'(x)}.$$ 

Now define

$$\forall z \in \mathbb{R}, \quad G(z) \equiv u(z) - \lambda v(z).$$

It remains to show that $G', G'' \leq 0$. Note that

$$\forall z \in \mathbb{R}, \quad G'(z) = u'(z) - \lambda v'(z) = v'(z)\left[\frac{u'(z)}{v'(z)} - \lambda\right] \leq 0,$$

and

$$\forall x \in \mathbb{R}, \quad G''(x) = u''(x) - \lambda v''(x) = v''(x)\left[\frac{u''(x)}{v''(x)} - \lambda\right] \leq 0,$$

where the last inequality follows because $v'' \leq 0$. This proves statement (ii).
(ii) ⇒ (i): Since $u$ and $v$ are related by
\[ u(z) = \lambda v(z) + G(z), \quad \forall z \in \mathbb{R}, \]
we have, because $G' \leq 0$,
\[ u'(z) = \lambda v'(z) + G'(z) \leq \lambda v'(z), \quad \forall z \in \mathbb{R}, \]
and, because $G'' \leq 0$,
\[ u''(x) = \lambda v''(x) + G''(x) \leq \lambda v''(x), \quad \forall x \in \mathbb{R}. \]
The last two relations then imply
\[ \frac{u''(x)}{v''(x)} \geq \lambda \geq \frac{u'(z)}{v'(z)}, \quad \forall x, z \in \mathbb{R}. \]
Hence we have
\[ \inf_{z \in \mathbb{R}} \frac{u''(z)}{v''(z)} \geq \sup_{x \in \mathbb{R}} \frac{u'(x)}{v'(x)}, \]
which is statement (i). \|
then we are done. Recall that for some \( \lambda > 0 \) and some \( G(\cdot) \) with \( 0 > G', G'' \), \( u'(x) = \lambda v'(x) + G'(x) \) for all \( x \in \mathbb{R} \). Hence we have

\[
E[u'(1 + \tilde{r}_B + a \tilde{z}) \tilde{z}]
= E[\lambda v'(1 + \tilde{r}_B + a \tilde{z}) \tilde{z} + G'(1 + \tilde{r}_B + a \tilde{z}) \tilde{z}]
= E[G'(1 + \tilde{r}_B + a \tilde{z}) \tilde{z}]
= E[E[G'(1 + \tilde{r}_B + a \tilde{z}) \tilde{z}\mid \tilde{r}_B]]
\]

where the second equality follows from V’s first-order condition, the third equality from the law of iterated expectations, the first inequality from the fact that \( E[\tilde{z}\mid \tilde{r}_B] \geq 0 \) and \( G' \leq 0 \) and the last inequality from the fact that \( G'(\cdot) \) is weakly decreasing and \( a \geq 0 \).\(^{28}\)

\[^{28}\]The last inequality uses the following lemma: If \( f : \mathbb{R} \to \mathbb{R} \) is decreasing and \( \tilde{x} \) and \( f(\tilde{x}) \) are both random variables with finite variances, then

\[
\text{cov}(f(\tilde{x}), \tilde{x}) \leq 0.
\]

To see that the lemma holds true, note that by definition

\[
\text{cov}(f(\tilde{x}), \tilde{x}) = E[[f(\tilde{x}) - E[f(\tilde{x})]][\tilde{x} - E[\tilde{x}]])
\]

\[
= E[f(\tilde{x})][\tilde{x} - E[\tilde{x}]])
\]

\[
= E[[f(\tilde{x}) - f(E[\tilde{x}])][\tilde{x} - E[\tilde{x}]]) \leq 0.
\]

To see that \( a \geq 0 \), recall that \( a \) satisfies

\[
V'(a) \equiv E[v'(1 + \tilde{r}_B + a \tilde{z}) \tilde{z}] = 0,
\]

where \( V(a) \equiv E[v(1 + \tilde{r}_B + a \tilde{z})] \) is concave in \( a \). Observe that

\[
V'(0) = E[v'(1 + \tilde{r}_B) \tilde{z}] = E[E[v'(1 + \tilde{r}_B) \tilde{z}\mid \tilde{r}_B]]
\]

\[
= E[v'(1 + \tilde{r}_B) E[\tilde{z}\mid \tilde{r}_B]] \geq E[v'(1 + \tilde{r}_B) \cdot 0] = 0 = V'(a).
\]

Since \( V'(\cdot) \) is weakly decreasing, and since \( V'(0) \geq V'(a) \), we conclude that \( a \geq 0 \).
4 Comparing Riskiness of Investment Projects

1. There are several criteria that have been used by financial economists to rank risky assets. The first criterion is called *first-degree or first-order stochastic dominance*, and it is defined as follows. Consider two risky assets $h$ and $g$ of which the rates of return $\tilde{r}_h$ and $\tilde{r}_g$ take values in a common support, say the unit interval $[0, 1]$. (The compact-support assumption is made only to ease the exposition, and the following main results hold for general distributions.) Let the distribution functions of $\tilde{r}_h$ and $\tilde{r}_g$ be respectively $H$ and $G$, where $H(z) = G(z) = 0$ if $z < 0$ and $H(z) = G(z) = 1$ if $z \geq 1$.

Now imagine that we are able to invite all investors whose VNM utility functions’ first derivatives are non-negative to rank these two distribution functions. Assume that all investors are endowed with a non-random initial wealth, say 1 dollar. (Again, the one-dollar assumption is made only for ease of demonstration.) Let $U_1$ be this set of VNM utility functions. We say that $H$ (stochastically) dominates $G$ in the first degree (or in the first order), written as $H \geq_{FSD} G$, if and only if every investor contained in $U_1$ (weakly) prefers $H$ to $G$; that is, if and only if

$$\int_{[0, 1]} u(1 + z) dH(z) \geq \int_{[0, 1]} u(1 + z) dG(z), \ \forall u \in U_1.$$

**Theorem 6** The following statements are equivalent.

(i) The distribution functions $H, G$ are such that $H(z) \leq G(z)$ for all $z \in \mathbb{R}$.

(ii) $H \geq_{FSD} G$.

(iii) There exists a random variable $\tilde{e} \geq 0$ such that $\tilde{r}_h$ and $\tilde{r}_g + \tilde{e}$ have the same distribution function; that is, $\tilde{r}_h$ and $\tilde{r}_g + \tilde{e}$ are equal in distribution.\(^{29}\)

*Proof.* We shall prove the equivalence of (i) and (ii). The reader can find the proof for the equivalence between (iii) and (i) in Rothschild

---

\(^{29}\)Suppose that at time 0 two agents $U$ and $V$ are faced with 3 possible time-1 states, $\omega_1, \omega_2, \omega_3$. Let us call $\Omega \equiv \{\omega_1, \omega_2, \omega_3\}$ the sample space. Let $P$ and $Q$ be $U$’s and $V$’s subjective probabilities regarding the 3 possible time-1 states. A (real-valued) random variable (r.v.) is roughly a function mapping from $\Omega$ into $\mathbb{R}$. Consider the probability measures $P$ and $Q$ and the 6 random variables described in the table below.
The table says that $U$ assigns zero probability to the event $\{\omega_3\}$ and $V$ assigns zero probability to the event $\{\omega_1\}$. Apparently, $x : \Omega \to \mathbb{R}$ and $y : \Omega \to \mathbb{R}$ are the same function, and hence we say the two random variables are equal, and we write $\tilde{x} = \tilde{y}$. Although $x : \Omega \to \mathbb{R}$ and $z : \Omega \to \mathbb{R}$ are not the same function, but from $U$'s perspective, the event $\{\omega \in \Omega : x(\omega) \neq z(\omega)\} = \{\omega_3\}$ may occur only with probability zero. Hence we say that $\tilde{x} = \tilde{z} P$-almost surely. Similarly, from $V$'s perspective, $\tilde{z}$ and $\tilde{s}$ are nearly the same (the two functions differ only on an event that, from $V$'s perspective, may occur with zero probability), and hence we say that $\tilde{z} = \tilde{s} Q$-almost surely. Finally, neither $U$ nor $V$ would regard $\tilde{t}$ and $\tilde{w}$ as the same function, but note that from $U$’s perspective, these two random variables share the same distribution function; that is, for each and every $r \in \mathbb{R}$, we have

$$P(\{\omega \in \Omega : t(\omega) \leq r\}) = P(\{\omega \in \Omega : w(\omega) \leq r\}),$$

and hence we say that from $U$’s perspective, $\tilde{t}$ and $\tilde{w}$ are equal in distribution. Note that given any constant $W_0$ and given any von Neumann-Morgenstern utility function $u(\cdot)$, we have $E[u(W_0 + \tilde{t})] = E[u(W_0 + \tilde{w})]$ as long as $\tilde{t}$ and $\tilde{w}$ are equal in distribution. Finally, observe that $\tilde{x} = \tilde{y}$ implies that $\tilde{x} = \tilde{y}$ $P$-almost surely for any probability measure $P$ defined on $\Omega$; and if $\tilde{x} = \tilde{y}$ $P$-almost surely under some probability measure $P$, then under $P$, $\tilde{x}$ and $\tilde{y}$ must share the same distribution function.

30While the necessity of (iii) is difficult to prove, the sufficiency of (iii) is easy to establish. Suppose that $u$ is any element of $U_1$ and suppose that (iii) holds. Then for each realizations $(r_g, e)$ of the random variables $(\tilde{r}_g, \tilde{e})$ we must have, by $e \geq 0$ and $u' \geq 0$,

$$u(1 + r_g + e) \geq u(1 + r_g),$$

so that, upon taking expectation on both sides of the last inequality, we have

$$E[u(1 + \tilde{r}_g + \tilde{e})] \geq E[u(1 + \tilde{r}_g)].$$

Now, since the two random variables $\tilde{r}_h$ and $\tilde{r}_g + \tilde{e}$ share the same distribution function, the two random variables $u(1 + \tilde{r}_h)$ and $u(1 + \tilde{r}_g + \tilde{e})$ must also share the same distribution.
(i)⇒ (ii). Assume that \( H(z) \leq G(z) \) for all \( z \in \mathbb{R} \). We must show that

\[
\int_{[0,1]} u(1+z)d[H(z) - G(z)] \geq 0, \quad \forall u \in U_1.
\]

By definition, the integral can be expanded into

\[
u(1+0)[H(0) - G(0)] + \int_{0}^{1} u(1+z)d[H(z) - G(z)],
\]

which, by integration by parts, is equal to

\[
u(1+0)[H(0) - G(0)] + \int_{0}^{1} u(1+z)[H(z) - G(z)]du(1+z)
\]

\[=
u(1+0)[H(0) - G(0)] - u(1+0)[H(0) - G(0)] + u(1+1)[H(1) - G(1)]
\]

\[- \int_{0}^{1} [H(z) - G(z)]du(1 + z),
\]

which, by the fact that \( H(1) = G(1) = 1 \), is equal to

\[- \int_{0}^{1} [H(z) - G(z)]u'(1 + z)dz \geq 0,
\]

where the last inequality follows from the fact that \( H(z) - G(z) \leq 0 \) for all \( z \in (0, 1) \), and \( u' \geq 0 \). Thus we have shown that every investor contained in \( U_1 \) prefers risky asset \( h \) to risky asset \( g \), and hence \( H \geq_{FSD} G \).

(ii)⇒ (i). Suppose that \( H \geq_{FSD} G \), so that for all \( u \) with \( u' \geq 0 \),

\[
\int_{[0,1]} u(1+z)dH(z) \geq \int_{[0,1]} u(1+z)dG(z).
\]

function, and hence the latter two must have the same expected value; that is,

\[
E[u(1 + \tilde{r}_h)] = E[u(1 + \tilde{r}_g + \tilde{e})].
\]

It thus follows that

\[
E[u(1 + \tilde{r}_h)] \geq E[u(1 + \tilde{r}_g)].
\]

The last inequality being true for an arbitrarily chosen \( u \in U_1 \), we conclude that \( H \geq_{FSD} G \).
We must show that $H(z) \leq G(z)$ for all $z \in \mathbb{R}$. Suppose instead that at $x \in [0, 1)$, $H(x) > G(x)$. Since $H(\cdot)$ and $G(\cdot)$ are right-continuous, for each point $y$ contained in the small interval $[x, x + e]$, where $e > 0$ is very small, it remains true that $H(y) > G(y)$. Now, define a VNM utility function $u$ as follows.

$$u(W) = \begin{cases} 
0; & W \leq 1 + x; \\
W - x - 1; & W \in [1 + x, 1 + x + e]; \\
e; & W \in (1 + x + e, +\infty). 
\end{cases}$$

Note that this utility function is strictly increasing only on the interval $[1 + x, 1 + x + e]$, and since $H(y) > G(y)$ for all $y \in [x, x + e]$, our proof for the assertion that (i) $\Rightarrow$ (ii) then implies that

$$E[u(1 + \tilde{r}_h)] - E[u(1 + \tilde{r}_g)] = -\int_0^1 [H(z) - G(z)]u'(1 + z)dz < 0,$$

so that an investor endowed with 1 dollar and with the above special VNM utility function will strictly prefer risky asset $g$ to risky asset $h$, which is a contradiction to the assumption that $H \geq_{FSD} G$ holds. Hence we have shown that, if $H \geq_{FSD} G$ holds, then it is necessary that $H(z) \leq G(z)$ for all $z \in \mathbb{R}$.$^{31}$

We can similarly invite all investors whose VNM utility functions are weakly concave to rank those two distribution functions. Let $U_2$ be the set of weakly concave VNM utility functions; i.e. those $u(\cdot)$ with $u'' \leq 0$ almost everywhere on $(1, 2)$. We say that $H$ (stochastically) dominates $G$ in the second degree (or in the second order), denoted by

$^{31}$One may feel uncomfortable about the fact that the $u(W)$ that we construct is not everywhere differentiable; $u'(1 + x)$ and $u'(1 + x + e)$ are not defined. This is all right actually. Since the expected utility is defined as a Lebesgue integral, for the expected utility to be well-defined, we only need $u'$ to be defined almost everywhere on $(1, 2)$. Here, a property holds almost everywhere on $\mathbb{R}$ if the set of points where that property fails has zero Lebesgue measure on $\mathbb{R}$; you can think of the Lebesgue measure of a subset of $\mathbb{R}$ as the length of that set. The Lebesgue measure for an interval $[a, b]$ is $b - a$, where $b > a$. The Lebesgue measure of a single-point set is zero. The set of integers also has a zero Lebesgue measure, for each integer has a zero Lebesgue measure, and there are only countably infinite integers.
\( H \geq_{SSD} G \), if and only if every investor with her VNM utility function \( u \) contained in \( \mathcal{U}_2 \) (weakly) prefers \( H \) to \( G \); that is, if and only if

\[
\int_{[0,1]} u(1 + z)dH(z) \geq \int_{[0,1]} u(1 + z)dG(z), \quad \forall u \in \mathcal{U}_2.
\]

**Theorem 7** The following statements are equivalent.

(i) The distribution functions \( H, G \) are such that (A) \( S(y) = \int_0^y H(z) - G(z)dz \leq 0 \) for all \( y \in \mathbb{R} \); and (B) \( \tilde{r}_h \) and \( \tilde{r}_g \) have the same expected value.

(ii) \( H \geq_{SSD} G \).

(iii) There exists a random variable \( \tilde{e} \) such that for each and every realization of \( \tilde{r}_h \), \( E[\tilde{e}|\tilde{r}_h] = 0 \), and such that \( \tilde{r}_g \) and \( \tilde{r}_h + \tilde{e} \) have the same distribution function.

**Proof.** We shall prove the equivalence of (i) and (ii). The reader can find the proof for the equivalence between (iii) and (i) in Rothschild
and Stiglitz (1970).\textsuperscript{32}

(i)$\Rightarrow$ (ii). Assume that (A) and (B) in statement (i) hold. Note that $S(0) = 0$. Note that $S(1) = \int_0^1 [(1 - G(z)) - (1 - H(z))]dz = E[\tilde{r}_g] - E[\tilde{r}_h]$, and hence by (B), we also have $S(1) = 0$. Now we must show that

$$\int_{[0,1]} u(1 + z)d[H(z) - G(z)] \geq 0, \quad \forall u \in U_2.$$ 

Recall that the above integral can be expanded into

$$-\int_0^1 [H(z) - G(z)]u'(1 + z)dz = -\int_0^1 u'(1 + z)dS(z),$$

\textsuperscript{32}While the necessity of (iii) is difficult to prove, the sufficiency of (iii) is easy to establish. Suppose that $u$ is any element of $U_2$ and suppose that (iii) holds. Then since the two random variables $u(1 + \tilde{r}_g)$ and $u(1 + \tilde{r}_h + \tilde{e})$ share the same distribution function, they must have the same expected value; that is,

$$E[u(1 + \tilde{r}_g)] = E[u(1 + \tilde{r}_h + \tilde{e})].$$

The above right side, by the Law of Iterated Expectations (LIE), is equal to

$$E[u(1 + \tilde{r}_h + \tilde{e})] = E[E[u(1 + \tilde{r}_h + \tilde{e})|\tilde{r}_h]].$$

Note that given any realization $r_h$ of $\tilde{r}_h$, $\tilde{e}$ is a fair gamble in the sense that

$$E[\tilde{e}|\tilde{r}_h = r_h] = 0,$$

so that we have, by Jensen’s inequality and $u'' \leq 0$,

$$E[u(1 + r_h + \tilde{e})|\tilde{r}_h = r_h] \leq u(1 + r_h + E[\tilde{e}|\tilde{r}_h = r_h]) = u(1 + r_h).$$

The last inequality being true for any realization $r_h$ of $\tilde{r}_h$, we have, upon taking expectation on both sides of the last inequality,

$$E[E[u(1 + \tilde{r}_h + \tilde{e})|\tilde{r}_h]] \leq E[u(1 + \tilde{r}_h)].$$

It follows that

$$E[u(1 + \tilde{r}_g)] = E[u(1 + \tilde{r}_h + \tilde{e})] \leq E[u(1 + \tilde{r}_h)].$$

The last inequality being true for an arbitrarily chosen $u \in U_2$, we conclude that $H \geq_{SSD} G$. 

40
which, by integration by parts, is equal to
\[ \int_{0}^{1} S(z) du'(1 + z) - \{u'(1 + z)S(z)\}_{0}^{1}, \]
which, by the fact that \( S(0) = S(1) = 0 \) when (B) holds, is equal to
\[ \int_{0}^{1} S(z) du'(1 + z) = \int_{0}^{1} S(z)u''(1 + z)dz \geq 0, \]
where the last inequality follows from the fact that (A) holds. Hence when (A) and (B) both hold, every investor contained in \( \mathcal{U}_{2} \) prefers risky asset \( h \) to risky asset \( g \), and hence \( H \geq_{SSD} G \).

(ii) \( \Rightarrow \) (i). Suppose that \( H \geq_{SSD} G \), so that for all \( u \) with \( u'' \leq 0 \),
\[ \int_{[0,1]} u(1 + z)dH(z) \geq \int_{[0,1]} u(1 + z)dG(z). \]
We must show that (A) and (B) both hold. Note that \( u(W) = W \) and \( v(W) = -W \) are both members of \( \mathcal{U}_{2} \), and since \( H \geq_{SSD} G \) holds, investors with either of these two VNM utility functions must weakly prefer risky asset \( h \) to risky asset \( g \). Since these investors only compare expected wealth, we must have \( E[\tilde{r}_{g}] = E[\tilde{r}_{h}] \), which is (B).

It remains to show that (A) holds. Suppose that (B) holds, but at some \( x \in [0,1] \), \( S(x) > 0 \). Since \( S(\cdot) \) is continuous on \((0,1)\), for each \( y \) contained in a small interval \([x - e, x + e]\), where \( e > 0 \) is very small, it must remain true that \( S(y) > 0 \). Now we can construct a concave VNM utility function that is strictly concave only on the interval \([1 + x - e, 1 + x + e]\) as follows.
\[
u(W) = \begin{cases} 
(1 + x - e)^2 - (1 + x - e)W, & W \leq 1 + x - e; \\
\frac{(1+x-e)^2 - W^2}{2}, & W \in [1 + x - e, 1 + x + e]; \\
\frac{(1+x-e)^2+(1+x+e)^2}{2} - (1 + x + e)W, & W \in (1 + x + e, +\infty). 
\end{cases}
\]
Our proof for the assertion that (i) \( \Rightarrow \) (ii) then implies that
\[ E[u(1 + \tilde{r}_{h})] - E[u(1 + \tilde{r}_{g})] = \int_{0}^{1} S(z)u''(1 + z)dz < 0, \]
and hence for the investor endowed with the above special concave VNM utility function, risky asset $g$ is preferred to risky asset $h$, which is a contradiction to the assumption that $H \geq_{SSD} G$ holds. Thus, when $H \geq_{SSD} G$ holds, (A) and (B) are both necessary.}

\textbf{Proposition 2} If $\tilde{x} \geq_{SSD} \tilde{y}$, then $E[\tilde{x}] = E[\tilde{y}]$ and $\text{var}[\tilde{x}] \leq \text{var}[\tilde{y}]$.

\textit{Proof.} By the preceding theorem, we know that for some $\tilde{e}$ with $E[\tilde{e}|\tilde{x}] = 0$ the two random variables $\tilde{y}$ and $\tilde{x} + \tilde{e}$ have the same distribution function, and hence they have the same variance and expected value also. Now, observe that

$$E[\tilde{y}] = E[\tilde{x}] + E[\tilde{e}] = E[\tilde{x}] + E[E[\tilde{e}|\tilde{x}]] = E[\tilde{x}],$$

where the second equality follows from the \emph{law of iterated expectations}. Hence $E[\tilde{e}] = 0$. Now, observe also that

$$\text{cov}(\tilde{e}, \tilde{x}) = E[\tilde{e}\tilde{x}] - E[\tilde{x}]E[\tilde{e}]$$

$$= E[\tilde{e}\tilde{x}] = E[E[\tilde{e}\tilde{x} | \tilde{x}]] = E[\tilde{x}E[\tilde{e}|\tilde{x}]]$$

$$= E[\tilde{x} \cdot 0] = 0,$$

so that

$$\text{var}[\tilde{y}] = \text{var}[\tilde{x} + \tilde{e}] = \text{var}[\tilde{x}] + \text{var}[\tilde{e}] + 2\text{cov}(\tilde{e}, \tilde{x})$$

$$= \text{var}[\tilde{x}] + \text{var}[\tilde{e}] \geq \text{var}[\tilde{x}],$$

where the inequality follows from $\text{var}[\tilde{e}] \geq 0$. \hfill \|
over $H$ if $H \geq_{SSD} G$. Based on this idea, empirical methods have been developed to investigate whether there coexist two traded assets of which the rates of return satisfy the relation of second-degree stochastic dominance. If such a pair of traded assets can be found, it will be taken as evidence that financial markets are less than fully efficient. Note that second-degree stochastic dominance really does not render the latter implication, for it is only natural that investors also take positions in other risky assets besides the pair of assets under comparison, and in that case an investor may rationally hold the stochastically dominated asset for hedging reasons.

5 Applications of the EU Theory

1. Our first application of the EU theory is to asset pricing. Suppose in a two-period (dates 0 and 1) economy there is one single consumption good (i.e., money) and one single investor with date-0 endowment $e_0$ (her initial wealth) and date-1 endowment $\tilde{e}_1$. Assume that there are

$$E[u(1 + \tilde{r}_g)] = E[u(1 + \tilde{r}_h + \tilde{e})] = E[E[u(1 + \tilde{r}_h + \tilde{e})|\tilde{r}_h]]$$

$$\leq E[E[u(1 + \tilde{r}_h + \tilde{e}|\tilde{r}_h)]] = E[u(1 + \tilde{r}_h)],$$

where the first equality follows from the fact that $E[u(\tilde{w})] = E[u(\tilde{z})]$ whenever $\tilde{w}$ and $\tilde{z}$ are equal in distribution, the second equality follows from the law of iterated expectations, and the inequality follows from Jensen’s inequality for conditional expectations. Note that when the initial wealth is random, the conclusion is no longer true. For example, let $k$ be a positive constant, $\tilde{r}_g = k - \tilde{z}$, $\tilde{r}_h = k$, and let $\tilde{z}$, a fair gamble, be the investor’s random initial wealth. In this case, $H \geq_{SSD} G$ obviously, but the expected utility from taking project G is $u(k)$, and the expected utility from taking project H is $E[u(k + \tilde{z})] < u(E[k + \tilde{z}]) = u(k)$, where the inequality follows from Jensen’s inequality. Thus project G should be chosen over H for the risk-averse investor whose random initial wealth is $\tilde{z}$. The idea is that, a (more) risky project may turn out to be a better instrument for hedging, when a risk-averse investor is already suffering from an endowment risk (we call it a background risk).

$^{34}$This follows from statement (iii) in the preceding theorem. Without loss of generality, assume that the initial wealth is one dollar, and the VNM utility function $u$ is such that $0 > u''$. We have

possible date-1 states, denoted by $\omega_1, \omega_2, \ldots, \omega_n$. Let $\Omega$ be the set containing these $n$ states. The outcome or realization of $\tilde{e}_1$ in state $\omega_i$ is written $e_1(\omega_i)$, or simply $e_{1i}$. At date 0, financial markets open, and assets are available for trading. First we consider an ideal situation where there are $n$ assets traded in the date-0 markets, where one unit of the $j$-th asset will generate 1 dollar in state $\omega_j$ but nothing in other states, and it is called the state-$\omega_j$ Arrow-Debreu security. We say that the date-0 markets are complete because all the $n$ Arrow-Debreu securities are available for trading at date 0. The date-0 price for the state-$\omega_j$ Arrow-Debreu security is denoted by $\phi_j$, or simply $\phi_{j}$. The single investor seeks to maximize the sum of her expected utilities obtained at dates 0 and 1:

$$\max_{c_0, c_1(\omega)} u_0(c_0) + \sum_{j=1}^{n} \pi_j u_1(c_1(\omega_j)),$$

subject to

$$c_0 + \sum_{j=1}^{n} \phi_j c_1(\omega_j) = e_0 + \sum_{j=1}^{n} \phi_j e_1(\omega_j),$$

where (i) for $t = 0, 1$, $u_t(\cdot)$ is the investor’s date-$t$ VNM utility function, with $u_t' > 0 \geq u_t''$; (ii) $\forall j = 1, 2, \ldots, n$, $\pi_j > 0$ is the probability for the event that the date-1 state of the world is $\omega_j$; (iii) $c_0$ is the amount of money that the investor decides to spend at date 0; and (iv) $c_1(\omega_j)$, or simply $c_{1j}$, is the amount of money she decides to spend at date 1 when the date-1 state is $\omega_j$. Note that, since 1 unit of the $j$-th Arrow-Debreu security pays 1 dollar in and only in state $\omega_j$, the investor must hold $c_1(\omega_j)$ units of the $j$-th Arrow-Debreu security from date 0 to date 1 in order to get $c_1(\omega_j)$ dollars at date 1 when state $\omega_j$ occurs. Note also that the investor is a price-taker: she believes that $\{\phi_j; j = 1, 2, \ldots, n\}$ will not be affected by her buying and selling those $n$ securities.

Our task here is to characterize the set of equilibrium prices $\{\phi_j; j = 1, 2, \ldots, n\}$. By equilibrium, we mean competitive equilibrium or Wal-

---

The presence of a complete set of Arrow-Debreu securities allows an investor to transfer purchasing power back and forth between $t = 0$ and $t = 1$, and between any two states at $t = 1$. Let me give an example. Suppose that $n = 2$, and $\phi_1 = 0.8$, $\phi_2 = 0.1$. Suppose that your endowed income is as follows:

$$e_0 = 2000, \; e_1(\omega_1) = 1000, \; e_1(\omega_2) = 100.$$
rasian equilibrium, in which (i) given the prices \( \{ \phi_j; j = 1, 2, \cdots, n \} \), the amounts that the investor chooses to spend at date 0 and date 1, \( \{ c_0, c_1; j = 1, 2, \cdots, n \} \), are expected-utility-maximizing; and (ii) given the prices \( \{ \phi_j; j = 1, 2, \cdots, n \} \), the demands for the \( n \) assets, which are again \( \{ c_0, c_1; j = 1, 2, \cdots, n \} \), equal the supplies of the \( n \) assets;\(^{37}\) that is, the \( n \) asset markets clear under the prices \( \{ \phi_j; j = 1, 2, \cdots, n \} \).

Replacing \( c_0 \) by \( e_0 + \sum_{j=1}^{n} \phi_j [e_1(\omega_j) - c_1(\omega_j)] \) into the objective function \( u_0(c_0) + \sum_{j=1}^{n} \pi_j u_1(c_1(\omega_j)) \), we have a maximization problem with a concave objective function, for which the first-order condition is necessary

(i) Suppose that you only want to consume in state \( \omega_1 \) at \( t = 1 \). How much can you spend in state \( \omega_1 \) at \( t = 1 \)? The answer is

\[
\frac{1}{\phi_1} [e_0 + e_1(\omega_2) \times \phi_2] + e_1(\omega_1).
\]

You have to carry out the following time-0 transactions in order to make this consumption plan possible.

- At first, you have to transfer your time-1 \( \omega_2 \)-state income back to time 0. This can be done by short selling 100 units of the 2nd Arrow-Debreu security: borrowing 100 units of the 2nd Arrow-Debreu security and selling at \( t = 0 \), you will immediately get 100\( \phi_2 \) at \( t = 0 \), but you will have to return 100 in (and only in) state \( \omega_2 \) at \( t = 1 \), which leaves you with \( e_1(\omega_2) - 100 = 0 \) in state \( \omega_2 \) at time 1.

- Next, you can spend \( e_0 + 100\phi_2 \) on the first Arrow-Debreu security. That is, you can purchase \( \frac{e_0 + 100\phi_2}{\phi_1} \) units of the first Arrow-Debreu security at \( t = 0 \), and carry them into time 1. Then, you will have no money left at \( t = 0 \), implying that your consumption at time 0 is zero, and you will also have zero consumption in state \( \omega_2 \) at time 1. However, you can consume more than \( e_1(\omega_1) \) in state \( \omega_1 \) at time 1. Indeed, the total consumption in state \( \omega_1 \) at \( t = 1 \) becomes

\[
\frac{e_0 + 100\phi_2}{\phi_1} + e_1(\omega_1).
\]

Now, as another exercise, determine how much you can spend at time 0, if you only want to consume at time 0. Detail the needed transactions.

\(^{37}\)What is the supply of the \( j \)-th Arrow-Debreu security? The answer is \( e_1(\omega_j) \). To see this, note that there is only one investor in this economy, and being endowed with \( e_1(\omega_j) \) dollars in state \( \omega_j \) at date 1 is the same as holding \( e_1(\omega_j) \) units of the \( j \)-th Arrow-Debreu security at date 0.
and sufficient for the optimal solution:

\[ \frac{\pi_j u'(c_1(\omega_j))}{u'_0(c_0)} = \phi_j, \quad \forall j = 1, 2, \ldots, n. \]

Now, the markets clearing condition requires that

\[ c_0 = e_0, \quad c_1(\omega_j) = e_1(\omega_j), \quad \forall j = 1, 2, \ldots, n. \]

(Since there are no other people in this economy, markets clear if and only if the prices of the \( n \) assets adjust in such a way that the investor finds it optimal to consume her endowments at each date in each state.) It follows that

\[ \frac{\pi_j u'(e_1(\omega_j))}{u'_0(e_0)} = \phi_j, \quad \forall j = 1, 2, \ldots, n. \]

Note that every conceivable asset generating cash flows at date 1 can be represented by a vector \( x \in \mathbb{R}^n \), where \( x_j \), the \( j \)-th element of \( x \), denotes the cash flow generated by one share of asset \( x \) in state \( \omega_j \) at date 1. (We have assumed that investors care about money and nothing else.) Now observe an important fact: every conceivable asset that generates cash flows at date 1 is a portfolio of the \( n \) Arrow-Debreu securities. For example, an asset that pays a per-unit cash flow \( x(\omega_j) \) in state \( \omega_j \) can be regarded as a portfolio that consists of \( x(\omega_1) \) units of the first Arrow-Debreu security, \( x(\omega_2) \) units of the second Arrow-Debreu security, and so on. For example, suppose that \( n = 3 \). Observe that

\[
\begin{bmatrix}
  x(\omega_1) \\
  x(\omega_2) \\
  x(\omega_3)
\end{bmatrix} = x(\omega_1) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x(\omega_2) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x(\omega_3) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.
\]

Now let \( P_x \) denote the price of the asset on the above left-hand side. (The right-hand side is a portfolio of the three traded Arrow-Debreu securities.) Note that in equilibrium \( P_x \) must equal the date-0 value of the portfolio on the above right-hand side: if not, then the investor would still want to alter her equilibrium positions in the \( n \) traded assets because an arbitrage opportunity has shown up, which contradicts the
assumption that the economy has reached an equilibrium. (We shall have more to say on this in Lecture 3.) It follows that

\[ P_x = \sum_{j=1}^{n} \phi_j x(\omega_j) = \sum_{j=1}^{n} \frac{\pi_j u_1'(e_1(\omega_j))}{u'_0(e_0)} x(\omega_j) \]

\[ = E[\frac{u_1'(\tilde{e}_1)}{u'_0(e_0)} \tilde{x}]. \]

Note that the above pricing formula holds for any conceivable asset \( x_{n \times 1} \). In particular, a pure discount bond with face value equal to one dollar is an asset that promises to pay \( x(\omega_j) = 1, \forall j = 1, 2, \cdots, n \). This is obviously a riskless asset. Its date-0 price, according to the above pricing formula, is

\[ E[\frac{u_1'(\tilde{e}_1)}{u'_0(e_0)}]. \]

Now, recall that \( r_f \) is the rate of return on any riskless asset. Since the pure discount bond is a riskless asset, we must have

\[ \frac{1}{1 + r_f} = E[\frac{u_1'(\tilde{e}_1)}{u'_0(e_0)}] \Rightarrow r_f = \frac{1}{E[\frac{u_1'(\tilde{e}_1)}{u'_0(e_0)}]} - 1. \]

Thus, for an asset that pays a per-unit cash flow \( x(\omega_j) \) in state \( \omega_j \), its date-0 price can be further written as

\[ P_x = E[\frac{u_1'(\tilde{e}_1)}{u'_0(e_0)} \tilde{x}] \times \frac{E[\frac{u_1'(\tilde{e}_1)}{u'_0(e_0)}]}{E[\frac{u_1'(\tilde{e}_1)}{u'_0(e_0)}]} \]

\[ = \frac{E[\frac{u_1'(\tilde{e}_1)}{u'_0(e_0)} \tilde{x}]}{1 + r_f} = \frac{E[\tilde{\xi} \tilde{x}]}{1 + r_f}, \]

where the random variable

\[ \tilde{\xi} \equiv \frac{E[\frac{u_1'(\tilde{e}_1)}{u'_0(e_0)}]}{E[\frac{u_1'(\tilde{e}_1)}{u'_0(e_0)}]}, \]

47
Because \( u'_0, u'_1 > 0 \), the realizations of \( \tilde{\xi} \) are all positive; that is, \( \tilde{\xi} \) is a positive random variable.

Recall that for all \( j = 1, 2, \ldots, n \), \( \pi_j \) is the real probability for state \( \omega_j \).

We can define a set of martingale probabilities as follows.

\[
\forall \omega_j \in \Omega, \quad \pi^*_j = \pi_j \xi(\omega_j).
\]

We claim that (A) \( \forall j = 1, 2, \ldots, n, \pi^*_j > 0 \); and (B) \( \sum_{j=1}^{n} \pi^*_j = 1 \). Hence \( \{\pi^*_j; j = 1, 2, \ldots, n\} \) are indeed well-defined probabilities. To see that (A) is true, note that \( \pi^*_j = \pi_j \xi(\omega_j) \), with \( \pi_j > 0 \) and \( \xi(\omega_j) > 0 \). To see that (B) is true, note that

\[
\sum_{j=1}^{n} \pi^*_j = \sum_{j=1}^{n} \pi_j \xi(\omega_j) = \frac{1}{E[\tilde{\xi}]} \sum_{j=1}^{n} \pi_j \frac{u'_1(\epsilon_1(\omega_j))}{u'_0(\epsilon_0)} = \frac{E[\tilde{\xi}]}{E[\tilde{\xi}]} = 1.
\]

Now, recall that for an asset that pays a per-unit cash flow \( x(\omega_j) \) in state \( \omega_j \), its date-0 price can be further written as

\[
P_x = \frac{E[\tilde{\xi}x]}{1 + r_f} = \frac{1}{1 + r_f} E[\tilde{\xi}x] = \frac{1}{1 + r_f} \sum_{j=1}^{n} \pi_j \tilde{\xi}(\omega_j) \hat{x}(\omega_j)
\]

\[
= \frac{1}{1 + r_f} \sum_{j=1}^{n} \pi^*_j \hat{x}(\omega_j) = \frac{1}{1 + r_f} E^*[\hat{x}] = \frac{E^*[\hat{x}]}{1 + r_f},
\]

where \( E^*[\cdot] \) is the expectation that is taken using the new probabilities \( \{\pi^*_j; j = 1, 2, \ldots, n\} \). The formula

\[
P_x = \frac{E^*[\hat{x}]}{1 + r_f}
\]

is called the martingale pricing formula for any asset \( x \).

If you ask someone what would be the fair date-0 price for an asset that pays \( \hat{x} \) at date 1, you would probably get the answer \( E[\hat{x}] \).
theory will tell you that two things may go wrong with this answer: first, one has to pay the price at date 0, and then wait to get the cash flow $\tilde{x}$ at date 1, and there is a time value for money; and second, the price that one pays at date 0 is a sure amount of money, and in return the investor will get an uncertain cash flow $\tilde{x}$ at date 1. The above pricing formula tells us that, to take care of the time value of money (that is, 1 dollar today is usually worth more than 1 dollar tomorrow), the expected date-1 cash flow must be discounted (i.e., divided by $(1 + r_f)$) before it transforms into the date-0 price; and moreover, the original probabilities $\{\pi_j; j = 1, 2, \ldots, n\}$ must be replaced by a set of new probabilities $\{\pi^*_j; j = 1, 2, \ldots, n\}$ in order to reflect the fact that $u''_t \leq 0$ (that is, an investor typically hates uncertainty). In fact, if $u''_t = 0$ so that the investor is risk-neutral and no longer bothered by the presence of uncertainty, then the old and the new probabilities will coincide with each other. In that case, the date-0 price of an asset is equal to its expected date-1 cash flow divided by $(1 + r_f)$. If the investor is risk-neutral and has no time preferences (i.e., she feels indifferent about 1 dollar today and 1 dollar tomorrow), then discounting can also be spared, and the fair date-0 price of asset $\tilde{x}$ for this investor is indeed $E[\tilde{x}]$.

Example 4

Suppose that there are two investors in the two-period frictionless economy, where for $i = 1, 2$, investor $i$ has von Neumann-Morgenstern utility function

$$u(z) = -e^{-\rho_i z}, \quad \forall z \in \mathbb{R},$$

where

$$\rho_2 > \rho_1 > 0.$$ 

There are one risky asset and one riskless asset available for trading at date 0. The riskless rate of return is $r_f$. The risky asset is a common stock, which has 2 shares outstanding, with the two investors each holding one share before trading starts at date 0. Nobody is endowed

\[\text{Example 4}\]

\[\text{Example 4}\]

\[\text{Example 4}\]

\[\text{Example 4}\]
with the riskless asset, so that the riskless asset is in zero net supply. Let \( \tilde{x} \sim N(\mu, \sigma^2) \) be the date-1 cash flow generated by one share of the common stock, where

\[
\mu > \frac{2\sigma^2}{\frac{1}{\rho_1} + \frac{1}{\rho_2}} > 0.
\]

(i) Let \( P \) be the date-0 stock price. Let \( D_i(P) \) be investor \( i \)'s demand for the common stock at date 0, given that the stock price is \( P \). Find \( D_1(\cdot) \) and \( D_2(\cdot) \).

(ii) Write down the markets-clearing condition, and obtain the equilibrium stock price \( P^* \).

(iii) Plug \( P^* \) into \( D_1(\cdot) \) and \( D_2(\cdot) \) and determine which investor is buying the stock, and which investor is selling the stock in equilibrium at date 0.

(iv) Which one between the two investors is borrowing in equilibrium at date 0? Which one is lending? Why?

Solution. Consider part (i). Investor \( i \)'s problem is

\[
\max_{D_i \in \mathbb{R}} E[-e^{-\rho_i D_i \tilde{x}} + (1 - D_i)P(1+rf)] = e^{-\rho_i [(1-D_i)P(1+rf)]} E[-e^{-\rho_i D_i \tilde{x}}].
\]

Note that the random variable

\[-\rho_i D_i \tilde{x} \sim N(-\rho_i D_i \mu, \rho_i^2 D_i^2 \sigma^2),\]

so that, by the formula of moment generating function for a Gaussian random variable,\(^{39}\) we have

\[E[-e^{-\rho_i D_i \tilde{x}}] = e^{-\rho_i D_i \mu + \frac{1}{2} \rho_i^2 D_i^2 \sigma^2}.\]

Thus investor \( i \) seeks to

\[
\max_{D_i \in \mathbb{R}} -e^{-\rho_i [D_i \mu + (1-D_i)P(1+rf)] - \frac{1}{2} \rho_i^2 D_i^2 \sigma^2}
\]

\(^{39}\) It says that for \( \tilde{z} \sim N(e, V) \),

\[
M_{\tilde{z}}(t) \equiv E[e^{\tilde{z}t}] = e^{te + \frac{t^2}{2}V}.
\]
\[
\begin{align*}
&= u(D_i \mu + (1 - D_i) P(1 + r_f) - \frac{1}{2} \rho_i D_i^2 \sigma^2),
\end{align*}
\]
where recall that \( u(\cdot) \) is strictly increasing. Thus in searching for the utility-maximizing \( D_i \), the investor can simply solve the following maximization problem:
\[
\max_{D_i \in \mathbb{R}} W(D_i) \equiv D_i \mu + (1 - D_i) P(1 + r_f) - \frac{1}{2} \rho_i D_i^2 \sigma^2.
\]
In finance literature, \( W(D_i) \) is referred to as the certainty equivalent of \( W \) for the investor, induced by the investment strategy \( D_i \).

The last maximization problem involves only a quadratic objective function, and hence is easy to solve. The necessary and sufficient first-order condition gives
\[
D_i(P) = \frac{\mu - P(1 + r_f)}{\rho_i \sigma^2}, \quad i = 1, 2.
\]
This demand function exhibits several interesting properties. At first, investor \( i \) takes a long position (i.e., \( D_i > 0 \)) if and only if the expected rate of return on the risky asset is greater than \( r_f \), and that position shrinks when the investor becomes more risk-averse (\( \rho_i \) gets higher) or the riskiness of the risky asset rises (\( \sigma^2 \) gets higher). This finishes part (i).

Now, for part (ii), the markets-clearing condition requires that at the equilibrium price \( P^* \),
\[
D_1(P^*) + D_2(P^*) = 2 \Rightarrow P^* = \frac{\mu - \bar{\rho} \sigma^2}{1 + r_f},
\]
where
\[
\bar{\rho} = \frac{2}{\frac{1}{\rho_1} + \frac{1}{\rho_2}}
\]
is the harmonic mean of \( \rho_1 \) and \( \rho_2 \). Note that \( P^* \) equals \( \frac{\mu}{1 + r_f} \) (the martingale pricing formula for risk neutral investors!) if the riskiness of the common stock, measured by \( \sigma^2 \), vanishes, or if \( \bar{\rho} \) vanishes (which happens if either \( \rho_1 \) or \( \rho_2 \) is equal to zero). Thus the term \( \bar{\rho} \sigma^2 \) represents
a risk discount: the two investors are actually risk averse, not risk neutral. Note that our condition

\[ \mu > \frac{2\sigma^2}{\frac{1}{\rho_1} + \frac{1}{\rho_2}} \]

ensures that both investors will take long positions in stock.

Next, consider part (iii). It is easy to see that

\[ D_i(P^*) = \frac{\rho_i}{\lambda_i}, \quad i = 1, 2, \]

and hence

\[ D_1(P^*) > 1 > D_2(P^*). \]

This is a very intuitive result. Investor 2 is more risk averse than investor 1, since \( \rho_2 > \rho_1 \), and hence in equilibrium investor 1 must hold more of the risky common stock than investor 2 does. Note also that this result verifies our earlier conjecture that \( i - D_i(P^*) \neq 0 \) for \( i = 1, 2 \).

For part (iv), recall that the two investors start with the same initial wealth, and so the investor holding more of the risk common stock must also be the one that is borrowing in equilibrium. Hence investor 1 is borrowing and investor 2 is lending in equilibrium.

2. Our second application is to the design of Pareto efficient sharing rules. Consider two partners U and V facing the set \( \Omega \) (defined in our first application above) of date-1 uncertain states. Their endowed date-1 incomes are respectively \( a_j \) and \( b_j \) in state \( \omega_j \), and we define \( C_j = a_j + b_j \) to be the aggregate consumption in state \( \omega_j \) for the two people. Their date-1 VNM utility functions are respectively \( u(\cdot) \) and \( v(\cdot) \), where \( u', v' > 0 \geq u'', v'' \). Our task here is to find the optimal way for the two people to share their aggregate consumption in each and every date-1 state. Here “optimal” means Pareto optimality: an arrangement between the two people is not Pareto optimal or efficient, if there is another arrangement that can make both of them weakly happier, and at least one of them strictly happier. When the latter does not happen, then the original arrangement is Pareto optimal (or Pareto efficient).
Suppose that U and V sign a contract, which specifies that U should get \( s_j \) in state \( \omega_j \). This contract \( \{s_j; j = 1, 2, \cdots, n\} \) is feasible if and only if both U and V are happy to sign it. In other words, it is feasible if and only if

\[
\sum_{j=1}^{n} \pi_j u(s_j) \geq \sum_{j=1}^{n} \pi_j u(a_j) \equiv u_0;
\]

and

\[
\sum_{j=1}^{n} \pi_j v(C_j - s_j) \geq \sum_{j=1}^{n} \pi_j v(b_j) \equiv v_0.
\]

The latter two inequalities are referred to as respectively U’s and V’s \textit{individual rationality constraints}. We call \( u_0 \) and \( v_0 \) respectively U’s and V’s reservation utilities. Our question here is, if there is at least one feasible contract for U and V, which feasible contracts are Pareto efficient? What do they look like?

Suppose that U has the right to \textit{design} the contract; i.e., U has the right to choose \( \{s_j; j = 1, 2, \cdots, n\} \). Then U seeks to

\[
\max_{\{s_j; j = 1, 2, \cdots, n\}} f(s_1, s_2, \cdots, s_n) \equiv \sum_{j=1}^{n} \pi_j u(s_j)
\]

subject to

\[
g(s_1, s_2, \cdots, s_n) \equiv v_0 - \sum_{j=1}^{n} \pi_j v(C_j - s_j) \leq 0.
\]

It can be verified that \( f \) and \( g \) are respectively concave and convex functions of \( (s_1, s_2, \cdots, s_n) \), and hence we can apply the Kuhn-Tucker theorem to solve the problem. Let \( \mu \geq 0 \) be the Lagrange multiplier for the constraint, and the Kuhn-Tucker condition requires that at optimum

\[
\forall j = 1, 2, \cdots, n, \quad \frac{\partial f}{\partial s_j}(s_1, s_2, \cdots, s_n) = \mu \frac{\partial g}{\partial s_j}(s_1, s_2, \cdots, s_n).
\]

Using the definitions of \( f \) and \( g \), we obtain the following \textit{Borch rule} for Pareto optimal risk sharing:

\[
\forall j = 1, 2, \cdots, n, \quad u'(s_j) = \mu v'(C_j - s_j).
\]
Two interesting cases arise: (i) $u'' = 0 > v''$; and (ii) $u'' < 0 = v''$. In case (i), $u'$ is a constant, and the Borch rule implies that $v'(C_j - s_j)$ must be independent of $j$! Since $v'(-)$ is a strictly decreasing function, this can happen only when

$$C_1 - s_1 = C_2 - s_2 = \cdots = C_n - s_n,$$

that is, $V$’s date-1 income becomes riskless under the Pareto optimal risk sharing! The idea is quite simple: if $U$ is risk neutral and $V$ is risk averse, then it is efficient to let $U$ bear all the risk in aggregate consumption. Case (ii) is similar. In fact, given that $v'$ is a constant, $u'(s_j)$ has to be independent of $j$, and since $u'(-)$ is strictly decreasing, this implies that

$$s_1 = s_2 = \cdots = s_n.$$

Again, the risk-averse party must receive a fixed income at date 1!

The above analysis has implications about how two partners of a private firm should share profits. If exactly one of them is risk-averse, the risk-neutral guy should become the shareholder, who issues a riskless debt to her partner. The above analysis also tells us something about the design of an insurance policy. If the insurance company is risk-neutral while the insuree is risk-averse, then there should be no deductibles in an insurance policy: the insuree should have non-random wealth after the date-1 reimbursements are made.\(^{40}\)

**Example 5** Suppose that the above case (i) holds. Suppose further that $U$ is an insurance company, and $V$ is an insuree. Suppose that there are two equally likely date-1 states ($n = 2$), with $b_1 = 0$, $b_2 = 4$, and $v(z) = \sqrt{z}$. Now, an insurance contract is defined as a pair $(p, r)$, where $V$ must pay $U$ $p$ in both states at date 1 ($p$ is the insurance premium), and $U$ will reimburse $r$ to $V$ in and only in state 1. This example shows how $U$’s and $V$’s bargaining power may affect $p$ and $r$.

1. Assume that $U$ can choose $(p, r)$, which $V$ can only accept or reject.

\(^{40}\)Two questions arise at this point. First, why are there always deductibles in the real-world insurance contracts? Second, do the above results depend on who has the right to design $\{s_j; j = 1, 2, \cdots, n\}$?
Compute $p$ and $r$.\(^{41}\)

(2) Assume that $V$ can choose $(p, r)$, which $U$ can only accept or reject. Compute $p$ and $r$.\(^{42}\)

**Theorem 8** Suppose that $u', v' > 0 > u'', v''$. Then the Pareto optimal sharing contract $\{s_j; j = 1, 2, \cdots, n\}$ has the following property: if $C_i > C_j$, then $s_i > s_j$ and $C_i - s_i > C_j - s_j$.

**Proof.** If $C_i > C_j$, then necessarily either $s_i > s_j$ or $C_i - s_i > C_j - s_j$. First consider the case where $s_i > s_j$. By the fact that $u'(\cdot)$ is a strictly decreasing function, we have

$$\mu u'(C_i - s_i) = u'(s_i) < u'(s_j) = \mu u'(C_j - s_j),$$

and hence by the fact that $v'(\cdot)$ is a strictly decreasing function, we have

$$C_i - s_i > C_j - s_j.$$ 

Now it is easy to see that the other case leads to the same conclusion. \(\|\)

So far, we have assumed that the aggregate consumption $\bar{C}$ is a discrete random variable with $n$ possible outcomes. Here we consider the case where it is a continuous random variable with a compact support $[0, 1]$.

---

\(^{41}\)Hint: The optimal $(p, r)$ must make $V$ feel indifferent about buying and not buying the insurance, and it must also satisfy the Borch rule. Hence we have

$$\begin{cases} \frac{1}{2}\sqrt{0 + r - p} + \frac{1}{2}\sqrt{4 - p} = \frac{1}{2}\sqrt{0} + \frac{1}{2}\sqrt{4}, \\
0 + r - p = 4 - p \end{cases} \Rightarrow r = 4, \ p = 3.$$ 

In this case, $U$’s expected profit is $p - \frac{1}{2}r = 1$, and $V$’s expected utility is 1.

\(^{42}\)Hint: The optimal $(p, r)$ must make $U$ feel indifferent about selling and not selling the insurance, and it must also satisfy the Borch rule. Hence we have

$$\begin{cases} \frac{1}{2}(-r + p) + \frac{1}{2} \cdot p = 0, \\\n0 + r - p = 4 - p \end{cases} \Rightarrow r = 4, \ p = 2.$$ 

In this case, $U$’s expected profit is zero, and $V$’s expected utility is $\sqrt{2}$. 

55
Let us replace U and V by agents 1 and 2.\footnote{In the economics literature, an agent is simply a decision-maker.} Agent $j$ has the following CARA VNM utility function
\[ u_j(c) = -e^{-a_j c}, \quad j = 1, 2, \]
where $a_j > 0$ is agent $j$’s Arrow-Pratt measure for absolute risk aversion. A feasible sharing rule consists of two increasing functions $s_1(C)$ and $s_2(C) = C - s_1(C)$. When one agent (say agent 1) has the right to design $s_1(\cdot)$, she seeks to
\[
\max_{s_1(\cdot)} E[-e^{-a_1 s_1(\tilde{C})}]
\]
subject to
\[
E[-e^{-a_2 [C - s_1(\tilde{C})]}] \geq v_0.
\]
A necessary condition can be easily derived. Let $s(\cdot)$ is the optimal solution, and observe that for all $t(\cdot)$, $s(\cdot) + xt(\cdot) = s(\cdot)$ when $x = 0$, so that the following maximization problem has a solution at $x = 0$ given any $t(\cdot)$:
\[
\max_{x \in \mathbb{R}} H(x) \equiv E[-e^{-a_1 [xt(\tilde{C}) + s(\tilde{C})]}]
\]
subject to
\[
G(x) \equiv v_0 - E[-e^{-a_2 [\tilde{C} - xt(\tilde{C}) - s(\tilde{C})]}] \leq 0.
\]
Applying the Kuhn-Tucker theorem, we have, for some $\mu > 0$ and some wisely chosen $t(\cdot)$,
\[
H'(0) + \mu G'(0) = 0 \Rightarrow u_1'(s(C)) = \mu u_2'(C - s(C)),
\]
for all realizations $C$ of $\tilde{C}$. Thus, Borch rule holds again! Now, using the fact that $u_i(z) = e^{-a_i z}$, we can obtain from the above Borch rule
\[
s(C) = \alpha + \beta C,
\]
for some $\alpha \in \mathbb{R}$ and some $\beta \in (0, 1)$. That is, for the two CARA people, the optimal sharing rules are linear. If agent 1 is an employer and agent 2 is working for agent 1, then it is optimal for agent 1 to pay agent 2 a base salary $\alpha$, and give agent 2 a bonus in the form of a fixed proportion $\beta$ of the realized firm earnings $\tilde{C}$.\footnote{In the economics literature, an agent is simply a decision-maker.}
3. Our last application considers a firm’s investment decision.

**Example 6** Suppose that a firm has a shareholder $S$ and a debtholder $D$, where the firm is protected by limited liability. Suppose that the debt will mature at date 1, with face value $x > 0$. Right now, at date 0, the firm has 1-dollar cash at hand, and it can decide whether to invest that dollar in a risky project, which will generate a random cash flow $\tilde{y}$ at date 1. Assume that both $S$ and $D$ are risk-neutral without time preferences, and that $\tilde{y}$ is equally likely to be $\mu + e$ and $\mu - e > 0$, where $\mu, e > 0$.

(i) Suppose $\mu = 1$. If $S$ has the right to make the investment decision, will $S$ choose to spend the dollar on that risky project?

(ii) Suppose $\mu = 1$. If $D$ has the right to make the investment decision, will $D$ choose to spend the dollar on that risky project?\(^{44}\)

Let $S_I$ and $S_N$ denote the shareholder’s expected date-1 cash flow in respectively the events where the risky project is and where the risky project is not adopted. Correspondingly, let $D_I$ and $D_N$ denote the debtholder’s expected date-1 cash flow in respectively the events where the risky project is and where the risky project is not adopted. Then we have, given $\mu = 1$,

- if $x \geq 1 + e$, then $S_I = 0 = S_N$, $D_I = 1 = D_N$;
- if $1 \leq x < 1 + e$, then $S_I = \frac{1 + e - x}{2} > 0 = S_N$, $D_I = \frac{x}{2} + \frac{1 - e}{2} < 1 = D_N$;
- if $1 - e \leq x < 1$, then $S_I = \frac{1}{2}(1 + e - x) \geq 1 - x = S_N$, $D_I = \frac{x}{2} + \frac{1 - e}{2} < x = D_N$;
- if $x < 1 - e$, then $S_I = \frac{1 + e - x}{2} + \frac{1 - e - x}{2} = 1 - x = S_N$, $D_I = x = D_N$.

Hence $S$ tends to take the risky project but $D$ would prefer not to. This happens because the random cash flow $\tilde{y}$ generated by the risky project is second-order stochastically dominated by the sure one dollar (which the firm will obtain if it gives up the risky project), and because the functions $f(\tilde{y}) \equiv \max(\tilde{y} - x, 0)$ and $g(\tilde{y}) \equiv \min(\tilde{y}, x)$ are respectively convex and concave functions.
(iii) Discuss what would happen if $\mu < 1$?45

6 Some Basic Ideas of Prospect Theory

4. Before we end this note, we shall briefly review the main ingredients of prospect theory. Unlike the EU theory, which was derived from a collection of axioms, prospect theory has been developed based on compelling experimental evidence. Consider the following mutually exclusive binary choices, where $(x, p)$ denotes a lottery that yields $x$ dollars with probability $p$ and nothing with probability $1 - p$, and $(z)$ denotes a sure lottery that yields $z$ dollars.

- Problem 1: Choose either
  A: $2,500$ with prob. 0.33, $2,400$ with prob. 0.66, and 0 with prob. 0.01.
  or
  B: $2,400$ for sure.∗ (82%)
- Problem 2: Choose either
  C: $2,500$ with prob. 0.33 and 0 with prob. 0.67∗ (83%)
  or
  D: $2,400$ with prob. 0.34 and 0 with prob. 0.66.
- Problem 3: Choose either
  A: $(4,000, 0.8)$ or B: $(3,000)^*$ (80%).
- Problem 4: Choose either
  C: $(4,000, 0.2)^*$ (65%) or D: $(3,000, 0.25)$.
- Problem 5: Choose either
  A: 0.5 chance to win a 3-week tour to England, France, and Italy.
  or
  B: one-week tour of England for sure.* (78%)

45 Note that in part (i) and part (ii), the firm value remains unchanged whether or not the risky project is taken. The different investment decisions do result in different values for the stock and for the debt, but the sum of the stock value and debt value is always equal to one, as long as investors are risk-neutral without time preferences. Here, things are different: adopting the risky project will result in a lower firm value because of $\mu < 1$. One can verify that S may still have an incentive to adopt the risky project, even if doing so will reduce the sum of the stock value and debt value!
Problem 6: Choose either C: 0.05 chance to win a 3-week tour to England, France, and Italy.∗ (67%)
or
D: 0.1 chance to win a one-week tour to England.

Problem 7: Choose either A: (6000, 0.45) or B: (3000, 0.9)*. (86%)

Problem 8: Choose either C: (6000, 0.01)* (73%) or D: (3000, 0.02).

Problem 9: You will be given 1000000 if you are willing to also accept either
A: (1000000, 0.5) or B: (500000)*.

Problem 10: You will be given 2000000 if you are willing to also accept either C: (-1000000, 0.5)* or (-500000).

Problem 11: Choose either A: (4000, 0.2)* (65%) or B: (3000, 0.25).

Problem 12: A two-stage game proceeds as follows. With prob. 0.75 this game may end at the first stage without payoffs, and with prob. 0.25 this game may reach the second stage, where you get to choose either C: (4000, 0.8) or D: (3000)*. To play this game, you must tell us your choice between C and D right before the game starts at the first stage.

5. The above lotteries were first presented to laboratory subjects by the two psychologists D. Kahneman and A. Tversky that subsequently developed the prospect theory. For each pair of the lotteries listed above, we have attached an asterisk (*) to the one that was preferred by a majority of subjects. Kahneman and Tversky concluded in their 1979 article that most people behaved in a way that was inconsistent with the predictions of the expected utility theory.

At first, the subjects tended to exhibit a preference toward sure lotteries over lotteries that were very likely to succeed, which appears at odds with the expected utility theory. For example, in Problem 1, B was a sure lottery, which was preferred to A by most subjects, but when B became an ingredient of a compound lottery, say D in Problem 2, subjects found it much less appealing than C. This phenomenon violates the independence axiom, and has been referred to as the Allais
paradox. A similar phenomenon appeared in the results of Problem 7 and Problem 8, where it is easy to see that lotteries in Problem 8 were simply compound lotteries consisting of those in Problem 7 and of a zero lottery. Kahneman and Tversky reported that, based on these findings, people tend to focus more on the amounts of cash flows (that is, 6000 and 3000 in Problems 7 and 8) and less on the probabilities of success when all the lotteries under consideration involve very small probabilities of success; and people also have a tendency of exaggerating the differences in the probability of success when the lotteries under consideration are either sure to succeed or very likely to succeed (this latter tendency is referred to as the certainty effect). These findings suggest that, if a subject’s preference for lotteries can be represented by some utility function, then that utility function cannot be linear in probabilities. (Recall how we derived an EU function from a preference in Theorem 1; proving that $H(p)$ is linear in $p$ is a key step.) Kahneman and Tversky therefore suggested that we should replace $p$ in the EU theory by a non-linear transform of $p$, denoted $w(p)$, and they referred to the latter as the weight function; see below.

Second, Kahneman and Tversky found that most subjects appeared to be risk seeking when facing losses (cash outflows). Note that, in terms of the terminal wealth, A in Problem 9 is identical to C in Problem 10, and B in Problem 9 is identical to D in Problem 10. Moreover, A (and hence C) is simply a lottery that consists of B (and correspondingly D) plus a fair gamble. Yet, most subjects exhibit a preference reversal when making choice between A and B, and between C and D. Kahneman and Tversky referred to the pattern that people tend to be risk averse about gains and risk seeking about losses as the reflection effect, which indicates that the utility function is convex and concave on respectively the negative and the positive orthants, making the origin a reflection point for the utility function (which will be termed a Kahneman-Tversky value function; see below). The above-mentioned preference reversal is taken by expected utility theorists to be evidence of human irrationality, because the subjects that exhibit preference reversal were apparently fooled by the different presentations of the same pair of lotteries. It seems to suggest that by framing or re-structuring the cash flows, one can manipulate people’s perception
and induce them to make totally opposite consumption or investment decisions. This implies that by carefully creating the *framing effect* for consumers or investors, firms or financial institutions can make more profits.

The third important argument of Kahnemand and Tversky’s prospect theory is that people tend to *edit* the lotteries under consideration before making choices, and the most important consequence of this editing is that common aspects of the lotteries under consideration would be deleted and disregarded. Recall that in Problem 9, regardless of the subject’s choice between A and B, the subject would get the extra 1000000, and hence getting the 1000000 was a common aspect of A and B. Facing such a common aspect, the subjects tended to totally disregard it, and to focus only on the differences. Similarly, in Problem 10, the 2000000 was a common aspect. The experimental findings show that by wisely creating a common aspect (framing!), one can manipulate a subject’s choice, knowing that the subject will first edit the lotteries and deleted common aspects. In fact, this observation also helps explain the Allais paradox. Note that in terms of the terminal wealth, A (respectively B) in Problem 11 is identical to C (respectively D) in Problem 12, and yet most subjects exhibited a preference reversal between Problem 11 and Problem 12. A close look into the sequential game in Problem 12 reveals that C and D have a common aspect, that the game might end in the first stage with probability 0.75. After deleting this common aspect from C and D, the certainty effect became important, which may have contributed to the observed preference reversal. At this point, we note that the same editing and deleting procedure did not seem to occur for one-stage lotteries. Why did the subjects not disregard the outcome 2400 (which might occur with prob. 0.66 no matter whether the subject chose A or B) when they compare A to B in Problem 1? Indeed, if they did, then their choices between A and B would have been consistent with their choices between C and D. Kahneman and Tversky referred to the pattern that people would first edit the lotteries and remove the common aspects before making choices as the *isolation effect*. The *certainty effect*, the *reflection effect*, and the *isolation effect* constitute the three cornerstones of the first generation of prospect theory proposed in Kahneman and Tversky.
(1979). Since then, the theory has been upgraded for several times, and it won a Nobel prize for the authors about a decade ago.

Two instruments, the *value function* and the *weight function*, are then proposed by the authors to further substantiate this theory. The value function is proposed to replace the von Neumann-Morgenstern utility function defined for terminal wealth. The argument in the value function is not the terminal wealth; rather, it is the difference between the realized terminal wealth and some reference wealth level (referred to as a reference point), and this difference is a gain (a loss) if it is positive (negative). The value function $v(x)$ is increasing, passing through the origin, and is concave on the positive orthant and convex on the negative orthant, which captures the aforementioned *reflection effect*. Moreover, $v'(x) \leq v'(-x)$ for all $x > 0$, a property termed *loss aversion*, which says that, roughly, the decision maker is risk averse around zero.

The weight function is proposed to replace the original probability appearing in an expected utility function. The weight function is an increasing function of the original probability. Although the weight function is increasing, for an event with probability close to zero, the weight is bounded below away from zero, and for an almost sure event, the weight is bounded above away from one (so that the weight function is discontinuous at 0 and 1). In plain words, the weight function says that a sure event is considered much more important than an almost sure event, a property that is consistent with the aforementioned *certainty effect*.

6. Prospect theory is merely one competing theory to EU theory. There are other preference theories developed in the past few decades that intend to generate investment behavior and asset price patterns that can fit empirical evidence better than EU theory does. In view of the developments of new theories, some defenses for the expected utility theory have also been made. For example, Machina (1982) has advanced a theory of local expected utility theory, which spares the independence axiom and is able to reproduce the main predictions of the expected utility theory (such as Theorems 6 and 7).

Some financial economists have questioned the relevancy of the exper-
imental findings presented by prospect theorists. They argued that if an investor may easily be fooled because of the framing effect, then he tends to lose money quickly in trading assets and securities. If irrational investors keep transferring wealth to rational investors, then in the end their behavior will cease to affect the financial markets equilibrium. That is, the equilibrium prices and volume of financial assets will eventually be determined by the decision-making behavior of rational investors (whose behavior is consistent with EU theory). This point of view, unfortunately, has been challenged in a series of new research, which shows that, contrary to the above argument, it is the irrational noise traders that may eventually dominate the financial markets when rational traders and irrational traders coexist in the markets. This happens because the irrational traders may be more willing to take positions in high-risk high-return assets, and their aggressive trading behavior may result in too much noise-trader risk that scares away rational risk-averse investors. Eventually, the proportion of aggregate wealth held by noise traders may even tend to one, and it is the rational traders who are forced out of the markets; see for example DeLong, Shleifer, Summers, and Waldman (1989, 1990a, 1990b, 1991). To learn more about the early developments in behavioral finance, see Thaler (eds.) (1993).

References


