Advanced Microeconomics (ECON 401)
Lecture 4
Choice Under Uncertainty

Teng Wah Leo

Our previous discussions on consumer choice are extensions from your intermediate Microeconomics, and remain couched in events that occur with certainty. We will now extend those ideas when consumer choices are over objects or events whose occurance are uncertain. Each event and their associated outcome occur with an objectively known probability, and the distribution is thus over all possible outcomes. These risky alternative outcomes are termed *lotteries*. We maintain the assumption of rationality for individuals, albeit now over lotteries, so that we will examine how individuals make their choice through maximizing their expected utility, in the process obtaining the *expected utility theorem*. You will realize that the key assumption that permits this analysis is the *independence axiom*.

We will then extend the discussion to *risk aversion* and their measurement, and also examine comparisons between alternative distributions of monetary returns associated with different lotteries.

1 Expected Utility Theory

Consider an individual who faces a choice over several risky alternatives, where each alternative may generate several outcomes, and where the realization of any outcome is uncertain when the choice needs to be made. Denote the set of all possible outcomes as $\mathcal{C}$. Let this number of outcomes be finite, indexing them by $k = 1, \ldots, K$.

Definition 1 *A simple lottery* $\mathbf{p}$ *is a list* $\mathbf{p} = [p_1, \ldots, p_K]'$ *where* $p_k \geq 0\ \forall k = 1, \ldots, K$*, and* $\sum_k p_k = 1$*, so that* p_k*'s are the probabilities of outcomes* k *occurring.*

This simple lottery should be distinguished from a *compound lottery* where the outcomes are themselves simple lotteries. In that case, you could think of the k indices as different states of the world, and given that k^{th} state of the world, you have N different possible outcomes.

Definition 2 *Given N simple lotteries $\mathbf{p}^n = [p_1^n, \ldots, p_K^n]'$ for $n = 1, \ldots, N$, with probabilities $q_n \geq 0$ with $\sum_n q_n = 1$, the compound lottery $(\mathbf{p}^1, \ldots, \mathbf{p}^N; q_1, \ldots, q_N)$ is the uncertain realization that gives you the simple lottery $\mathbf{p}^n$ with probability q_n for $n = 1, \ldots, N$.*

Of course such a compound lottery can be simplified into a *reduced lottery* where the state of the world k for $k = 1, \ldots, K$, occurs with probabiliity $p_k = \sum_n p_k^n q_n$. In turn, the reduced lottery $\mathbf{p} = \sum_n q_n \mathbf{p}^n$

We can now discuss the individual's preference over the lotteries. To do so, we assume that what matters is the reduced lotteries, since it is the final outcomes that are of consequence. First, define the set $\mathcal{L}$ as the set of all possible simple lotteries over the set of outcomes $\mathcal{C}$. Second, define a complete and transitive rational preference relation $\succsim$ on $\mathcal{L}$, so that these simple lotteries can be compared (it is in situations under uncertainty, and when the underlying alternatives become more complex, that our rationality assumption might become a stretch). Before we can begin our analysis, we need two additional assumptions.

Axiom 1 *The preference relation $\succsim$ on $\mathcal{L}$ is continuous if for $\mathbf{p}, \mathbf{p}', \mathbf{p}'' \in \mathcal{L}$, the sets*

$$\{\alpha \in [0,1] : \alpha \mathbf{p} + (1-\alpha)\mathbf{p}' \succsim \mathbf{p}''\} \subset [0,1]$$

and

$$\{\alpha \in [0,1] : \alpha \mathbf{p} + (1-\alpha)\mathbf{p}' \precsim \mathbf{p}''\} \subset [0,1]$$

are closed.

This last assumption essentially eliminates the possibility of minute changes in probability that could alter the preference orderings between two simple lotteries. As you would recall from the case of certainty, the continuity assumption iplies the the existence of a utility function that represents $\succsim$, a function $U : \mathcal{L} \to \mathbb{R}$, such that $\mathbf{p} \succsim \mathbf{p}' \Leftrightarrow U(\mathbf{p}) \geq U(\mathbf{p}')$. In addition, we also need the following controversial axiom so as to give $U(.)$ more "structure".

Axiom 2 *The $\succsim$ on $\mathcal{L}$ satisfies the independence axiom if* $\forall\ \mathbf{p}, \mathbf{p}', \mathbf{p}'' \in \mathcal{L}$ *and* $\alpha \in (0,1)$, *we have*

$$\mathbf{p} \succsim \mathbf{p}' \ \ \leftrightarrow \ \ \alpha\mathbf{p} + (1-\alpha)\mathbf{p}'' \succsim \alpha\mathbf{p}' + (1-\alpha)\mathbf{p}''$$

In other words, the addition of an additional alternative lottery will not affect the individual's preference ordering.

With this, we can now specify the form of U.

Definition 3 *The utility function $U : \mathcal{L} \to \mathbb{R}$ has the expected utility form, also known as the von Neumann-Morgenstern expected utility function, if there is an assignment of values $(u_1, \ldots, u_K)$ to the K outcomes associated with every $\mathbf{p} \in \mathcal{L}$ such that*

$$U(\mathbf{p}) = \sum_{k=1}^{K} u_k p_k$$

This leads on to a useful insight.

Proposition 3 *A utility function $U : \mathcal{L} \to \mathbb{R}$ has a expected utility form $\Leftrightarrow$ it is linear,*

$$U\left(\sum_{n=1}^{N} q_n \mathbf{p}^n\right) = \sum_{n=1}^{N} q_n U(\mathbf{p}^n)$$

for N lotteries $\mathbf{p}^n \in \mathcal{L}$, $n = 1, \ldots, N$, and $q_n \geq 0$, $\sum_n q_n = 1$.

Proof. Suppose $U(.)$ satisfies the linearity condition. Then we know we can write $\mathbf{p} = [p_1, \ldots, p_K]$ as a convex combination of degenerate lotteries $(\mathbf{p}_1, \ldots, \mathbf{p}_K)$, in other words, $\mathbf{p} = \sum_k p_k \mathbf{p}_k$. Then,

$$\begin{aligned} U(\mathbf{p}) &= U\left(\sum_k p_k \mathbf{p}_k\right) \\ &= \sum_k p_k U(\mathbf{p}_k) \\ &= \sum_k p_k u_k \end{aligned}$$

On the other hand, suppose $U(.)$ has the expected utility form, and a compound lottery

$(\mathbf{p}^1, \ldots, \mathbf{p}^N; q_1, \ldots, q_N)$ as defined prior, so that the reduced lottery is $\mathbf{p} = \sum_n q_n \mathbf{p}^n$.

$$\begin{aligned} U\left(\sum_n q_n \mathbf{p}^n\right) &= \sum_k u_k \left(\sum_n p_k^n q_n\right) \\ &= \sum_n q_n \left(\sum_k u_k p_k^n\right) \\ &= \sum_n q_n U(\mathbf{p}^n) \end{aligned}$$

■ Note that this expected utility property is a cardinal property with respect to the underlying utility functions, themselves defined on the space of lotteries. The following says that the expected utility form is preserved only by increasing linear transformations.

Proposition 4 *Define* $U : \mathcal{L} \to \mathbb{R}$ *as a von Neumann-Morgenstern utility function for* $\succsim$ *on* $\mathcal{L}$*. Then* $\widetilde{U} : \mathcal{L} \to \mathbb{R}$ *is another von Neumann-Morgenstern utility function for* $\succsim$ $\Leftrightarrow$ $\widetilde{U}(\mathbf{p}) = \alpha U(\mathbf{p}) + \beta$*, for* $\alpha > 0$*,* $\forall \beta$*, and* $\mathbf{p} \in \mathcal{L}$*.*

The implication of the above is that the differences between utility between outcomes have meanings. To see this, consider four outcomes, 1 to 4. If we say the difference in utility between outcomes 1 and 2, is larger than 3 and 4, then $u_1 - u_2 > u_3 - u_4$, for $u_i > 0$, so that $\frac{2}{3}u_1 + \frac{1}{3}u_4 > \frac{2}{3}u_2 + \frac{1}{3}u_3$, and we would prefer $\mathbf{p} = [2/3, 0, 0, 1/3]'$ to $\mathbf{p}' = [0, 2/3, 1/3, 0]'$.

Further, this also means that if *succsim* on $\mathcal{L}$ is representable by a von Neumann-Morgenstern utility, and since this function is linear, and thus continuous, then $\succsim$ is continuous on $\mathcal{L}$. And importantly, $\succsim$ must also satisfy the independence axiom.

The bottomline of the expected utility theorem is that if an individual's preferences over lotteries satisfy continuity and the independence axiom, then this preference is representable by a von Neumann-Morgenstern utility function.

Theorem 5 ***Expected Utility Theorem*** *Let* $\succsim$ *on* $\mathcal{L}$ *satistfy the continuity and independence axioms* $\Rightarrow$ $\succsim$ *can be represented by the expected utility form, in that each outcome may be assigned a value* u_k*,* $k = 1, \ldots, K$ *such that for two lotteries,* $\mathbf{p}$ *and* $\mathbf{p}'$*, we have*

$$\mathbf{p} \succsim \mathbf{p}' \;\;\Leftrightarrow\;\; \sum_{k=1}^{K} p_k u_k \geq \sum_{k=1}^{K} p_k' u_k \tag{1}$$

Although we will not be disccusing the proof due to its length, you may find it in the suggested references. There several interesting intuitive insights to this theorem nonetheless. Consider a three outcome lottery, so that the space of the lottery can now be represented by a $2-$simplex (triangle), so that indifference curves here are linear, since the expected utility form is linear in probability. Further, these linear utility will be parallel to each other. As it turns out, this is a consequence of the independence axiom.

Why can't the indifference curve be a curve. Consider two lotteries $\mathbf{p}$ and $\mathbf{p}'$ on the $2-$simplex. Suppose $\mathbf{p} \sim \mathbf{p}'$, then by linearity we have $\alpha\mathbf{p} + (1-\alpha)\mathbf{p}' \sim \mathbf{p}\ \forall \alpha \in [0, 1]$. Suppose not, and the indifference curve is depicted as a curve, then you might have the situation that $\alpha\mathbf{p} + (1-\alpha)\mathbf{p}' \succ \mathbf{p} = \alpha\mathbf{p} + (1-\alpha)\mathbf{p}$, so that you get a contradiction of $\mathbf{p}' \succ \mathbf{p}$.

Similarly, maintaining $\mathbf{p} \sim \mathbf{p}'$, and add the third lottery $\mathbf{p}''$. Suppose a lowest lottery being $\alpha\mathbf{p} + (1-\alpha)\mathbf{p}'$, so that $\mathbf{p}$ and $\mathbf{p}'$ are on the same indifference curve. If the indifference curves are non-parallel, for instance $\beta\mathbf{p} + (1-\beta)\mathbf{p}''$ and $\beta\mathbf{p}' + (1-\beta)\mathbf{p}''$ are on different indifference curves, and without loss of generality, suppose the latter is on a higher indifference curves than the former, you'd arrive at a contradiction that $\mathbf{p} \prec \mathbf{p}'$.

1.1 Some Caveats

There are nonetheless issues with this model of consumer behavior under uncertainty that continues to be an engaging field of research. We will discuss some of their issues here.

To begin with, let us realize what the advantages are in using the expected utility theorem. Firstly, it is convenient analystically, particularly with some prior statistical training, its use would seem perfectly logical, explaining its ubiquity. Secondly, since most individuals find judgements between riskly alternatives are difficulty, this provides reasonable guidance towards "optimal" action(s), particularly if the individual finds *the independence axiom* reasonable. However, that is not to say it works for all possible individuals. The fact we are even talking about caveats highlights how pervasive violations of expected utility theorem is!

Example 1 ***The Allais Paradox.*** *This paradox was first discovered by Allais (1953), and is a thought experiment. Consider three monetary alternatives:*

1^{st} ***Prize***	2^{nd} ***Prize***	3^{rd} ***Prize***
$\$2,500,000$	$\$500,000$	0

Suppose the individual has two choice tests: First is a choice between:

$$\mathbf{p}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{p}_2 = \begin{bmatrix} 0.1 \\ 0.89 \\ 0.01 \end{bmatrix}$$

Second, the choice is between:

$$\mathbf{p}_3 = \begin{bmatrix} 0 \\ 0.11 \\ 0.89 \end{bmatrix} \quad \mathbf{p}_4 = \begin{bmatrix} 0.1 \\ 0 \\ 0.9 \end{bmatrix}$$

It turns out that it is common that individuals exhibit the following preference ranking; $\mathbf{p}_1 \succ \mathbf{p}_2$ *and* $\mathbf{p}_4 \succ \mathbf{p}_3$*. And this actually violates expected utility theorem!*

To see that, the first choice states that the individual prefers $500,000 with certainty versus the lottery with a chance of a windfall. The second choice states that getting the windfall with probability 0.1 is better than getting $500,000 with a probability 0.11. (Note that in the first choice, $\mathbf{p}_2$ *has a expected winnings of $695,000. However, in the second, the expected winnings respectively are $55,000 and $250,000.) The violation is easier observed on a* $2-$*simplex.*

More precisely, denote the utility you obtain from each outcome as u_25, u_5 *and* u_0*. Then,*

$$\begin{aligned}
\mathbf{p}_1 &\succ \mathbf{p}_2 \\
\Rightarrow u_5 &> 0.1u_25 + 0.89u_5 + 0.01u_0 \\
\Rightarrow u_5 + (0.89u_0 - 0.89u_5) &> 0.1u_25 + 0.89u_5 + 0.01u_0 + (0.89u_0 - 0.89u_5) \\
\Rightarrow 0.11u_5 + 0.89u_0 &> 0.1u_25 + 0.9u_0 \\
\Rightarrow \mathbf{p}_3 &\succ \mathbf{p}_4
\end{aligned}$$

so that you have a contradiction! There are four main replies to this anomaly. Firstly, it has been noted that choosing between uncertain risky alternatives is a normative activity, so that mistakes such as these once noted can and should be corrected. Secondly, some has argued that the example has little application in economics since the alternatives that are extraordinary with probabilities close to the bounds of 0 and 1. But yet if you think about it, the example is closer to reality, and if we think of all the empirical financial outcomes, such are the scenarios often predicted. Thirdly, some has elaborated on the theory to *include considerations of what could have been forgone in individual choice, leading to Regret Theory. Finally, to completely forgo the independence axiom.*

Example 2 ***Machina's Paradox.*** *Consider three possible versions of an "adventure" offerred to you: "a trip to the Galápagos Islands" (G_1), "watching the premier of a documentary on the Galápagos Islands" (G_2), and "staying at home" (G_3). And it would not be a stretch to say your preference rankings would be $G_1 \succ G_2 \succ G_3$. Consider two possible lotteries. Let $\mathbf{p}_1 = (0.999, 0.001, 0)$, and $\mathbf{p}_2 = (0.999, 0, 0.001)$. Since you have noted that $G_2 \succ G_1$, the independence axiom says that $\mathbf{p}_1 \succ \mathbf{p}_2$. Yet, for most individuals, $\mathbf{p}_2$ is reasonable if they anticipate event of either G_2 or G_3, since if you envisage losing, that you might wish to reevaluate events G_2 and G_3, since losing involves significant disappointment, and watching the documentary just "rubs salt onto the wound". This is not that dissimilar to the "regret" arguement in* ***Allais Paradox****, and highlights the "psychological" impression the notion of "what could have been", thereby casting a shadow over the* ***independence axiom****.*

In addition, it should be noted that no all situations involving uncertainty is amenable to the application of expected utility theorem, particularly when the outcomes under uncertainty are affected by the actions of the individual. In that situation, we have the issue of *induced preferences*, and the effect that the action has needs to be incorporated into the model. To put it more succinctly, "*how we rank our preferences should not dependent on our ex ante actions* ". This is a particularly important caveat, since upon deeper consideration, you would realize that most outcomes under uncertainty, are affected by our actions, and consequently our preferences over the outcomes!

2 Risk Aversion

We will now focus on pecuniary/monetary outcomes, thereby begin first to consider the ideas of *risk aversion* and their measurement and implications. Assume money to be a continuous variable $x \in \mathcal{X} \subset\in \mathbb{R}^+ + \{0\}$, then the monetary lottery can be describe by a distribution function $F : \mathbb{R} + \{0\} \to [0, 1]$, and where the density function is $f(.)$, and $F(x) = \int_{\underline{x}}^{x} f(t)dt\ \forall x$.

Consideration using the distribution preserves the linear structure of lotteries we have discussed till now. To see this consider the compound lottery $(\mathbf{p^1}, \ldots, \mathbf{p^N}; \mathbf{q_1}, \ldots, \mathbf{q_N})$

with each lottery $\mathbf{p}_n$ having its associated distribution F_n, so that the final distribution of the compound lottery is just $F(x) = \sum_n q_n F_n(x)$.

To refine the description of the space, let it be such that $\mathcal{L}$ be the set of all distribution functions over nonnegative sums of money. The individual's preference relation, $\succsim$, is now then defined over $\mathcal{L}$. In the current context of expected utility function, it can now be written as,

$$U(F) \quad = \quad \int u(x)dF(x) \tag{2}$$

so that the assignment of utility is over the distribution function, so that any distribution function can be evaluated by a utility function. Further, observe that the utility function continues to be linear in $F(.)$. The import of this approach is that the evaluations of utility now is sensitive to variations in mean and all relevant moments between differing lotteries. Finally, we will from here on refer to $U(.)$ as the *von Neumann-Morgenstern utility function*, while $u(.)$ which is for the actual outcome realization is refered to as the *Bernoulli utility function.* This is a particularly important distinction, since the benefits to the expected utility approach is borne out through specifications of $u(.)$ that allows us to capture facets of economic attributes of individual choices, specifically to postulate the $u(.)$ is increasing and continuous in x (exception at $x = 0$ where $u(0) = -\infty$). Indeed at sensitivity to higher moments, such as when we extend towards discussions of stochastic dominance, we would add curvature conditions on $u(.)$.

Since *risk aversion* is pervasive in much economic discussion, we will define it here first.

Definition 4 *An individual is risk averse if for any lottery $F(.)$, the degenerate lottery that yields $\int xdF(x)$ with certainty is at least as good a a lottery $F(.)$.*

On the other hand, if the individual is indifferent between these two lotteries, then she is risk neutral.

Finally, if the indifference hold only when the two lotteries are the same, that is $F(.)$ is likewise degenerate, then she is strictly risk averse

Then risk aversion may be stated succinctly as an individual who abides by the following inequality $\Leftrightarrow$,

$$\int u(x)dF(x) \quad \leq \quad u\left(\int xdF(x)\right) \quad \forall F(.) \tag{3}$$

The above is also known as the *Jensen's inequality*, and essentially defines concavity of $u(.)$, so that *risk aversion is equivalent to the concavity of* $u(.)$. This in turn highlights that a risk neutral individual has a linear $u(.)$. For the rest of our discussions, the following concepts are useful.

Definition 5 *For a Bernoulli utility function* $u(.)$*,*

1. *The* ***certainty equivalent*** *of* $F(.)$*, denoted* $c(F, u)$*, is the amount of money that would make an individual indifferent between a lottery* $F(.)$ *and recieving with certainty* $c(F, u)$*, in other words,*

$$u(c(F, u)) = \int u(x)dF(x) \tag{4}$$

2. *For* x *and* $\epsilon > 0$*, the probability premium,* $\pi(x, \epsilon, u)$ *is the excess in winning probability over fair odds that makes the individual indifferent between the certain outcome* x *and a gamble betqween outcomes* $x + \epsilon$ *and* $x - \epsilon$*. That is,*

$$u(x) = \left[\frac{1}{2} + \pi(x, \epsilon, u)\right] u(x + \epsilon) + \left[\frac{1}{2} - \pi(x, \epsilon, u)\right] u(x - \epsilon) \tag{5}$$

This then leads to the following important equivalence relationships. For an individual who maximizes based on the expected utility, the following statements are equivalent.

1. The individual is *risk averse.*

2. The Bernoulli utility function is concave.

3. $c(F, u) \leq \int xdF(x)\ \forall F(.)$

4. $\pi(x, \epsilon, u) \geq 0\ \forall x, \epsilon > 0$.

It may be useful to have a measure for risk aversion, so that should you be an asset manager, for instance, you could make comparisons between your clients, in your attempt to meet their needs.

Definition 6 *For a twice continuously differentiable Bernoulli utility function* $u(.)$*, the* ***Arrow-Pratt coefficient of absolute risk aversion*** *at some value* x *is* $r_A(x) = -\frac{u''(x)}{u(x)}$

The intuition of the measure lies in the fact that risk neutrality occurs when the Bernoulli utility function is linear, which implies that $u'' = 0$. So that an increment away from that is an increase in risk aversion. It turns out that with this, we can compare the degree of risk aversion across individuals, particularly since the idea of an utility function is a subjective artifact.

Proposition 6 *The following definitions of "**more risk averse than**" relation are equivalent.*

1. $r_A(x, u_2) \geq r_A(x, u_1)\ \forall x$
2. *There exists an increasing concave function* $\phi(.)$ *such that* $u_2(x) = \phi(u_1(x))\ \forall x$.
3. $c(F, u_2) \leq c(F, u_1)\ \forall F(.)$.
4. $\pi(x, \epsilon, u_2) \geq \pi(x, \epsilon, u_1)\ \forall x$ *and* $\epsilon > 0$.

However, any such rankings using the r_A measure is only a *partial ordering*, since although the rankings are transitive, there are not complete since it is usually the case that the inequality may switch under other realizations of x. Nonetheless, r_A has the following interesting property that seem to be mirrored in reality.

Definition 7 *The Bernoulli utility function* $u(.)$ *exhibits decreasing absolute risk aversion if* $r_A(x, u)$ *is a decreasing function of* x.

In other words, risk taking behavior increases as income increases. r_A however is a weak measure, and is complemented by the stronger assumption of *nonincreasing relative risk aversion*. So whereas r_A is useful when comparisons are between absolute gains/losses, when the comparisons are between proportions/percentages, the concept of *relative risk aversion* is relevant.

Definition 8 *For a Bernoulli utility function* $u(.)$*, the* ***coefficient of relative risk aversion*** *at* x *is* $r_R = \frac{xu''(x)}{u'(x)}$.

To see this derivation from r_A, consider $\alpha > 0$, and defining $\widetilde{u}(\alpha) = u(\alpha x)$, so that at $\alpha = 1$ we have the original income position. Then differentiating the relationship with respect to α, we have the relationship $\frac{\widetilde{u}''(\alpha)}{\widetilde{u}'(\alpha)} = \frac{xu''(x)}{u'(x)}$.

Then taken together, for nonincreasing r_R , this would imply that individuals become less risk averse for gambles that are proportional to her wealth (gambles specified as proportions/percentages to her wealth) as her wealth rises. This is a stronger assumption because an individual exhibiting nonincreasing relative risk aversion will necessarily demonstrate decreasing absolute risk aversion, but the reverse needn't be true.

3 Stochastic Dominance

Where in the previous section, comparisons were made between the Bernoulli utility functions, we will examine lotteries more directly (less subjectively. It is noted here as "less subjectively" as opposed to "objectively" because even (prior) probabilities can be subjective, dependent on the individual's perception of the lottery.) by comparising the distributions of the payoffs/realizations. This idea of using distribution is borne of the idea that for most individuals, lotteries can be compared based either on their level of returns, or their variation. This gives rise to two concepts, first when a distribution F yields unambiguously higher returns that G (*first order stochastic dominance*), and when F offers a less risky gamble than G (*second order stochastic dominance*).

Definition 9 *$F(.)$ first order stochastically dominates $G(.)$, $F \succ_1 G$, if, for every nondecreasing function $u : \mathbb{R} \to \mathbb{R}$,*

$$\int u(x)dF(x) \geq \int u(x)dG(x)$$

Proposition 7 *$F \succ_1 G \Leftrightarrow F(x) \leq G(x)\ \forall x$.*

Proof. Let $x \in \mathcal{X} \subset \mathbb{R}$. Denote $H(x) = F(x) - G(x)$, and suppose that for some $x \leq \overline{x}$, $H(\overline{x}) > 0$. Then constructing a nondecreasing function u such that $u(x) = 1$ for $x > \overline{x}$ and $u(x) = 0$ for $x \leq \overline{x}$. Thus we have,

$$\begin{aligned}
\int u(x)dH(x) &= \int_{x\leq\overline{x}} u(x)dH(x) + \int_{x>\overline{x}} u(x)dH(x) \\
&= \int_{x>\overline{x}} u(x)dH(x) \\
&= \int_{x>\overline{x}} [F(x) - G(x)]dx \\
&= [(1 - F(\overline{x})) - (1 - G(\overline{x}))] = -H(\overline{x})
\end{aligned}$$

So that $\int u(x)dH(x) > 0 \Leftrightarrow H(\overline{x}) \leq 0$, and the "only if" part of the proposition follows.

For a twice differentiable $u(x)$ for $x \in \mathcal{X}$, realize that

$$\int_{\mathcal{X}} u(x)dH(x) \quad = \quad [u(x)H(x)]_{\mathcal{X}} - \int_{\mathcal{X}} u'(x)H(x)dx$$

Next observe that the first term on the right hand side is zero, so that for $\int_{\mathcal{X}} u(x)dH(x) \geq 0$, we would need $\int_{\mathcal{X}} u'(x)H(x)dx < 0$. Since $u'(x) \geq 0$, the expected value is greater than zero if $H(x) \leq 0$, and the "if" part of the proof follows. ■

Although this is a powerful insight, there are at least two important caveats. Firstly, it should be noted that the inequality is not strict. Secondly, just because the mean from one distribution is greater than the other does not mean that it first order dominates the other distribution. In deed, this concept highlights how "powerless" comparisons of means are.

What happens when first order stochastic dominance does not hold? We can examine whether second order stochastic dominance holds.

Definition 10 *Consider two distributions, $F(x)$ and $G(x)$ with the same mean. Then $F \succ_2 G$, or interpreted as F being less risky than G, if for every nondecreasing concave function $u : \mathbb{R}^+ \to \mathbb{R}$, we have,*

$$\int u(x)dF(x) \quad \geq \quad \int u(x)dG(x)$$

To see what this entails, lets define the following,

$$\begin{aligned} F^k(x) &= \int_{x \in \mathcal{X}} F^{k-1}(x)dx \\ F(x) &= \int_{x \in \mathcal{X}} f(x)dx \end{aligned}$$

And define the k^{th} derivative of the utility function as $u^{(k)} \equiv u^{(k)}(x)$. Then observe that through integrating by parts,

$$\begin{aligned} \int_{\mathcal{X}} u(x)dF(x) &= u(x)F(x)|_{\mathcal{X}} - \int_{\mathcal{X}} u^{(1)}F(x)dx \\ &= u(x)F(x)|_{\mathcal{X}} - \left[u^{(1)}(x)F^2(x)\big|_{\mathcal{X}} - \int_{\mathcal{X}} u^{(2)}F^2(x)dx \right] \\ &= \sum_{k=1}^{K} (-1)^{k-1} u^{(k-1)}(x)F^k(x) \Bigg|_{\mathcal{X}} + \int_{\mathcal{X}} (-1)^K u^{(K)} F^K(x)dx \end{aligned}$$

So that for the distributional difference, we can write,

$$\int_{\mathcal{X}} u(x)dH(x) \quad = \quad \sum_{k=1}^{K} (-1)^{k-1} u^{(k-1)}(x)H^k(x) \Bigg|_{\mathcal{X}} + \int_{\mathcal{X}} (-1)^K u^{(K)} H^K(x)dx$$

where,

$$
\begin{aligned}
F^k(x) - G^k(x) = H^k(x) &= \int_{x \in \mathcal{X}} H^{k-1}(x) dx \\
&= \int_{x \in \mathcal{X}} \left[F^{k-1}(x) - G^{k-1}(x) \right] dx
\end{aligned}
$$

From the above and based on our prior analysis, the first order stochastic dominnance is reliant on the sign of $-\int_{\mathcal{X}} u^{(1)} H(x) dx$. Extending that, using the expanded formula, you can see that for a second order stochastic dominance, $-u^{(1)} H^2(x) dx \geq 0$ and $\int_{\mathcal{X}} u^{(2)} H^2(x) dx \geq 0$. There are several key points to note here. This leads to the following proposition.

Proposition 8 *For two distributions F and G with the same mean, the following statements are equivalent.*

1. *$F \succ_2 G$.*
2. *G os a mean-preserving spread of F.*
3. *$H^2(x) \leq 0 \Rightarrow F^2(x) \leq G^2(x)$.*

The second point essentially means that you can reconstruct F from G by reducing its variance, and it is in that sense that you can think of G as a riskier lottery. This idea that we are now considering risk aversion can be seem from the fact that the we have added the additional constraint of the Bernoulli utility function being concave, $u^{(2)} \leq 0$, besides $u^{(1)} \geq 0$, thereby ensuring $\int_{\mathcal{X}} u^{(2)} H^2(x) dx \geq 0$. Indeed, we may extend these ideas to higher order stochastic dominance concerns, and with each increment in order, we need additional constraints on the higher order derivatives of the Bernoulli utility function, and the higher order integral of the distribution function. This theory has been applied to both Portfolio selection in the finance literature, and poverty measurement and comparison in the welfare literature.