ECO 317 – Economics of Uncertainty – Fall Term 2009
Notes for lectures
3. Risk Aversion

## Reminders

On the space of lotteries $L$ that offer a finite number of consequences $(C_1, C_2, \ldots C_n)$ with probabilities $(p_1, p_2, \ldots p_n)$, we established the existence of a utility function $u(L)$ such that:

[1] it represents the preferences, that is, it has the property that for any two lotteries $L^a$ and $L^b$,

$$u(L^a) > u(L^b) \quad \text{if and only if} \quad L^a \succ L^b$$

[2] has the expected utility property

$$EU(L) \equiv u(L) = \sum_{i=1}^{n} p_i \; u(C_i)$$

The proof was "constructive". On the bounded (and closed, if you are a mathematician) set of lotteries we let $\overline{L}$ be the best and $\underline{L}$ the worst. Then we showed the existence of a unique $p$ such that $L$ is indifferent to the lottery that yields $\overline{L}$ with probability $p$ and $\underline{L}$ with probability $(1-p)$. (This is often written for short as the lottery $p\,\overline{L} + (1-p)\,\underline{L}$; this is convenient but it should be understood that no sum in any usual number or vector sense is intended.) Then we simply defined $u(L) = p$, and verified that it had the two desired properties.

The same preferences can be represented by another utility function $\tilde{u}$ which also has the expected utility property if (and only if, although we did not prove this) $\tilde{u}$ is an increasing linear (a pedantic mathematician would say "affine") transform of $u$: there are constants $a$, $b$ with $b > 0$ such that $\tilde{u}(c) = a + b\,u(c)$ for all $c$.

Since each consequence $C_i$ is a degenerate lottery that yields this consequence with probability 1 and all other consequences $C_j$ with zero probabilities, the construction automatically gives a utility function $u(C_i)$ over consequences. We can think of the utility of one consequence, $u(C_i)$, as the utility of a degenerate lottery that yields $C_i$ with probability 1 and any other consequence $C_j$ with probability zero. We call this the von Neumann-Morgenstern utility function, to distinguish it from the expected utility function for a non-degenerate lottery. (A pedantic mathematician would create different symbols for the two.)

Many of our applications will be expressed in terms of actions $a$, possible states of the world $s$, and consequence functions $c = F(a, s)$. We can convert our theory of preferences over lotteries easily to this context by writing expected utility of an action as the expectation of the random variable namely the utilities of all possible consequences it might yield in different states of the world:

$$EU(a) = \sum_{j=1}^{m} Pr(s_j)\; u(F(a, s_j)\,)$$

## Risk aversion

In consumer theory without uncertainty, if $c$ is a positive scalar magnitude like money income or wealth or consumption, the utility functions $c$, $c^2$, $\sqrt{c}$, $e^c$, and $\ln(c)$ would all represent the same preferences (all reflecting the trivial property that more is better). But as components of expected utility, these are different. For example, if there are just two consequences with probabilities $p_1$, $p_2$, the three expected utility functions

$$p_1 \ c_1 + p_2 \ c_2, \qquad p_1 \ \ln(c_1) + p_2 \ \ln(c_2), \qquad \text{and} \qquad p_1 \ (c_1)^2 + p_2 \ (c_2)^2$$

represent very different preferences. (Just sketch indifference curves in $(c_1, c_2)$ space.) Specifically, they represent preferences with very different attitudes toward risk. We now develop this idea.

We will usually take the consequences $c$ to be monetary magnitudes such as income or wealth. If the underlying preferences are defined over quantities of goods, then we can work in terms of the indirect utility function of income or wealth, so long as the relative prices are constant or are not the focus of the analysis.

With this convention, if preferences can be represented by expected utility where the utility-of-consequences function is linear, so we can take $u(c) = c$ up to an increasing linear transformation, that means the decision-maker is indifferent between two alternatives that yield equal expected income or wealth, $\sum_i p_i \, c_i$, regardless of the variance or any other measure of dispersion of the distribution over consequences. In other words, this would be a risk-neutral decision-maker. But this is an exceptional case, and raises difficulties, one of which was the starting-point of this whole subject. So we begin there.

## St. Petersburg Paradox

The development of probability theory in the $17^{\text{th}}$ and $18^{\text{th}}$ centuries came from certain observations of gamblers (especially French aristocrats). The expected utility theory of choice under risk has the same origin. A friend of Nicholas Bernoulli proposed to him the following question: "Consider a lottery that works as follows. A fair coin is tossed until it comes heads up. If this requires $n$ tosses, you are paid $2^n$ ducats. How much would you be willing to pay to enter such a lottery?" The event that heads show up for the first time on the $n^{\text{th}}$ toss is $2^{-n}$. Therefore the expected monetary value of the lottery is

$$\sum_{n=1}^{\infty} 2^{-n} \ 2^n = \sum_{n=1}^{\infty} 1 = \infty$$

But no one seems willing to pay any very large sums, let alone unbounded sums, to play this game. This is the St. Petersburg Paradox. (For most of the $20^{\text{th}}$ century it was renamed the Leningrad Paradox :-) .)

Nicholas' brother Daniel Bernoulli offered the following resolution. "People's perceptions of money are logarithmic. Therefore the log of the value they place on the game equals

$$\sum_{n=1}^{\infty} 2^{-n} \ \ln(2^n) \ = \ \sum_{n=1}^{\infty} 2^{-n} \ n \ \ln(2) = \frac{\ln(2)}{2} \sum_{n=1}^{\infty} n \ 2^{-(n-1)}$$

$$= \frac{\ln(2)}{2} \left[ \sum_{n=1}^{\infty} 2^{-(n-1)} \right]^2 = \frac{\ln(2)}{2} \, 2^2 = 2 \, \ln(2) = \ln(2^2) = \ln(4)$$

So people would value the game at only 4 ducats."

We don't need to take Daniel Bernoulli's argument about logarithmic perceptions seriously, even though it may have some basis in psychology. We can instead regard this as an example of a general idea: the logarithm is just one of many possible utility-of-consequences functions.

But problems remain. First, the argument ignores the initial wealth a person brings to the game. (Without some such wealth, how would he pay any entry fee anyway?) If $W_0$ is initial wealth, then maximum entry fee that Bernoulli would be willing to pay to enter the St. Petersburg lottery is given by the $X$ that solves the equation

$$\ln(W_0) = \sum_{n=1}^{\infty} 2^{-n} \, \ln(W_0 - X + 2^n)$$

This is like the "compensating variation" in ECO 310 – it is the change in money income that compensates for, or cancels out, the effect of the lottery and leaves Bernoulli at the same level of utility as before. We could instead look for the "equivalent variation," namely the sure amount of money that would give Bernoulli the same utility as the expected utility he would get when given a gift of the lottery. Then we want the $Y$ that solves the equation

$$\ln(W_0 + Y) = \sum_{n=1}^{\infty} 2^{-n} \, \ln(W_0 + 2^n)$$

If you have elementary programming skills, try these out for a few values of $W_0$.

More importantly, Bernoulli's resolution of the paradox is unsatisfactory in a more fundamental way. Even with a logarithmic utility-of-consequences function, the paradox can be reconstructed by changing the reward if heads show up first on the $n^{\text{th}}$ toss from $2^n$ to $R_n = \exp(2^n)$. Then the utilities of consequences are $u(R_n) = \ln(\exp(2^n)) = 2^n$, and now the expected utility is infinite. But most people still would not be willing to pay very large sums up front for this prospect. The only sure way to avoid the paradox in this framework is to have a utility-of-consequences function that is bounded above, but that can create other problems. More realistically, perhaps people just don't believe that the prizes will actually be paid out if a large value of $n$ is realized, and such disbelief is justified since the prizes soon start to exceed the GDP of the US or of the whole world.

## Risk Aversion and Concavity of Utility

The general idea is that differently nonlinear Bernoulli utility (of consequences) functions yield expected utility that capture different attitudes toward risk.

Continue to work with scalar consequences, typically income or wealth (but could be the quantity of just one good that is the focus of the analysis). Denote them by $C$. Compare two situations: [1] $L_0$, which gives you $C_0$ for sure, and [2] $L$, which gives you $C_1 = C_0 + k$

and $C_2 = C_0 - k$ with probabilities $\frac{1}{2}$ each. Suppose the utility-of-consequences function $u$ is concave. Then, as is evident from Figure 1,

$$\tfrac{1}{2}\ u(C_1) + \tfrac{1}{2}\ u(C_2) < u(C_0) \tag{1}$$

In terms of expected utilities, this becomes $EU(L) < EU(L_0)$. (To be mathematically rigorous, I should define a concept of "strict concavity" that will yield strict inequalities and ordinary or weak concavity that will yield only weak inequalities, but I will leave this out; the textbook is likewise sloppy.)

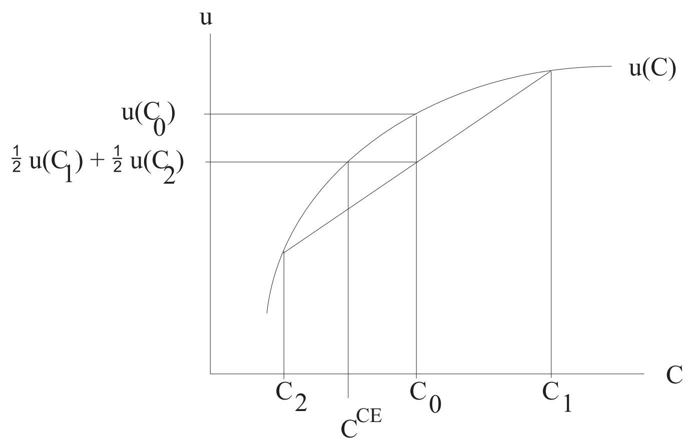


Figure 1: Concave Bernoulli utility implies risk-aversion

The figure also shows that there is a $C^{CE} < C_0$ (but $C^{CE} > C_2$) such that

$$\tfrac{1}{2}\ u(C_1) + \tfrac{1}{2}\ u(C_2) = u(C^{CE})$$

so the decision maker with this utility function is indifferent between the random prospect $L$ and the sure prospect $C^{CE}$. Then we call $C^{CE}$ the certainty-equivalent of $L$. Suppose this person starts at $C_0$ but is then confronted with the prospect of gaining or losing $k$ with equal probabilities $\frac{1}{2}$ each. We can think of $C_0 - C^{CE}$ as the highest insurance premium he is willing to pay to avoid the risk; it is called the risk premium for the decision maker in this initial situation facing this random prospect.

There are three different ways of characterizing a concave function:

[1] $u''(C) < 0$,

[2] $u'(C)$ is a decreasing function of $C$, and

[3] For any $C_1$, $C_2$ and any $p \in (0,1)$, $p\ u(C_1) + (1-p)\ u(C_2) < u(p\,C_1 + (1-p)\,C_2)$.

They have different degrees of generality: the first requires $u$ to be twice differentiable, the second requires it to be only once-differentiable, and the third does not require any differentiability at all, for example it works for piecewise linear functions. For most of our uses, we will work with twice-differentiable functions, so this distinction is not so important and we can use the three characterizations indifferently. There is another point where care is necessary: the inequalities in the above definitions can be weak, in which case we will speak of a “weakly concave” function, and distinguish it from a “strictly concave” function where the inequalities are strict. An expected-utility maximizer with a weakly concave utility-of-consequences function may be indifferent to some fair gambles, but can never be an actual risk-lover. This distinction too should be kept in mind, but for most of our uses it will not play a role.

## Jensen’s Inequality

Returning to the idea that the concavity of the $u$ function implies risk-aversion, we can develop it more generally. Consider a lottery whose outcomes are a random variable $c$ with a given distribution. Then $u(c)$ is another random variable. We show that

$$\text{if } u \text{ is concave, then} \quad \mathrm{E}[u(c)] < u(\mathrm{E}[c])\,, \tag{2}$$

that is, the expected utility of the lottery is less than the sure utility one would get from having the monetary expected outcome with certainty. (2) is called Jensen’s Inequality. Know and remember it well; it is used all the time in the economics of uncertainty. The (1) we started with is a special case with just two possible outcomes of the random $c$.

To prove it, write $\overline{c} = \mathrm{E}[c]$ for brevity, and use Taylor’s theorem with remainder to write

$$u(c) = u(\overline{c}) + (c - \overline{c})\; u'(\overline{c}) + \tfrac{1}{2}\; (c - \overline{c})^2\; u''(\widetilde{c})$$

for some $\widetilde{c}$ lying between $c$ and $\overline{c}$. Since $u$ is assumed to be concave, $u''(\widetilde{c}) < 0$ and so

$$u(c) < u(\overline{c}) + (c - \overline{c})\; u'(\overline{c})$$

Taking expectations and using $\mathrm{E}[c - \overline{c}] = \mathrm{E}[c] - \overline{c} = 0$, we get the result.

If you use the above no-derivatives formulation [3] of concavity, and extend it by induction to the case of $n$ values and then by taking limits to more general random variables, that also proves Jensen’s Inequality.

## Quantitative Measures of Risk Aversion

Consider a decision-maker with initial position (income or wealth) $C_0$, presented with a risk whose expected value is zero. That is to say, his final position is a random variable $C_0 + X$, where $X$ has zero expectation: $E[X] = 0$. For example, $X$ may take on values $X_i$ for $i = 1$, 2, …, $n$, with probabilities $p_i$, and such that $\sum_{i=1}^{n} p_i\, X_i = 0$. This person’s certainty equivalent is defined as the solution $C^{CE}$ to the equation

$$u(C^{CE}) = \sum_{i=1}^{n}\; p_i\, u(C_0 + X_i)\,, \tag{3}$$

and then the risk premium is $\Pi \equiv C_0 - C^{CE}$. For some functional forms of $u$, it may be possible to solve the equation explicitly. But most of the time we have to rely on numerical solution. Alternatively, for small risks we can use a Taylor approximation to get an expression that is due to Pratt and Arrow; it is called the Pratt or Arrow-Pratt approximation. Here is a heuristic derivation; a more rigorous justification requires more calculus but yields no better understanding.

Write (3) as

$$u(C_0 - \Pi) = \sum_{i=1}^{n} p_i \, u(C_0 + X_i) \, .$$

Expand both sides in Taylor series around $C_0$:

$$\begin{aligned}
u(C_0) - \Pi \; u'(C_0) + \ldots \;&=\; \sum_{i=1}^{n} p_i \; \Big[ \, u(C_0) + X_i \; u'(C_0) + \tfrac{1}{2} \; (X_i)^2 \; u''(C_0) + \ldots \Big] \\
&=\; u(C_0) \; \sum_{i=1}^{n} p_i + u'(C_0) \; \sum_{i=1}^{n} p_i \, X_i + \tfrac{1}{2} \, u''(C_0) \; \sum_{i=1}^{n} p_i \, (X_i)^2 + \ldots \\
&=\; u(C_0) \; * \; 1 + u'(C_0) \; * \; 0 + \tfrac{1}{2} \, u''(C_0) \; \mathrm{Var}[X] + \ldots \\
&=\; u(C_0) + \tfrac{1}{2} \, u''(C_0) \; \mathrm{Var}[X] + \ldots \, ,
\end{aligned}$$

where ... indicates higher-order (smaller for small risks) terms in the expansion.

Canceling $u(C_0)$ from both sides, and equating the leading terms that remain on each side (rigorous proof of the validity of this step is what requires more math), we have

$$-\Pi \; u'(C_0) = \tfrac{1}{2} \, u''(C_0) \; \mathrm{Var}[X] \, ,$$

or

$$\Pi = \tfrac{1}{2} \; \mathrm{Var}[X] \; \frac{- \; u''(C_0)}{u'(C_0)} \, .$$

This is quite intuitive: the risk premium is proportional to the variance of the random component of $C$, which can be thought of as a measure of the magnitude of the risk, and also proportional to a measure of the extent of concavity or curvature of the utility-of-consequences function. The latter factor then serves as a measure of the decision-maker's risk aversion for small risks around $C_0$:

$$A(C_0) = \frac{- \; u''(C_0)}{u'(C_0)} \, . \tag{4}$$

This is called the *coefficient of absolute risk aversion* to distinguish it from the *coefficient of relative risk aversion*, where the latter pertains to risks that are expressed as a proportion of the initial position, and the risk premium is likewise expressed as a proportion. Consider a decision-maker who starts from the position $C_0$ and faces the risk of going to $C_0 \, (1 + \widehat{X})$ where $\widehat{X}$ is a random variable with zero mean, for example taking on values $\widehat{X}_i$ with probabilities $p_i$ and satisfying $\sum_i \, p_i \, \widehat{X}_i = 0$. The relative risk premium, call it $\widehat{\Pi}$, solves the equation

$$u(C_0 \, (1 - \widehat{\Pi}) \, ) = \sum_{i=1}^{n} p_i \, u(C_0 \, (1 + \widehat{X}_i) \, ) \, .$$

We can obtain an Arrow-Pratt measure of relative risk aversion by calculations very similar to those above for the absolute case; I highly recommend that you try this yourself to improve your understanding and skills. The resulting measure is

$$R(C_0) = \frac{-\,C_0\; u''(C_0)}{U'(C_0)}\,. \tag{5}$$

This is just the elasticity with which marginal utility declines.

Here are some examples: [1] If

$$u(C) = 1 - \exp(-\,a\,C) \tag{6}$$

where $a$ is a positive constant, then

$$u'(C) = a\;\exp(-\,a\,C), \qquad u''(C) = -\;a^2\;\exp(-\,a\,C), \qquad A(C) = \frac{-\;u''(C)}{u'(C)} = a$$

so this function has constant absolute risk aversion, and $a$ is the parameter that measures this risk aversion. By reversing the steps, that is, by solving the differential equation $-\,u''(C)/u'(C) = a$, we can show that all utility-of-consequences functions with constant absolute risk aversion belong to the family (6), of course within an increasing linear transformation.

[2] Next, consider

$$u(C) = \begin{cases} \frac{1}{1-r}\;C^{1-r} & \text{for } r > 0,\ r \neq 1 \\ \ln(C) & \text{for } r = 1 \end{cases} \tag{7}$$

For this family of functions,

$$u'(C) = C^{-r}, \qquad u''(C) = -\;r\;C^{-r-1}, \qquad R(C) = \frac{-\;C\;u''(C)}{u'(C)} = r;$$

so this is the family of utility-of-consequence functions with constant relative risk-aversion. This is the family that was used, and the implied $r$ calculated, from your answers about risk premia in Question 4 of the questionnaire on the first day of class.

Observe that if $r > 1$, then the values of $u$ are negative. For example, if $r = 2$, we have $u(C) = -\,1/C$. That is OK; the utility function is arbitrary to within origin and scale, so the sign of utility itself is immaterial. What matters is that utility be increasing and concave. And that is true. We have $u'(C) = 1/C^2 > 0$, which is a positive and decreasing function. So $u$ starts out very negative, then gradually increases and flattens out, asymptoting to the horizontal axis as $C \to \infty$.

Similarly in the constant absolute risk aversion form (6), we could have left out the 1 and written utility as simply $-\;\exp(-\,a\,C)$.

Exercise: Try out the St. Petersburg Paradox example for a few utility functions other than the logarithmic.

## More on Cardinal v. Ordinal Utility

Suppose the utility-of-consequences function is logarithmic, so the expected utility of a lottery $L = (C_1, C_2; p_1, p_2)$ is

$$EU(L) = p_1 \ \ln(C_1) + p_2 \ \ln(C_2) \,.$$

Then

$$e^{EU(L)} = (C_1)^{p_1} \ (C_2)^{p_2} \,,$$

the famous Cobb-Douglas utility function from ECO 310. And for any two lotteries $L$, $L'$,

$$L \succ L' \quad \text{if and only if} \quad EU(L) > EU(L') \quad \text{if and only if} \quad e^{EU(L)} > e^{EU(L')} \,,$$

so we could have used the Cobb-Douglas utility function equally well to represent the same preferences. Only it would not have the "expected utility" form – it would not be the mathematical expectation of anything. (This approach is useful when comparing random prospects with common probabilities but different monetary consequences, as is done when we consider bets of different magnitude, or insurance coverage with different sizes of deductible, coinsurance, and indemnity, for a given event. We will discuss such applications in a couple of weeks.)

More generally, we can take any nonlinear increasing transformations of the whole of expected utility to represent the same preferences. In this sense we could say that utility is still ordinal. But the transforms do not have the expected utility form. If we want to preserve that, the only kind of transforms permissible are increasing linear (or pedantically, affine) transforms of the underlying utility-of-consequences function. So the utility-of-consequences function has to be cardinal if we want the expected utility form of the overall objective that is maximized.

## Convex and Mixed Utility Functions

We saw that a concave utility-of-consequences function yields risk-averse behavior. Conversely, a person with a convex utility-of-consequences function $u$ likes risk. In general, the function $u$ does not have to be either concave over its whole domain or convex over its whole domain. It can be concave over one part and convex over some other part of the domain. A person with such a function will then be averse to gambles confined to the concave part, and will like gambles confined to the convex part; his attitude toward gambles that span the two parts will depend on the exact details.

How might such behavior arise? Consider a person choosing to gamble with his job or investments or whatever. If the outcome is above a certain level $C_L$, he keeps what he has. But if the outcome falls below $C_L$, a social or governmental safety net prevents the person's consumption from falling below $C_L$. Even if the person is risk-averse, his utility-of-consequences function over his own outcomes $C$ will be as shown in Figure 2. Then it is easy to find initial situations $C_0$ that are only slightly above $C_L$, and substantial risks, such that the person will take that risk hoping to enjoy the upside and knowing that the safety net means that there is little effective downside. This can happen even when the risk is statistically unfavorable in the sense that the expected value of the money amount is negative. Similarly, if a firm is close to bankruptcy, owners or managers with substantial equity stakes may find it attractive to take excessive risks, knowing that the downside will fall on other lenders or bondholders: for the equity holder, going bankrupt for a million dollars is no worse than going bankrupt for a thousand.

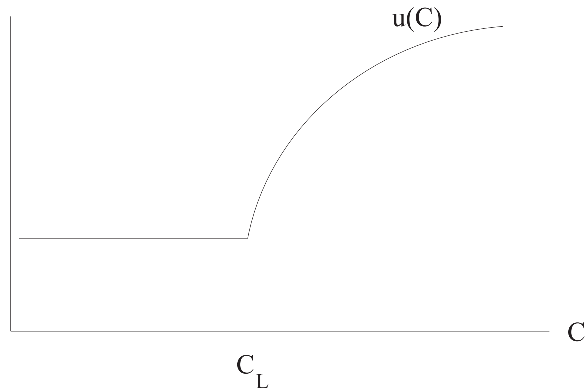


Figure 2: Utility of consequences with safety net or bankruptcy

While a convex utility-of-consequences function (at least over some intervals of wealth or income) can explain gambling behavior in an expected utility framework, probably a more realistic explanation is that people enjoy the act of gambling for its own sake, as a form of entertainment. However, this goes against a basic assumption underlying expected utility theory, namely the assumption that compound lotteries can be collapsed into their simple form without affecting preferences. If you enjoy gambling, you care about the process by which uncertainty is resolved. This violates the compound lottery axiom. (There is some very recent work that studies some consequences of such behavior.)

Many other kinds of utility-of-consequences functions have been suggested. Friedman and Savage argued that there would be risk aversion for very low and very high levels of wealth, and a middle range of risk preference. This is somewhat the opposite of the safety net idea described above.

Kahneman and Tversky proposed the "prospect theory" of choice under risk, in which the status quo plays a special role. The utility-of-consequences function is generally concave for gains and convex for losses, with a discontinuity of slope (kink) at the status-quo point. Such people would be especially averse to small gambles around the status quo, but may like some larger gambles. Figure 3 shows these examples. Rabin studied some consequences of this kink, and we will consider this in more detail later.

The utility-of-consequences function may be different in different states of the world, if something about the state (for example the state of your health, or whether your favorite sports team wins) makes you value the consequences (for example money) differently. Suppose in state $s_j$ for $j = 1, 2, \ldots m$, the utility-of-consequences function is $u_j(c)$. Then the

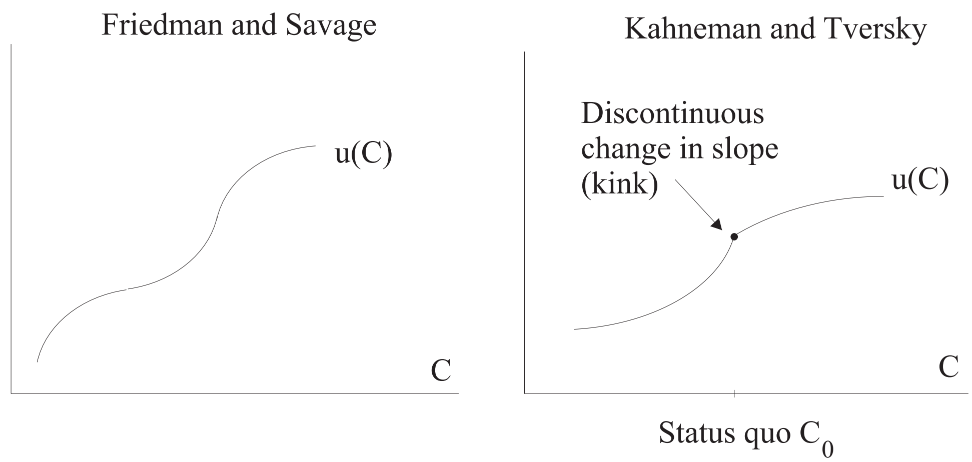


Figure 3: Some other possibilities

corresponding expected utility function of actions is

$$EU(a) = \sum_{j=1}^{m} Pr(s_j)\ u_j(F(a, s_j))$$

where $c = F(a, s)$ is the consequence function. The exact form of a state-dependent utility is very specific to its context, so it is not possible to develop any useful general theory of this. But we will use it from time to time in problems and applications.