0. PREFACE

This thesis is a study in the economics of uncertainty. The literature in this field has grown so rapidly that even a survey of the field as a whole would require more space than is available here. Nevertheless, I have aimed at a kind of completeness. My object has been to present an integrated development from basic notions of choice and uncertainty to theoretical and policy applications. My central claim is that Expected Utility theory has been superseded by more general models which retain its desirable properties such as transitivity and preservation of dominance while being consistent with behavior which is proscribed by Expected Utility theory but frequently observed in practice. One such general model, Anticipated Utility theory, is developed in detail in the thesis.

The thesis draws heavily on my own previously published work in the economics of uncertainty, but also makes a number of new contributions. First, there a number of new theoretical results, notably in Chapter 5, but also in the sections on risk-preference and regret theory. Second, the economic applications in Chapter 6 are completely new. Third, the agricultural policy applications in Chapter 7 have been reformulated in an Anticipated Utility framework.

The attempt to integrate a wide range of previous work has not been without difficulties. In particular, there has been the problem of notation. In general, I have sought internal consistency rather than the maintenance of the notation used in the original publication. However, where ambiguity seems unlikely, I have used certain symbols for different purposes in different parts of the thesis. For example, W and Y are used in section 7.6 to denote wool and yarn respectively, whereas in the rest of the thesis they are spaces of outcomes and of prospects over outcomes. As regards orthography, I have made my own compromise between American and English spelling conventions. I have generally followed English conventions, but have replaced "our" endings with "or" in line with the American practice.

I would like to thank my supervisors, Jock Anderson, Tony Chisholm and John Dillon as much for encouraging me to undertake this project in the first place as for the help and encouragement they have given me during its progress. Despite initial misgivings, I have found that the discipline of organising a fairly disparate body of work into a coherent whole is worthwhile in itself as well as being a great stimulus to creative thought. I would also like to thank my fellow-workers in the field of generalised expected utility theory. Having initially developed my ideas in isolation, I have found exposure to their ideas, and feedback on my own, both stimulating and encouraging.

1. UNCERTAINTY

1.0 Introduction

The object of this thesis is to present and develop a framework for the analysis of economic choice under uncertainty. Before examining choice under uncertainty, it is necessary to consider the nature of uncertainty itself, and to develop some tools for analysis. In this Chapter, some basic issues concerning uncertainty are discussed.

The term 'uncertainty' is used in the most general possible sense while, the terms 'instability', 'risk' and 'ambiguity' are used to cover specific aspects of uncertainty. Instability is examined in section 1.2. Essentially, it refers to objective variation in variables of interest. The subjective notion of risk, referring specifically to unpredictable future occurrences is examined in section 1.3. It was at one time common in the literature to draw a distinction between 'risk' and 'uncertainty', following Knight (1921). In this thesis, the term ambiguity (Ellsberg 1961) is used to refer to Knightian 'uncertainty'. Ambiguity is discussed in section 1.4, and an attempt to draw together the notions of instability, risk and uncertainty is made in section 1.5.

1.1 Fundamental sources of uncertainty

Uncertainty pervades all aspects of life. Before examining economic theories relating to uncertainty it is worth giving some consideration to the fundamental sources of uncertainty. One way to begin this analysis is with the philosophical debate over free will and determinism. One side of this debate holds that determinism is inconsistent with human freedom of action and that the world must therefore be pervaded by uncertainty in a fundamental sense. An important modern representative of this viewpoint is Koestler (1965). At the other extreme are those, such as Skinner (1971), who see no problem in exorcising consciousness altogether from science. An intermediate viewpoint suggests that behaviour may be causally determined without ceasing to be "freely chosen" . This viewpoint is clearly more comfortable for practitioners of an economic mode of reasoning involving attempts to predict human behaviour while stressing the notion of choice.

A second point of importance is the denial by many modern physicists that the world is, in fact, deterministic in nature. Concepts of determinism, powerfully bolstered in the 18th and 19th Century by the Newtonian model of a "clockwork" universe, have been undermined by the discoveries of modern particle physics. In the view of theorists such as Heisenberg (1930) events at the subatomic level are fundamentally uncertain (the dissenting view, attributed to Einstein, that "God does not play dice with the Universe" should be noted). Despite the profound philosophical implications of the Heisenberg Uncertainty Principle, statistical mechanics provides a straightforward reconciliation with Newtonian determinism at the macroscopic level of everyday life. Thanks to the Law of Large Numbers, it is "almost" certain that the average behavior of the huge number of particles making up say, a desk, will be that predicted by Newtonian mechanics. Nevertheless, these modern discoveries have cast doubt on the confident assertion that mental phenomena can ultimately be reduced to the deterministic product of chemical and physical processes.

For economists, however, the fundamental source of uncertainty must be sought not in physical or philosophical concepts, but in human ignorance. Uncertainty relating to events that are clearly determinate (is there a pool of oil underneath this well?) can be just as important as that relating to events that may be uncertain in some fundamental sense (will the government decide to raise taxes?). Indeed, in many important cases of uncertainty, some participants in a given economic setting may be uncertain with respect to a given variable, while others are fully informed. Uncertainty is essentially a subjective state of mind rather than an objective property of the world.

In order to deal with uncertainty, it is necessary to examine the ways in which people collect, process and use information in decision-making. This task may be commenced by considering a number of different aspects of the phenomenon of uncertainty.

1.2 Instability and variability

Uncertainty is essentially subjective in nature. However, it has an objective counterpart in the form of instability or variability in the values taken by a particular variable over time and space. To the extent that uncertainty can be modelled in terms of random drawings from a population, instability or variability is a necessary condition for the existence of uncertainty. For example, universal physical constants such as the mass of the hydrogen atom cannot constitute a source of uncertainty for decision-makers, although if these constants took different values the implications would be profound. Instability in a given variable is not, however, a sufficient condition for the existence of uncertainty regarding that variable in the minds of decision-makers. Take, as an example, the number of Easter eggs sold in March of a given year. This varies sharply from year to year, but the variation is largely determined by the date on which Easter falls. Thus, the value for, say, 1988, may be predicted quite well.

Because instability may be observed objectively, it is susceptible to measurement. This topic will be given a rigorous mathematical treatment later, but an informal development of the relevant ideas will be useful now. Consider first the case where a given variable θ can take only a finite range of values $x_1, x_2, \ldots x_n$. Then it is possible to associate to each x_i a probability of occurrence p_i. This probability may be interpreted as the frequency of occurrence of outcome x_i, expressed as a proportion, in a large number of realisations of the variable θ. The entire (population) distribution may be written as a **prospect**

$$(1.2.1)\ \theta = \{\mathbf{x};\mathbf{p}\} = \{(x_1, x_2, .. x_n);(p_1, p_2, \ldots p_n)\} \quad ,$$

where $p_i > 0$ and $\Sigma_i p_i = 1$.

It will normally be assumed that the x_i are members of an ordered set, and in this case the subscripts will be chosen such that $x_1 \leq x_2 \leq \ldots x_n$. In the case of a variable y taking values in an ordered set of outcomes, it is possible to define the **cumulative distribution function**

$$(1.2.2)\ F_\theta\colon\ x \to \Pr\{ y \leq x\}$$

This definition is meaningful whether or not the outcome set is finite . For the case of a finite outcome set, it is apparent that

$$(1.2.3)\ F_\theta(x_j) = \Sigma_{i \leq j}\, p_i \quad .$$

If the distribution function F is differentiable, then it is further possible to define the **density function** $f\colon x \to \partial F/\partial x$.

While there are no major mathematical problems in moving from the finite to the infinite case, there are some difficulties associated with the maintenance of the "frequentist" interpretation of probability mentioned above. In general, it is no longer meaningful to talk of the frequency with which a given outcome occurs, but only of the frequency with which the outcome lies in a given set.

In the case of real-valued variables, well-known measures such as the mean and variance can be used to describe the characteristics of a population of realisations of a given random variable. The mean is given by

(1.2.4a) $E[x] = \Sigma\, p_i x_i$, and

(1.2.4b) $E[x] = \int x\, f(x)\, dx$,

for the finite and infinite cases, respectively. Similarly, the variance is given by

(1.2.5a) $\sigma_x^2 = \Sigma_i (x_i - E[x])^2$,and

(1.2.5b) $\sigma_x^2 = \int (x - E[x])^2 f(x)\, dx$ respectively.

Another useful concept is that of the coefficient of variation which is the ratio of the standard deviation to the mean. This measure of variability is useful only for variables which are always positive.

The mean and variance are also referred to as the first moment and second central moment of a distribution. More generally, it is possible to define the k-th moment of a distribution as

(1.2.6) $E[X^k]$,

and the k-th central moment as

(1.2.7) $E[(X-E[X])^k]$

It may be shown, using the binomial theorem, that knowledge of the first n moments is sufficient to infer the values of the first n central moments (Ash, 1972, p.226). In practice, at least in economics, only the first three moments of a distribution are important. The third central moment measures the skewness of the distribution, with a value of zero arising in the case when the distribution is symmetric about the mean. An important special case of a symmetric distribution is the normal distribution, which has the notable property that it is fully characterised by its first two moments, the mean and variance. This property of the normal distribution had a significant influence on the early development of theories relating to economic behaviour under uncertainty.

Although the moments are extremely useful, they can be somewhat intractable in a number of contexts. It is frequently useful to use an alternative characterisation in terms of **cumulants**. It is possible to define a characteristic function f(t) for any distribution, such that the moments are the coefficients fo the power series expansion of f. The cumulants may then be defined as the coefficients of the power series expansion of log(f). These have the property that they are invariant under linear transformations.

In much economic analysis, it is desirable to compare and rank distributions, on the assumption that the ordering of the outcomes corresponds to some preference orderings. The simplest case is when the outcome set is a subset of the real line representing levels of income or wealth. One approach is to compare the moments of the distribution. It is obvious, for example, that, *ceteris paribus,* the higher the mean the more desirable the distribution. On the other hand, a lower variance is generally preferable. This development can be carried somewhat further (for example, it may be argued that positive skewness is desirable) but it is difficult to obtain an intuitive feeling for the desirable values of the fourth and higher moments. This problem does not arise, however, if attention is confined to normal distributions, where the mean and variance give all the necessary information. Since one distribution may have a higher mean and a higher variance than another, it is impossible to give a total ordering on the basis of this mean-variance analysis. It is, however, possible to draw indifference curves and estimate marginal rates of substitution. Markowitz (1959) developed an extensive system of portfolio analysis on this basis and further extensions were made by Sharpe (1964), Lintner (1965) and others.

The mean-variance approach works well for normal distributions, but has proved unsatisfactory in general. A more sophisticated approach is based on the concept of **stochastic dominance**, developed by Quirk and Saposnik (1962), Hadar and Russell (1969) and Hanoch and Levy (1969). Given two cumulative distribution functions F_1 and F_2, F_1 is said to first stochastically dominate F_2 (written F_1 FSD F_2) if

$$(1.2.8)\quad F_1(x) \leq F_2(x) \quad \forall\, x \ .$$

This definition means that for any outcome **x**, the probability of a worse outcome is greater under the second distribution. It is clear that the first distribution must have a higher mean then the second. It is intuitively plausible that if F_1 FSD F_2 then rational individuals should prefer F_1. This requirement provides a useful test for theories of choice under uncertainty, as many proposed theories have been shown to imply violations of (first stochastic) dominance. For example, the mean-variance approach can yield violations of dominance if it is applied to classes of non-normal distributions such as the uniform distributions. (Consider two uniform distributions without any overlap. These are related by first stochastic dominance, but if their variances differ sufficiently, mean-variance theory may lead to the selection of the dominated distribution.)

However it is rare for practical decision problems to be characterised by the existence of one alternative which first stochastically dominates all others. For this and other reasons, considerable attention has been paid to the weaker concept of second stochastic dominance. F_1 second stochastically dominates F_2 (F_1 SSD F_2) if

$$(1.2.9) \int_{-\infty}^{x} F_1(t)\, dt \leq \int_{-\infty}^{x} F_2(t)\, dt \ \forall x.$$

Rothschild and Stiglitz (1970) develop a related concept of "increasing risk", which is equivalent to second stochastic dominance with the additional condition that the two variables should have the same mean.

The concept of increasing risk may be characterised in several different ways, which Rothschild and Stiglitz show to be logically equivalent. The first relates to the concept of **simple mean preserving spreads.** This concept is a strict version of the intuitively plausible requirement that "one distribution have more weight in the tails" than the other. Let F_y and F_z be two distributions with equal mean. Then F_z is derived from F_y by a simple mean preserving spread if the difference between the two distributions displays the **single crossing property**, that is, there exists an x* such that

$$(1.2.10)\ x \leq x^* \Leftrightarrow F_y \leq F_z$$

The single crossing property is illustrated in Figure 1.1. The spread is represented as a transfer of weight form the centre of the density function to the tails.

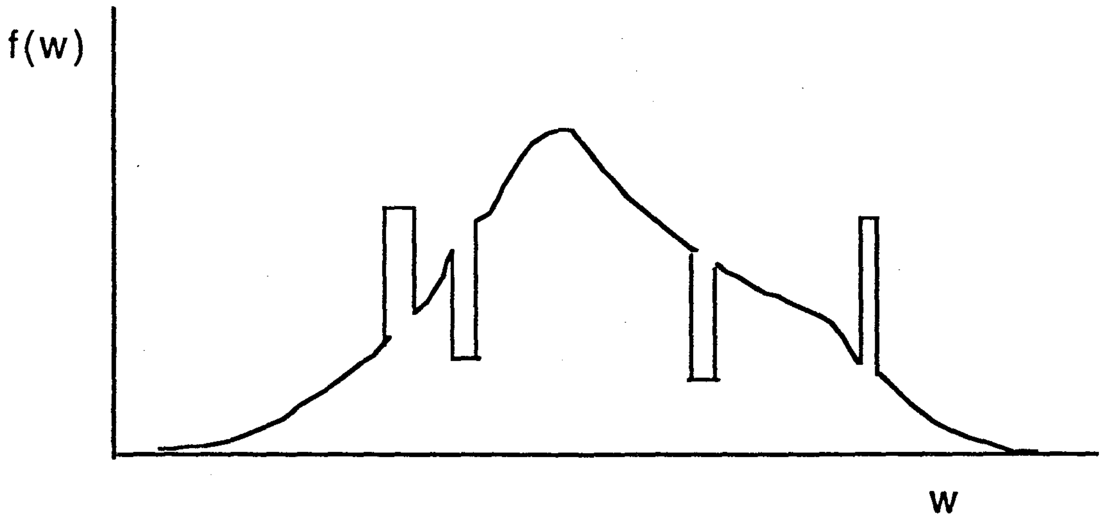


Figure 1.1a Change in density function under a simple spread

It is obvious that the single crossing property implies second stochastic dominance. Rothschild and Stiglitz show further that F_z is riskier than F_y if and only if it can be generated from F_y as the limit of a sequence of simple mean preserving spreads. [1] Another equivalent characterisation is the requirement that the second distribution may be derived from the first by the addition of "noise" in the form of an independently distributed random variable with zero mean. That is, F_y SSD F_z if and only if there exists w independent of y and such that $F_z=F_{y+w}$. Stochastic dominance conditions are discussed further in section 2.4, in relation to the theory of expected utility and its generalisations.

[1] The requirement that the spreads be mean-preserving is crucial here. Meyer(1975) proposes a definition of increasing risk which dispenses with this requirement and finds that it is equivalent to imposing the minimax rule that only the worst possible outcome should be considered in comparing prospects.

1.3 Risk

The concept of instability refers to the objectively observed distribution of a variable over time or perhaps space. By contrast, the term **risk** will be used to refer to subjective beliefs held by particular decision- makers concerning particular events, including individual realisations of a given random variable. This distinction between risk and instability was articulated by Quiggin and Anderson (1979, p.192) :

> "It is necessary to distinguish between decisions that must be made without knowing what realisation a given random variable will take and those that are made with this knowledge. People making decisions of the first type are subject to both risk and instability while those making decisions of the second type are subject to instability only."

This distinction makes it fairly clear that risk, rather than instability *per se*, is the major economic problem associated with uncertainty. It is not, however, immediately clear how subjective risk should be modelled. The most forthright answer to this question is given by Bayesian decision theorists, who argue that subjective uncertainty should be modelled using the concepts of probability distributions outlined in section 1.2 . This approach implies a rejection of the "frequentist" interpretation of probability outlined above in favour of one based on beliefs about states of the world. A mathematical development of such an interpretation is given below, while a good general outline of the approach is given by Anderson, Dillon and Hardaker (1977).

One of the most important aspects of subjective risk is the way in which the subjective probability distribution associated with a particular event changes with the acquisition of new information. A crucial tool in this connection is Bayes' theorem, originally developed by the English clergyman, Thomas Bayes in the 18th Century (Bayes 1763). The theorem may be developed from the concept of conditional probability. Intuitively, the conditional probability of A relative to B, written $Pr(A | B)$ is the probability of A occurring , given that B has occurred. If the set of possible events after which A can occur is $B_1, B_2, \ldots B_n$, then the probability of A is $\Sigma_i Pr(A | B_i) Pr(B_i)$. Working backwards, it is now possible to determine the probability that B_i has occurred given that A has occurred .

$$(1.3.1)\ Pr(B_i | A) = Pr(A | B_i)\, Pr(B_i) / \Sigma_i\, Pr(A | B_i)\, Pr(B_i) \ .$$

This approach shows how, given a **prior** subjective probability distribution, the acquisition of a new item of information may be used to develop a new **posterior** distribution. A particularly important case is that which arises when the new information consists of a random drawing from the population under consideration. However, other forms of information may also be incorporated using the theorem, provided that conditional probabilities can be assigned.

This approach has had a number of critics. Defenders of the frequentist interpretation of probability have attacked the concept of prior probability distributions, and hence denied meaning to Bayes' theorem. Bayesians have responded by arguing that the frequentist notion of probability is so limited as to have no practical value, since it can only apply to events which are members of an infinite series of drawings from a given population. Thus, the frequentist approach cannot be used in relation to events which are in any way unique. Cornfield (1967, p44) argued:

> "If we ask the probability that the 479th digit in the decimal expansion of π is a 2 or a 3, most people would say 2/10, but the frequentist, if he answers the question at all, must say 0 or 1, but that he does not know which."

Other writers have sought to limit the scope of the concept of subjective probability, by arguing that such probabilities cannot be assigned in many cases. These attacks, centred around concepts of ambiguity and uncertainty are considered below.

This thesis will be developed using subjective probability concepts, on the pragmatic basis that no meaningful alternatives have been developed that can act as either a basis for positive analysis of decisions under uncertainty or as a normative guide to decision-makers. However, the criticisms of this approach will be examined closely and it will be shown that the decision theoretic approaches now in use can be modified so as to overcome many of these criticisms. This development will be assisted by a more rigorous mathematical formulation.

A measure-theoretic approach

In this section the concept of probability will be formalised using the tools of measure theory and real analysis. Ash (1972) has been used as the main source, but the analysis differs slightly in notation and substantively in the attempt to give an interpretation in terms of subjective probability. The central concept is that of the set Ω of states of the world. An "event " A is represented as a subset of Ω consisting of all those states $w \in \Omega$ such that a particular statement is true. For example, the event "Party Z wins the next Australian election" consists of all states of the world in which Party Z wins, regardless of what happens to any other variable. The analysis will deal with a family of events ,S,which is assumed to be a σ-field, that is, to satisfy the following axioms:

(1.3.2a) $\Omega \in S$;

(1.3.2b) If $A \in S$, then so does the complement of A in Ω;

(1.3.2c) A is closed under the operations of countable union and intersection.

It is now possible to define a probability measure over S as a non-negative real-valued function $P: S \to \Re$, such that $P(\Omega)=1$ and, for any finite or countable collection of events $A_1, A_2 \ldots$,

$$(1.3.3) \quad P(\cup(A_i)) = \Sigma P(A_i).$$

The measure P denotes the subjective probability associated with any given event A. In the case of a finite set Ω, we may take the family consisting of all subsets of Ω as the σ-field S and the measure P is determined simply by setting $P(\omega)$ equal to the probability of the state ω for each $\omega \in \Omega$.

A random variable θ is a function from Ω to some set X of outcomes. Thus, θ associates with each state of the world $\omega \in \Omega$, an outcome $x \in X$. Of particular interest is the case when X is a subset of the real line, representing, say, possible prices for wheat, or levels of real income. Such a function is said to be (Borel) measurable if the inverse image of every set of the form (a,b] is a member of S. Thus, it is possible to derive the probability that the outcome $\theta(\omega)$ lies in any Borel subset of the real line. The distribution function F defined in equation (1.3) above may now be restated for any real-valued θ as

$$(1.3.4) \quad F(x) = P\{\omega: \theta(\omega) \leq x\}.$$

More generally, let Ω_1 and Ω_2 be sets with associated σ-fields S_1 and S_2. Then a function $\eta: \Omega_1 \to \Omega_2$ is measurable (with respect to S_1 and S_2) if, for any $A \in S_2$, $\eta^{-1}(A) \in S_1$. From now on, it will be assumed that all random variables are measurable.

The mean and variance of random variables may be calculated as above using integration, though it is now necessary to work in terms of the more general Lebesgue-Stieltjes integral (see Ash, 1972, pp. 53-7). Finally, it is possible to give a rigorous development of the concept of conditional probability. It may be observed that the analysis described above works only for a finite set of events with non-zero probabilities. However, it is clearly desirable to have a concept of conditional probability which works in the case of continuous random variables for which $\Pr\{\theta=x\} = 0$ for all values of x.

Consider two real-valued random variables θ_1 and θ_2. Then, by the use of product measures, it is possible to define a variable (θ_1,θ_2) taking values in $\Re^2$. In particular, for any measurable subsets A,B of $\Re$, it is possible to define $\Pr\{A \times B\}= \Pr\{\theta_1 \in A, \theta_2 \in B\}$. In order to obtain a conditional probability measure, it is necessary to define for each $x \in \Re$ a probability measure $\Pr\{x,\bullet\}$ which may be regarded as defining the probability distribution for θ_2 given $\theta_1 = x$. In order for this measure to be regarded as a conditional probability measure, it is necessary that

$$(1.3.5) \quad \Pr(A \times B) = \int_A P(x,B)\, dF(x) \quad \forall A,B.$$

Ash (1972, Ch 6) demonstrates that there is a unique (in the sense that any two such measures are derived from functions which are equal almost everywhere) way of constructing a conditional probability measure satisfying this condition.

1.4 Ambiguity

A significant group of economists following Knight (1921) have maintained that the 'frequentist' approach is the only legitimate basis for the assignment of numerical probabilities. Knight (1921, p232) referred to the case when this can be done as 'risk' and to other cases as 'uncertainty'. He argued that only uncertainty is economically relevant.

> "An uncertainty which can by any method be reduced to an objective, quantitatively determinate probability can be reduced to complete certainty by grouping cases. The business world has evolved several organization devices for effectuating this consolidation, with the result that when the technique of business organization is fairly developed, measurable

uncertainties do not introduce into business any uncertainty whatever."

Others, notably Keynes (1920, 1936), have accepted a concept of subjective probability, but sought to limit its scope. Keynes (1920) argued that, for many events, no numerical assignment, or even pairwise comparison, of probabilities was possible. Keynes suggested that a partial order could be imposed on the set of possible events. This partial order would contain both a minimal element (event with zero probability) and a maximal element (event with probability 1). These ideas were further developed in Chapter 12 of the General Theory. Keynes (1936) argued that in the absence of sufficient knowledge to form a reasonable subjective probability, the *status quo* will be given more weight than can be justified on objective grounds.

> "We are assuming, in effect that the existing market valuation, however arrived at, is uniquely correct in relation to our existing knowledge and that it will only change in proportion to changes in this knowledge; though, philosophically speaking, it cannot be uniquely correct, since our existing knowledge does not provide a sufficient basis for a calculated mathematical expectation. In point of fact, all sorts of considerations enter into the market valuation which are in no way related to the prospective yield. "(Keynes 1936, p152)."

In both cases the implications of the argument have gone well beyond probability theory. Knight's arguments have been developed by a number of writers, particularly those in the Austrian tradition, who have sought to emphasize the role of human powers of exploration, discovery and decison-making, particularly in relation to entrepreneurial activity, and to attack the neoclassical reliance on optimisation and equilibrium. The 'classical' writers in this tradition were Hayek (1937) and von Mises (1951).More recently, these views have been elaborated by Lachmann (1977), Kirzner (1979) and Moldofsky (1982).

In general, these arguments lead to a strongly laissez-faire policy perspective, in which the various efficiency-based arguments for government intervention developed by economists in the neoclassical tradition remain essentially outside the universe of discourse. Moldofsky (1982, p.161) states:

> "Thus, while implying a strong case for an unhampered market system, the 'Austrian' theoretical framework shuns all

claims that the system is perfect. In an uncertain human world the market system, however free, cannot possibly attain full coordination and a state of equilibrium continues to remain elusive. Nevertheless according to this view, for the purposes of communicating new knowledge, attaining better coordination of diverse, even contradictory, plans and effectively allocating resources, the market system, however imperfect, still seems to stand out as the best available device known to man."

Writers in the post-Keynesian tradition have used the concept of uncertainty very differently. While the mainstream neoclassical synthesis of Keynes' ideas, largely due to Hicks (1937), effectively discarded Keynes' concerns with uncertainty, post-Keynesian writers, such as Minsky (1975), Chick (1983) and especially Shackle (1968), have pursued them vigorously. These writers stress the idea, developed in Chapter 12 of the General Theory, that the uncertainty associated with long-run investment decisions is such that these decisions cannot be based on rational optimisation, but only on a set of expectations which are essentially conventional in nature. The rapid revision of these expectations is a major element of economic crises. Hence, full-employment equilibrium is inherently unstable in the absence of extensive government intervention.

The Keynesian and Austrian treatments of uncertainty have in common the idea that uncertainty is outside the scope of the rational optimising methods that characterise neoclassical concepts such as profit-maximisation. By contrast, much of the work discussed and presented in this thesis involves an attempt to extend these methods to the case of decisions under uncertainty. It is obviously important to consider the extent to which this feasible and the limits which must be placed on decision theory.

Ellsberg and ambiguity

In general, the criticisms of Keynes and Knight have not had much impact on the users of subjective probability concepts. The most common response has been to use the concept of betting In order to estimate, a person's subjective probability that a particular event will occur, they may be asked would you bet on that event occurring at even money? At 5 to 1? At 10 to 1? etc. Since, for most people, there is some set of odds at which they are willing to bet on any given event, it would seem, at least, that all events may be compared with certain simple risks. In order to assimilate these events to subjective probability theory, it is necessary to go further and examine the least favorable odds at which a person would be willing to bet on the occurrence of a given event. It has often been incautiously asserted that this is the person's subjective probability for that event. This need not be the case even if well-defined subjective probabilities exist. In this section, the main concern will be with the implicit assumption of "fair betting" (Gottinger 1974) that is, that once the odds fall below the cut-off point described above, the person would be willing to bet aginst the event occurring.The first direct criticism of this assumption was made by Ellsberg (1961) who introduced the term **ambiguity** to cover Knight's concept of uncertainty. Ellsberg offered a number of decision problems indicating that the assumption might be violated by reasonable people. The most striking was as follows. Consider an urn containing 30 red balls and a total of 60 black and yellow balls, the latter in unknown proportion. Which of the following bets do you prefer :

(1.4.1a) $100 if a red ball is drawn, nil otherwise; or

(1.4.1b) $100 if a black ball is drawn, nil otherwise.

Most people choose (1.4.1a) indicating, on the orthodox view, that the subjective probability of drawing a red ball is greater than that for a black ball. Ellsberg then offered the problem of choosing between

(1.4.2a) $100 if a red or yellow ball is drawn, nil otherwise;

and

(1.4.2b) $100 if a black or yellow ball is drawn, nil otherwise.

Most people now choose (1.4.2b) which has the known probability 2/3, thus apparently indicating that their subjective probability for a black ball is greater than that for a red ball.

Ellsberg argued that, in situations of known probabilities, people will tend to follow decision rules of the type advocated by von Neumann and Morgenstern (see Chapter 2). By contrast, as confidence in probability estimates declines, maximin or minimum-regret rules of the type advocated by Shackle will be given a greater weight.

Raiffa (1961) argued that the subjective probability approach is normatively superior as a guide to rational behaviour, even if it does not predict well. Raiffa pointed out an apparent inconsistency in the choices of Ellsberg's subjects. Consider the following options. In option A a fair coin is tossed and prospect (1.4.1a) is taken if heads come up, (1.4.2b) if tails come up. In option B heads yields (1.4.1b), and tails (1.4.2a) . Dominance suggests that option A should be preferred. But an analysis of the prospects shows that the two options are objectively identical. Whichever ball is drawn, both options yield a fifty per cent chance of winning.

One explanation of the Ellsberg 'paradoxes', is related to Akerlof's (1970) 'lemons' model. Akerlof argued that the difference in price between new and almost new cars could be explained by observing that the very fact of an almost-new car being offered for sale increased the likelihood that it was a 'lemon'. Similarly, the fact that a bet is offered on an uncertain event raises the possibility that the person offering the bet has information indicating that their side of the bet is favourable. Thus it is quite rational to attach a subjective probability of a black ball coming up of less than 1/3 if you are invited to bet on it, as in 1.4.1b, and less than 1/3 if you are invited to bet against it, as in 1.4.2a . If you are free to choose which outcome to bet on or can split bets as in Raiffa's example, then it is reasonable to assume the two outcomes to be equi-probable. An argument of this kind was presented by Brewer (1963).

This point can be illustrated even more sharply by the following pair of examples:

(1.4.3a) A fair coin is tossed once. If it turns up heads you receive \$2, otherwise lose \$1; and

(1.4.3b) Same as (1.4.3a) except that the coin is biased (but you don't know in which direction).

This choice is ,in formal terms, very similar to that in the urn problem, but it brings out more clearly the reason for preferring the bet at known odds, namely the fact that the person offering the bet will normally know the nature of any bias.

Thus, the reactions to 'ambiguity' observed by Ellsberg can be explained in terms of the differential information sets possessed by the parties to a gamble. Two interpretations of this result are possible. The first is that ambiguity is a non-problem, and that once the 'lemons' effect is taken into account, subjective probability concepts provide an adequate basis for analysis. The second, which will be explored in the following section, is that the different aspects of uncertainty, described above as instability, risk and ambiguity, must be analysed in terms of information sets.

1.5 Risk, instability and ambiguity - an informational approach

The discussion in section 1.4 centred on informational asymmetries between individuals. It was argued that individuals might distinguish between 'ambiguous' and 'risky' bets on the basis that, in the former case, others might have better information about the likelihood of the different outcomes. It is not immediately clear whether such a distinction can be extended to the case of 'games with Nature', that is, decision problems where the outcome is not affected by the choices of other individuals. In this section, an attempt to draw such a distinction will be made, and its implications considered.

Consider the following definition:

> "A person is faced by risk in relation to a particular decision if there is no information available at 'reasonable' cost to them or anyone else which would alter the subjective probability distribution for the event in question and by ambiguity if such information is available".

This distinction clearly covers the case of the fair and biased coins. In the case of a fair coin, the probability that a given toss will yield heads is 1/2, and no amount of observation of previous tosses will change this. On the other hand, it is, *a priori*,. equally likely that a biased coin will be biased towards heads or tails (unless there is a particular preference for double-headed pennies) so that to someone who knows nothing more than the fact that the coin is unfair, the subjective probablity of a head is still 1/2. In this case, however, more information is available either by observing a number of tosses or by subjecting the coin to physical examination. [2]

This distinction may also be extended to the case of 'games with Nature'. Variables which generate risk are typically those which may be modelled as random drawings from a population on which many observations have already been made. An example might be the monthly rainfall for July 1990. Ambiguity is present when our subjective probability distribution for the variable is determined by a

---

[2] It is worth noting that the concept of 'additional information' need not be confined to new empirical data. In some cases, notably that of a number of 'bar bets', no new data is required, but additional information can be obtained through knowledge and application of the relevant probability theory.

model (formal or informal) which could be improved by the addition of further information. An example might be the probability of rain at the coming weekend.

These examples reveal an important feature of the distinction between risk and ambiguity. Improved information may yield a shift from risk to ambiguity or vice versa. Whereas the probability of rain on a given day well into the future is a matter of pure risk, the probability of rain tomorrow may be subject to ambiguity. For example, in the case of tomorrow's weather, a person with no 'intuitive' forecasting ability would be in a situation of pure risk after examining, say, the Meteorological Bureau's previous record of predictive accuracy (or otherwise) in similar situations.

It is useful to compare the concepts of risk and ambiguity with the Bayesian notions of tight and diffuse prior distributions. There is an important correspondence in that situations of diffuse priors and of ambiguity share the characteristic that additional observations may have a significant impact on the subjective probability distribution whereas in the case of risk or tight priors they will not. The most important difference is that the degree of risk or ambiguity is unrelated to the shape of the (prior) subjective probability distribution, whereas diffuseness of priors is generally associated with a close to uniform distribution. For example, the probability distribution for a roll of a fair die is uniform, just as it is for a die with unknown bias, although the first situation is one of risk and the second one of ambiguity. A second difference is that the acquisition of additional information normally acts to make Bayesian priors tighter, but there is no such uniform tendency in the case of the risk/ambiguity distinction.

The most important problem associated with the distinction between risk and ambiguity is the extent to which decision-theoretic methods can be applied in situations of ambiguity. An important special case is that where the additional information consists of sample observations of a random variable. In this case, Bayesian decision theory provides a fairly complete solution. Given any prior distribution for a given random variable, it is possible to compute the difference between the expected value yielded by a decision made before the variable's value is known and one made after its value is known (that is, between a choice under risk and one under instability). This difference is referred to as the cost of imperfect information. A similar calculation is possible for the posterior distributions arising from any possible value for a given sample observation. Thus it is possible to compute the expected value of additional sample information.

If the term 'reasonable cost' in the definition of ambiguity given above is used to denote a cost less than the expected value of the information concerned, then it is apparent that, in this case, Bayesian decision theory can be used to convert a situation of ambiguity into one of risk, by taking additional sample observations until the cost would exceed the expected benefit.

A second case of interest is that where additional information becomes available over time. It is thus necessary to decide on when to make a decision, such as a choice of output levels. The benefits of new information must be balanced against the costs of deferring decisions. Some special cases of this problem have been analysed in detail. An example is the literature on irreversibilities in cost-benefit analysis (A good summary of the issues is provided by Resources for the Future 1982).

The general problem of optimal timing of decisions in the face of a gradual, but unpredictable and not necessarily monotonic, convergence of the subjective probability distribution is exceedingly complex, and, at this stage, some reliance on heuristic decision rules seems unavoidable. The main heuristic that can be derived from the discussion above is the need to make allowances for future flexibility in the evaluation of decisions. Thus, a decision that appears superior on the basis of the current subjective probability distribution may be rejected in favour of one which allows greater opportunities for adjustment to new information.

The discussion above permits the delineation of three aspects of uncertainty. Instability refers to objectively observed variation in a population of realisations of a given random variable, and in particular to variations observed over time. Risk refers to subjective uncertainty concerning a given realisation of a random variable relevant to a particular decision. Finally, ambiguity refers to the case where the information set on which subjective probability distributions are formed is not fixed for a given decision, but may be altered ,for example, by market transactions or by the additional sample observations.

The remainder of this thesis is devoted primarily to the analysis of behaviour under risk. However, the distinctions that have been outlined should be kept in mind, and will prove especially important in

the discussion of policy issues such as agricultural price stabilisation and underwriting.

2. UTILITY

2.0 Introduction

In Chapter 1, concepts of uncertainty were developed without the use of any formal theory of choice, although the existence of a preference ordering over outcomes was implicitly assumed. The basic objective of the present Chapter is to describe the theory of choice under uncertainty which is currently predominant in economic analysis, the Expected Utility (EU) theory of von Neumann and Morgenstern (1944).

In section 2.1, a discussion of the historical background of utility theory is given. Among other things, this discussion indicates that utility theory has historically been closely related to questions such as gambling and social welfare attitudes, neither of which receives much attention in the modern literature on choice under uncertainty. In section 2.2, an outline of EU theory is presented, along with the basic notions of risk aversion and risk preference. In section 2.3, it is argued that risk preference in EU theory implies highly implausible patterns of behavior and that attention may therefore be confined to risk aversion. This argument does not appear to have been developed previously. Section 2.4 describes some of the major tools of analyis for EU theory, such as coefficients of risk aversion and stochastic dominance concepts. Applications of the theory are discussed in general in section 2.5 and the specific case of the theory of the firm under uncertainty is examined in more detail in section 2.6. Finally, in section 2.7 some issues from the philosophy of science are discussed in order to lay the groundwork for a comaprison between EU theory and competing research programs.

The objective of this chapter is to demonstrate the power and intellectual appeal of EU theory. In subsequent chapters, it is argued that, despite these qualities, EU theory has outlived its usefulness and should be replaced by more general and less restrictive models of choice under uncertainty.

2.1 Background

The early development of probability theory was based on the analysis of gambling by Pascal and Fermat. In this context, it was fairly natural to base the analysis on the concept of maximising expected returns, since it is precisely by securing a positive expected return on each of a large number of gambles that the successful gambler operates. The first writer to challenge the appropriateness of profit maximisation as a basis for economic decisions under uncertainty was Bernoulli (1738).

Bernoulli argued that "the determination of the value of an item must not be based on its price, but rather on the utility (emolumentum) it yields. Given a choice between a prospect yielding 20000 ducats with probability 1/2 and nothing otherwise, and one yielding 9000 ducats with certainty, a poor man would be well advised to choose the latter and a rich man the former". In modern terms, this suggests a theory of positive but declining absolute risk-aversion (see section 2.4). Bernoulli made this more precise by asserting that "the utility resulting from any small increase in wealth will be inversely proportional to the quantity of goods previously possessed". On the basis of this assumption (implying, in modern terms, constant relative risk-aversion), he derived a logarithmic utility function, illustrated its property of declining absolute risk aversion, and pointed out the advantages of diversification for a risk-averse individual.

Bernoulli's final argument against profit-maximisation was the "St. Petersburg paradox" . This is a prospect which offers \$1 with probability 1/2, \$2 with probability 1/4 and generally $\$2^{n-1}$ with probability 2^{-n} (it may be specified in terms of a payment depending on the number of successive heads a player can toss with a fair coin) . The expected value of the prospect is infinite, yet introspection suggests that most people would not be willing to pay more than \$20 for it. {The subsequent history of the St. Petersburg paradox has been chequered. Menger (1967) and Arrow (1974) used it to argue that the utility function must be bounded. Shapley (1977a) on the other hand argued that it was a "con game". His point was that the game had a value in excess of \$50 only for people who actually believe they will be paid $\$2^{51}$ (more than the entire wealth of the world). Aumann (1977) responded with a less extreme version of the paradox. In his response, Shapley agreed that the utility function should be bounded but argued that the

St. Petersburg paradox was neither necessary to reach this conclusion nor convincing as a basis for it.

The main theoretical development in the 19th Century was the adoption of explicit concepts of utility, first by Bentham and the "Philosophic Radicals" and then by the mainstream classical and neoclassical economists. With the development of marginal utility analysis by Jevons and others, the concept of utility became indispensable to economists. Moreover, the 19th Century economists did not display any greater doubts about the real existence of utilities than do modern physicists about their equally unobservable "atoms". Thus, the argument that declining marginal utility of wealth provided a strong argument (on utilitarian grounds) for egalitarian income redistribution, had a considerable impact. Since such redistributive measures ran counter to the generally laissez-faire tenor of neoclassical economics, utility theory was regarded as something of a mixed blessing.

The problem was resolved by Robbins (1938) who proclaimed that interpersonal utility comparisons were unscientific and should be avoided. This declaration was followed by an surge of interest in revealed preference theory, a means by which the troublesome concept of utility could be exorcised altogether. While concepts of ordinal utility might still be used for convenience, the dangerous ground of cardinal utility would be strictly avoided.

Ironically, just as this approach achieved its greatest success with Samuelson's (1947) recasting of welfare economics in terms of revealed preference, the concept of cardinal utility theory was revived by von Neumann and Morgenstern in their analysis of behaviour under risk.

2.2 Expected Utility theory

The development of Expected Utility theory (hereafter referred to as EU) by von Neumann and Morgenstern (1944) provided the basis for the analysis of economic behaviour under uncertainty. While the EU model has been subject to many criticisms, the main result has been to generate modifications and generalisations of the model. The basic tool for EU analysis is the utility function $U: W \rightarrow \Re$, a real-valued function on some outcome set W, which is normally a subset of the real line. Outcomes $w \in W$ may be interpreted as levels of wealth, income, or in some cases, consumption. (A utility function may be defined for outcomes expressed in other terms, but the main line of development is that described here). For any prospect {**w** ; **p**}, the **expected utility** of the prospect is defined by a functional V

$$(2.2.1)\ V(\{\mathbf{w} ; \mathbf{p}\}) = \Sigma_i\, p_i\, U(w_i) = E\,[U(w)] \quad .$$

More generally, if Y is a set of measurable random variables on W, a functional V:Y→W may be defined, such that for any random variable y ∈ Y with cumulative distribution function F,

$$(2.2.2)\ V(y) = \int U(w)\, dF(w).$$

In the standard case where W is a subset of the real line representing wealth levels, this approach yields a simple representation of the concept of **risk-aversion**. Preferences are said to exhibit risk-aversion if the function U is globally concave. It follows immediately from Jensen's inequality that a risk-averse person will always find a risky prospect less attractive than the certainty of receiving the expected value of the prospect. A converse analysis applies to the convex case ('risk-seeking'or 'risk-preferring' behaviour).

In view of the discussion of section 2.1, it is normally assumed that U is bounded. If U is also globally concave, this implies that W must be bounded below, an assumption which is intuitively plausible, since it is difficult to give a meaningful interpreation of large negative levels of wealth or income and clearly meaningless to talk of negative consumption. On the other hand, there is no obvious upper bound to levels of wealth which can at least be imagined, if not realised in practise. For this reason it is frequently assumed that W is a set of the form $[0, \infty)$ where the zero corresponds to some concept of bankruptcy.

In addition to presenting the EU function, Von Neumann and Morgenstern developed an axiomatic basis for EU theory. For any person whose preferences over risky prospects satisfies the von Neumann-Morgenstern axioms, they showed that there exists a function U (unique up to a linear transformation) such that one random variable, y_1, is preferred to a second, y_2, if and only if $V(y_1) \geq V(y_2)$. The appropriate choice of an axiomatic basis has been a central feature of many subsequent debates over utility theory. The axiomatic basis for EU theory consists of a series of requirements on an individual's preference relationship P, and the associated indifference relationship I.

EU1 (Completeness) - $\forall\, y_1, y_2 \in Y$, either $y_1 \, P \, y_2$ or $y_2 \, P \, y_1$

EU2 (Transitivity)- $\forall\, y_1, y_2, y_3 \in Y$, if $y_1 \, P \, y_2$ and $y_2 \, P \, y_3$, then $y_1 \, P \, y_3$

EU3 (Continuity) -$\forall\, w_1, w_2, w_3 \in W$, $w_1 \, P \, w_2 \, P \, w_3$
$\exists\, p \in [0,1]$, $w_2 \, I \, \{(w_3, w_1);(p,1-p)\}$

EU4 (Independence) - $\forall\, y_1, y_2, y_3 \in Y$, $p \in [0,1]$
$y_1 \, P \, y_2 \Rightarrow \{py_1+(1-p)y_3\} \, P \, \{p\, y_2+(1-p)\, y_3\}$

It should also be noted that the way in which prospects have been formulated presupposes certain assumptions about risk attitudes. For example, it is assumed that people are concerned only about outcomes and not about states of the world *per se* . Another implication of the formulation relates to the 'reduction of compound lotteries'. It is assumed that prospects initially expressed with outcomes which are themselves risky prospects are evaluated in terms of the ultimate probability of each outcome.

All of the axioms proposed by von Neumann and Morgenstern have been subject to some criticism, but most attention has been focused on the independence axiom. It has been reformulated in a number of ways, most notably in Savage's (1954) 'sure thing principle'. Savage's formulation involves arranging the prospects as a two-stage lottery. The prospect $py_1+(1-p)\, y_3$ is arranged as a lottery in which the first stage yields prize y_1 with probability p and prize y_3 with probability (1-p). The prospect $p\, y_2+(1-p)\, y_3$ is arranged in the same way except that y_1 is replaced by y_2. Since y_1 is preferred to y_2, Savage argues that acceptance of the independence axiom may be justified on the same grounds as acceptance of the dominance axiom, that is, that in some sense a preferred outcome is a 'sure thing'.

This argument has been strongly criticised by writers such as Allais (1979). It may be noted that the 'sure-thing' formulation of the independence axiom depends on the choice of a particular arrangement of the possible outcomes and no obvious justification for this arrangement is given. In the case of first stochastic dominance, there is a natural arrangement of the outcomes which illustrates why the dominated prospect should be rejected. This is the arrangement of the outcomes as a distribution ranging from worst to best. First stochastic dominance implies that say , the outcome at the 57th percentile of the dominating distribution will be preferred to the corresponding outcome for the dominated distribution. This natural arrangement cannot be generalised to support the 'sure-thing' principle. It can, however, be used to support a weaker independence axiom. Weaker independence axioms are developed in Chapter 5.

A second, more intuitive criticism, may be developed as follows. The Independence Axiom implies a strict separation between the evaluation of different future states, which is rational enough in terms of an Arrow-Debreu style view of the world, but which is uncongenial to a less reductionist approach to life. This separation acts to eliminate human phenomena such as hope. For example, in evaluating a segment of a compound lottery in which one prize is a trip to Paris, the independence axiom forbids us to consider whether this is our only chance; that is, whether there exists another second-stage lottery with prizes which would enable us to make the trip. For many people, it seems reasonable to evaluate the set of possible outcomes as a whole rather than partitioning it into separate states which may then be added up to form an expectation.

Criticisms of this kind were made from the early days of utility theory. However, criticism of the axiomatic foundations of a theory is rarely very effective unless there is an alternative in the field. The EU approach initially faced strong competition from mean-variance analysis, exemplified by the work of Markowitz (1959) on portfolio analysis, but the logical foundations of this approach were far more dubious than those of expected utility theory, of which it constituted a special case. Practitioners of the mean-variance approach generally argued that, while expected utility analysis might be theoretically superior, it was of little practical use, since its formulation was so general as to prevent the derivation of sharp results. For example, Tobin (1969, p14) responded to the criticisms of mean-variance theory by Borch (1969) and Feldstein (1969), saying

> " The [mean-variance approach] was never advertised as a complete job or the final word and I think that its critics in 1968 owe us more than demonstrations that it rests on restrictive assumptions. They need to show us how a more general and less vulnerable approach will yield the kind of comparative static results economists are interested in. This need is satisfied neither by the elegant but nearly empty existence theorems of state preference theory nor by normative prescriptions to the individual that he should consult his utility and his subjective probabilities and then maximise."

EU theorists countered with the observation that (viewed in terms of EU theory) mean-variance analysis implied either that all random variables were normally distributed or that the utility function was quadratic. Given the bizarre properties of the quadratic utility function and the obvious non-normality of many random variables, neither of these alternatives seemed attractive.

The resolution of this debate was not due to a preference for theoretical elegance over practical usefulness, but rather to a gradual realisation that EU theory was, in fact, a more powerful basis for economic analysis than the mean-variance approach. This realisation depended, to a large extent on the development of analytical tools within the EU framework which permitted the derivation of sharper, and more accurate, results than those available from the mean-variance approach. As most of these tools relate to concepts of risk-aversion, it is worth giving some consideration to the relative prevalence of risk-aversion and risk-preference. In the following section, it is argued that EU theory implies that risk-aversion must be effectively universal.

2.3 The impossibility of risk preference

The basic argument against the possibility of risk preference in EU theory may be stated as follows. Suppose there are two individuals, both global risk-preferrers. Then they would mutually benefit from a bet in which each of them staked their entire wealth, with the odds being actuarially fair. The resolution of the bet would leave one bankrupt and the other considerably more wealthy. More generally, if there are n risk-preferrers, mutually desirable betting transactions among them will be available until all but one are bankrupt. Thus we should not expect to observe global risk-preference except possibly among bankrupts and billionaires.

This argument can even be carried over to the case of a single risk-preferrer in a community of risk-averters who face irremovable risks. The risk-preferrer will clearly benefit from offering to insure the risk-averters at actuarially fair rates. Even if, as seems likely, the rate struck is actuarially favourable to the risk-preferrer, the probability of bankruptcy will approach unity as the size of the available risk increases.

By contrast, it may be noted that in a finite community of risk-averters, even perfect risk-pooling cannot generally move everyone to their preferred portfolio (given a particular expected value), namely the riskless portfolio, at which point risk-aversion becomes irrelevant. Only two risk-preferrers are required to achieve this outcome.

A broadly similar argument applies to people with both concave and convex segments in their utility functions. Mutually beneficial gambles will always be available to people whose present wealth lies in a concave segment of the curve. Even in a concave segment of the curve, people will tend to gamble until they reach a point at which they are risk-averse in the sense that they would refuse all risky gambles with expected outcome zero or less. This point is developed further in Chapter 3, in connection with the analysis of gambling put forward by Friedman and Savage (1948).

Thus, risk-preference is an unstable state. Because it is easy and cheap to satisfy a desire for increased risk, and difficult and costly to reduce risk, only risk-aversion presents a major economic problem. Risk-preferrers will eliminate themselves from economic systems either by going bankrupt or by achieving a level of wealth at which they are effectively risk-averse.

These arguments do not, of course, demonstrate that risk-preferrers are irrational. They do show, however, that, except for bankrupts and billionaires, observed behaviour which appears to be inconsistent with risk-aversion cannot be explained within the EU framework by invoking the possibility of local or global risk-preference.

2.4 Tools of analysis for Expected Utility theory

The mean-variance approach to decision theory relies on a simple ranking of prospects. If one prospect has a higher mean and a lower variance than another, then it must be preferred. If the first prospect has a higher mean, but also a higher variance than the ranking will depend on specific risk attitudes. It is easy to see that this is unsatisfactory. For example, it does not unequivocally suggest that one should accept a free lottery ticket, since the alternative of rejection yields a mean and variance of zero.

The development of a more acceptable approach to the problem of ranking prospects using EU theory was a major factor in its success. The crucial tools were the Arrow-Pratt coefficients of risk-aversion and the the concepts of stochastic dominance described in section 1.2 . The crucial results for stochastic dominance are

Theorem 2.4.1 Let $y_1, y_2 \in Y$, y_1 FSD $y_2 \Leftrightarrow V(y_1) \geq V(y_2)$ for any EU functional V.

Theorem 2.4.2 Let $y_1, y_2 \in Y$, y_1 SSD $y_2 \Leftrightarrow V(y_1) \geq V(y_2)$ for any EU functional V such that the associated utility function U is globally concave.

Combining this result with those from section 1.2 gives the major result on increasing risk developed by Rothschild and Stiglitz (1970). Rothschild and Stiglitz demonstrated the equivalence of a number of possible definitions of 'increasing risk'.

Theorem 2.4.3 Let $E[y_1] = E[y_2]$. Then the following are equivalent:

(i) y_1 SSD y_2 (ie y_2 has more weight in the tails);

(ii) $V(y_1) \geq V(y_2)$ for any EU risk-averter; and

(iii) $y_2 =_d y_1 + y_3$ for some y_3 such that $E[y_3 \mid y_1] \equiv 0$.

It is of equal interest to note that some plausible definitions of increasing risk are not equivalent to (i)-(iii). The fact that not all increases in variance fit into this category has already been noted. There are also more restrictive definitions of an increase in risk. The first, which has proved useful in the theory of the firm under uncertainty (see section 2.6) is that of a multiplicative spread about the mean. The second is the requirement that an increase in risk should lead a risk-averse investor to purchase less of the risky asset in the standard one safe asset, one risky asset portfolio problem. Rothschild and Stiglitz (1971) examine this problem but are unable to find any simple conditions on the utility function under which a spread satisfying (i)-(iii) will imply a reduction in purchases of the risky asset.

The power of stochastic dominance analysis was greatly enhanced because of the earlier development of coefficients of risk-aversion. While risk-aversion is equivalent to the condition that $U''(w) < 0 \ \forall \ w$, the actual values of U'' convey little information since they are not scale- independent. Arrow (1963) and Pratt (1964) were the first to develop useful measures of risk-aversion with the coefficients of absolute and relative risk-aversion. The **coefficient of absolute risk-aversion** is defined by

(2.4.1) $r_a(w) = -U''(w)/ U'(w)$.

This coefficient is a local measure of aversion to risk in a fairly straightforward sense. Given two individuals with different levels of r_a (at their respective current wealth levels) then for 'sufficiently small' bets, the individual with the higher value of r_a will always have a higher risk-premium (be less willing to accept the bet). The coefficient can also be used as a basis for global comparisons of risk-aversion. If one individual has a higher $r_a(w)$ than another for every w, then the first may be shown to be more risk-averse, in the sense of having a lower certainty equivalent for any risky prospect.

It seems intuitively obvious that $r_a(w)$ will decline with w; that is, that an increase in wealth will increase an individual's willingness to accept a given bet. This hypothesis of **declining absolute risk aversion** (DARA) was mentioned above in connection with the writing of Bernoulli. It has received considerable support from empirical studies. Among the implications of DARA it is important to note the requirement that $U'(w) > 0 \ \forall w$.

Arrow and Pratt also proposed an alternative measure of risk aversion which is in some sense independent of wealth levels. This is the coefficient of relative risk aversion

(2.4.2) $r_r(w) = - wU''(w)/U'(w)$.

This coefficient measures willingness to accept bets expressed as a proportion of current wealth. A number of plausible utility functions, including the logarithmic function, display constant relative risk-aversion. In contrast to the widespread acceptance of DARA there is no general consensus on the likely behavior of $r_r(w)$ as w increases although Arrow (1974a) has argued that r_r will tend to increase.

Theorems 2.4.1 and 2.4.2 suggest an extension of stochastic dominance conditions. Whitmore (1970) defines third stochastic dominance as follows:

> Given two random variables, y_1 and y_2, y_1 third stochastically dominates y_2 if and only if
>
> (i) $\int(y_1 - y_2)\, dF(w) \geq 0$ and
>
> (ii) $\int w\, (\int t\, (y_1(z) - y_2(z)\, dz)\, dt) \geq 0 \ \forall w$. Whitmore proves that y_1 TSD y_2 if and only if $E[U(y_1)] \geq E[U(y_2)]$ for all U such that $U'''(w) > 0, \ \forall w \in W$.

The requirement that U''' be positive may be motivated by noting that this is a necessary condition for DARA. This additional condition, relative to the Hadar and Russell theorem for the SSD case, indicates that more distributions which can be ordered by the TSD criterion than by the SSD criterion. An even closer relationship between third stochastic dominance and DARA is obtained by Bawa (1975). He shows that for distributions with the same mean, third stochastic dominance is precisely equivalent to preference by all utility functions displaying DARA.

The relationship between the concepts of stochastic dominance and the Arrow-Pratt analysis may be summarised by saying that stochastic dominance relates to increases in the riskiness of situations while the Arrow-Pratt analysis relates to increases in the risk-aversion of individuals. This point is developed more fully by Diamond and Stiglitz (1974) who extend and integrate the increasing risk analysis of Rothschild and Stiglitz (1970,1971) with the Arrow-Pratt analysis of the characteristics of utility functions. They show that a type of duality exists between increases in risk and in risk aversion so that results relating to one concept can be paired with similar results relating to the other.

This duality is also apparent in the concept of risk-aversion with respect to a function developed by Meyer (1977a). Meyer's analysis works in terms of the risk-aversion coefficients rather than the utility functions themselves. The basic idea is to choose a function r(w) and to consider the set of agents with a utility function more risk-averse than that defined by r(w). If one prospect is preferred to another by all agents in this set then it is said to exhibit second stochastic dominance with respect to the function r(w). A refinement offered in Meyer (1977b) is to choose two functions $r_1(w)$ and $r_2(w)$, such that $r_1(w) \geq r_2(w) \; \forall \; w$ and then to consider the set of agents whose utility functions lie between r_1 and r_2 . It is easy to show that this approach includes FSD and SSD as special cases, but it does not appear to include the concept of third stochastic dominance. It should also be noted that Meyer defines r(w) to be the coefficient of absolute risk-aversion which makes the use of constant functions highly implausible. A logically equivalent, but simpler, approach would be to use the coefficient of relative risk-aversion.

Despite its power, the Arrow-Pratt characterisation of risk aversion suffers from a major weakness, in that it yields strong results only in the case of comparisons between a risky prospect and a certain alternative. Ross (1981) examines the case of a choice between two prospects y_1 and y_2, where y_2 is riskier than y_1 but has a higher return, that is, $y_2 = y_1 + \theta$, where θ is a random variable uncorrelated with y_1, but with positive mean. It is obviously desirable that if a given individual prefers the less risky option, the same should be true of any other person who is more risk-averse. However, as Ross shows, the Arrow-Pratt characterisation does not guarantee this. Ross offers a stronger characterisation of increasing risk-aversion, as follows.

Let A and B be two utility functions. Then A is **strongly more risk-averse** than B if and only if

$$(2.4.4)\ \inf_w\ A''(w)/B''(w) \geq \sup_w\ A'(w)/B'(w)$$

Ross demonstrates that this requirement is strictly stronger than the Arrow-Pratt characterisation, and derives stronger versions of concepts of decreasing absolute risk-aversion and increasing relative risk-aversion.

This section has given a summary of some the most important features of the vast literature relating to stochastic dominance between prospects and comparisons of risk-aversion between utility functions. A bibliography containing an extensive set of references on stochastic dominance has been provided by Bawa (1982), while a more general summary of the major developments in EU theory is given by Machina (1983a).

2.5 Applications of Expected Utility theory

EU theory has been applied to a wide range of problems within and outside economics and the range of applications continues to grow. Since almost all real-world problems involve some degree of uncertainty, the potential scope of the theory is enormous. It is useful to divide the economic applications of EU theory into three main categories.

The first class of applications consists of problems for which there is a well developed economic theory based on the assumption of perfect information (For most purposes, these theories can be extended to the case of uncertainty if actors are risk neutral, either by nature or because of the availability of costless risk-spreading through "perfect" capital markets). In these applications, interest centres on the extent to which standard results of theory under certainty may be carried over to the EU framework. One major area of interest has been the theory of saving and investment. In addition to modifying the theory of individual saving and investment behaviour, the application of EU theory has generated sharp debates over public investment policy. Whereas most economists working in an efficiency framework have taken the view that the return on public investments under certainty should be the same as that for competing private investments, there has been a strong argument that this is no longer true under uncertainty. Arrow and Lind (1970) argued that, because of the very great risk-spreading capacity of governments, discount rates for public investments should not include a premium for risk. Thus, assuming that private discount rates do include risk premiums, the public rate should be lower than the private one. This viewpoint has been sharply criticised by numerous writers. The debate is examined in Chapter 7. Another important application of this kind has been the theory of the firm under uncertainty, which is discussed in detail in section 2.6.

A second category of applications consists of problems which cannot be treated in any reasonable fashion without taking account of risk and risk-aversion. An obvious category is insurance. Since the expected value of an insurance contract is normally negative for the buyer (that is, the premium must cover both expected payouts and administrative costs), no theory which does not include a concept of risk-aversion can say anything useful about the demand for insurance. The general problem of insurance overlaps with the rapidly developing field of health economics, as health insurance is a major public policy issue in most Western countries. Arrow (1963) was a leading early contributor in this debate.

A third category of applications are those motivated directly by the development of EU theory. These include attempts to estimate utility functions for individuals and groups. For example, Officer and Halter (1968) and Bond and Wonder (1980), along with many others, have sought to estimate coefficients of risk aversion using questionnaires while Just (1975), Gallagher (1978) and others have derived econometric estimates of supply response under risk. There is also an extensive practical literature on stochastic dominance, employing the ideas of higher level stochastic dominance and dominance with respect to a function developed above.

2.6 The firm under uncertainty

The analysis of the behavior of the firm under uncertainty has been one of the most fruitful areas of application of EU theory. The case of the firm under output price uncertainty has been examined by Baron (1970), Sandmo (1971), Leland (1972) and Coes (1977). Hartman (1976) and Stewart (1978) have examined the impact of factor price uncertainty and Pope and Just (1977) the problem of production uncertainty. The problem of the existence of a competitive optimal output has been examined in Quiggin (1982b).

A wide range of maximisation problems have been considered for the firm under uncertainty. In this thesis, formal analysis is confined to single-output competitive firms although non-competitive firms have been examined by Leland and multi-output firms by Quiggin (1982b). The following general model covers most of the cases of interest. The technology is given by

(2.6.1) $x = f(\zeta, z, \theta)$

where:

x	is the firm's output;
ζ	is a vector of inputs variable in the short run;
z	is an input fixed in the short run; and
θ	is a random variable reflecting production uncertainty.

The firm's profit function is

(2.6.2) $\pi = rx - \mathbf{c} \bullet \zeta - B$,

where:

r is the output price;
$\mathbf{c}$ is a vector of input prices; and
$B = bz$ is the level of fixed costs .

All of the exogenous variables r,$\mathbf{c}$,b and θ are subject to variation over time. However, the crucial characteristics of the model depend on the distinction between instability and uncertainty (Quiggin and Anderson 1979). A variable is subject to instability if it varies over time, but to uncertainty only if its realisation (for a given time-period) is unknown at the time a particular decision is made. Thus, the characteristics of the model described above depend on which of the exogenous variables are subject to uncertainty. Three basic models may be distinguished in the literature. First, there is the case of the firm under output price uncertainty, examined by Baron, Sandmo,Coes and Quiggin. The only variable subject to uncertainty is the output price r and the firm's decision problem is to choose the output level x so as to maximise $E[U(\pi)]$, given cost-minimising levels of inputs. Second, there is the "short-run" case of input uncertainty, where θ and/or **c** are subject to uncertainty when inputs are chosen, but where z is already fixed. Thus, the problem is to choose ζ so as to maximise $E[U(\pi)]$. Third, there is the case where z must be chosen under uncertainty relating to some or all of the exogenous variables, but ζ is chosen after this uncertainty has been resolved.

The output price uncertainty model may be presented in a slightly simpler form than that given above. Since input choices are determined by cost-minimisation, it is possible to derive a cost function C(x) and write

(2.6.3) $\pi = rx - C(x) - B$.

It is assumed throughout that the firm's utility function U is globally concave in π. The output price r is a random variable and except where noted otherwise, it is assumed that it takes the form $r=\mu+k\gamma$ for some random variable γ with $E[\gamma]=0$. This notation, which permits the representation of additive shifts in the entire distribution of prices, and of linear increases in risk, is a slight modification of that used by Sandmo and Coes. Unless stated otherwise, it is assumed that $C''(x)>0$, and that a finite output exists which maximises expected profit.

The assumption of decreasing absolute risk-aversion (DARA) will also be used frequently. A utility function displays DARA if the coefficient of absolute risk-aversion, $r_a(\pi) = -U'(\pi)/U''(\pi)$, is decreasing in π. The firm's problem is to determine the optimal output, x^*, which maximises $E[U(\pi)]$.

The following are some of the main results in the literature:

Theorem 2.6.1 (Sandmo) The optimal output x^* is less than the profit-maximising output.

Theorem 2.6.2 (Sandmo) Assume that the function U is characterised by DARA. Then
(i) $\partial x^*/\partial B \leq 0$; and
(ii) $\partial x^*/\partial \mu \geq 0$.

Theorem 2.6.3 (Coes) Assume that the function U is characterised by DARA. Then $\partial x^*/\partial k \leq 0$.

For the final theorem the assumptions that C is convex and that a finite expected profit-maximising output exists are dropped. It is necessary to define:

(2.6.4a) $U_1 = \lim_{\{\pi \to \infty\}} U'(\pi)$,

(2.6.4 b) $U_2 = \lim_{\{\pi \to -\infty\}} U'(\pi)$,

(2.6.4c) $C^* = \lim_{\{x \to \infty\}} C'(x)$

Then the criterion for existence of a finite optimum is given by

Theorem 2.6.4 (Quiggin):

Let

$$(2.6.5)\ K = U_1 \int_{-\infty}^{C^*} (p - C^*)\, dF(p) - U_2 \int_{C^*}^{\infty} (p - C^*)\, dF(p)$$

Then a finite optimum exists if $K < 0$ and not if $K > 0$.

Corollary: A finite optimum exists whenever $\Pr\{p < C\} > 0$ and $\lim_{\{\pi \to \infty\}} U'(\pi) = 0$.

A number of other results could be examined. For example, the theorems listed here are mainly concerned with comparisons between probability distributions. It would also be possible to compare firms with different levels of risk-aversion as in the work of Diamond and Rothschild. This yields a range of results dual to some of those listed here. For example, Theorem 2.6.1 may be rephrased as saying that the output of a risk-averse firm will be less than that of a risk-neutral one. However, the set of results listed here are sufficient to give the flavour of much of the work which has been done in this field, and the analysis given below will suggest ways of generalising other results.

2.7 Expected Utility theory as a scientific research program

The discussion in this section is based on Lakatos' (1970) concept of "scientific research programs". This concept is related to Kuhn's (1962) notion of paradigms, but the two are distinguished by Kuhn's insistence, based on gestalt ideas, that paradigms are fundamentally incommensurable, and that it is impossible to think in terms of two paradigms at once. While the paradigm idea has some appeal for large-scale conflicts between scientific theories (eg between neoclassical and Marxist economics), the concept of scientific research programs seems more appropriate to the problems of utility theory.

Lakatos points out that a scientific theory can never be refuted solely by contradictory empirical observations. Such observations may always be explained by faulty experimental techniques or the failure of ceteris paribus conditions. Lakatos argues that theories must be assessed in terms of a scientific research program consisting of a "hard core" of maintained hypotheses which are not susceptible to refutation within the framework, and a "protective belt" of testable hypotheses which are adjusted in the light of new empirical evidence. Research programs are characterised by a negative heuristic which excludes hypotheses inconsistent with the core, and by a (less clearly articulated) positive heuristic which suggests the type of work which should be done to generate and test refutable hypotheses. Lakatos distinguishes sharply between progressive and degenerating research programs. In the former, the empirical content of the theory tends to increase over time and refutations of individual predictions are met with problem-shifts which yield unexpected and novel predictions. In the latter, reaction to evidence consists largely of ad hoc auxiliary theories which serve as immunising strategies, protecting the theory from any possibility of refutation. However, a research program, even a degenerating one, will never be abandoned unless there is a competing research program in the field which appears to offer better prospects for progress.

For followers of the EU approach, the hard core consists of the concepts of rationality embodied in the EU axioms, with the possible addition of the postulate of risk-aversion, and the negative heuristic consists of a rejection of approaches inconsistent with these concepts. The protective belt consists of hypotheses such as that of declining absolute risk-aversion and empirical theories relating to such phenomena as insurance and portfolio choice. The positive heuristic involves such ideas as " examine situations for which there is an established economic theory of behavior under certainty and see how far these results carry over to the case of uncertainty" and "seek conditions on utility functions and probability distributions under which sharp behavioral predictions can be made".

There is no doubt that the EU approach has generally been a progressive research program. Analytical tools such as stochastic dominance conditions and characterisations of different forms of risk aversion have been developed steadily. The extension of EU analysis to problems such as the theory of the firm has led to the generation of a number of novel and testable hypotheses. In addition, EU theory has permitted the extension of economic analysis to important problems such as insurance for which no adequate treatment was previously available. Empirical research based on the hypotheses suggested by EU analysis has been extensive and has given support to important hypotheses such as decreasing absolute risk aversion.

Despite this progress, however, there has been a steady increase in the number of 'anomalies' associated with EU theory and particularly with the independence axiom. The auxiliary hypotheses used by followers of the EU approach to explain these problems have become increasingly threadbare. The way is, therefore open for an alternative research program. The remainder of this thesis is devoted to consideration of the anomalies which have developed within EU theory and proposals for alternatives.

3. ANOMALIES IN EXPECTED UTILITY THEORY

3.0 Introduction

The object of this chapter is to give an account of the observed violations of EU theory, and to describe the main responses made by users of EU theory. It will be argued that violations of the theory are widespread and consistent and that, in a number of important cases, the responses of EU theorists are content-decreasing. Thus, despite the vigorous work going on within the EU framework, it may be regarded in certain senses as a degenerating research program. Problems in EU theory fall into three main categories. First, there are observed patterns of market behavior which appear to contradict the predictions of the theory. The best-known example is that of the coexistence of gambling and insurance. However, a number of other aspects of observed insurance are, at least arguably, inconsistent with EU theory. Second, there are choice problems used in experimental settings for which many subjects give responses inconsistent with the EU axioms. The best-known and most widely confirmed of these is the "Allais paradox", produced as a counterexample to the EU independence axiom. A wide range of similar problems has been observed. Third, there are problems which have emerged in the elicitation of utility functions using questionnaires. Some methods for constructing utility functions have produced thoroughly inconsistent results, while others have shown evidence of consistent bias. All of these problems are discussed in this chapter. Explanations of the apparent anomalies in terms of EU theory will be examined and it is argued that, at least in some cases, these cannot be considered adequate. The independence axiom will play a major role in the discussion. Many of the most intractable problems associated with EU theory are closely related to the independence axiom, and much of the debate over EU theory has turned on this axiom. It is important to establish the weakness of this axiom on both normative and descriptive grounds and to consider possible alternatives.

3.1 Gambling

One of the first major problems to emerge in EU theory derived from observed behavior in markets involving uncertainty such as those for insurance and gambling. Gambling, as well as being the inspiration for the development of probability theory, is an economically significant aspect of uncertainty. According to the Australian Bureau of Statistics Household Expenditure Survey (ABS 1983), Australians spend almost as much on gambling (net of winnings) as they do on insurance (gross). Yet gambling is inconsistent with EU theory unless the very plausible hypothesis of risk-aversion is abandoned, and the coexistence of gambling and insurance purchase is very difficult to fit into the EU framework.

Some types of gambling can be explained easily enough. For example, some betting on horse races may reflect divergent beliefs about the outcome. This is reflected in the large amounts of effort devoted to collecting and analysing information about the quality of horses, jockeys, tracks etc. On the other hand, there are a large number of racetrack bettors who collect no information at all, and bet on an essentially random basis.

Some types of gambling may be explained as entertainment activities rather than as financial decisions. For instance, given the virtual certainty of losing on slot machines in the medium to long run, it seems reasonable to assume that this activity is undertaken largely for entertainment. Similar explanations may be advanced for participation in various time-consuming gambling games such as bingo.

In respect of lottery tickets, however, the only plausible reason for betting is the chance of winning a large amount of money. The predominance of this motive is confirmed by psychological studies (Walker 1984). The purchase of lottery tickets by people who are generally risk-averse constitutes a significant problem for EU theory, which is very difficult to explain without resorting to the content-decreasing measure of excluding gambling from the realm of rational behavior.

At least some elements of race-track betting would appear to fall into the same category. A series of studies from Griffith (1949) to Ali (1977) has shown that bettors are willing to pay a higher margin for long-priced horses. That is, the expected loss from consistently backing long-priced horses is quite large (similar to the loss invovled in buying lottery tickets) while the expected loss for short-priced horses is very small (about 2 cents for every dollar invested). Indeed, some studies have actually suggested that a strategy of routinely backing horses at odds-on would lead to small positive profits, though this seems unlikely.

Friedman and Savage (1948) attempted to explain the coexistence of gambling and insurance using the concept of an S-shaped utility function . The basic idea is that people will be risk averse with respect to changes in wealth within a neighborhood of their current wealth level but may be risk-seeking with respect to prospects which may take them into a higher social class. Thus, the utility function may be regarded as concave for low income levels and convex for some higher incomes. Friedman and Savage introduced a third concave segment, at still higher income levels, as a response to the observation that lotteries typically have multiple prizes whereas convexity throughout the upper range would imply that a lottery with a single large prize would be preferred. (Figure 3.1).

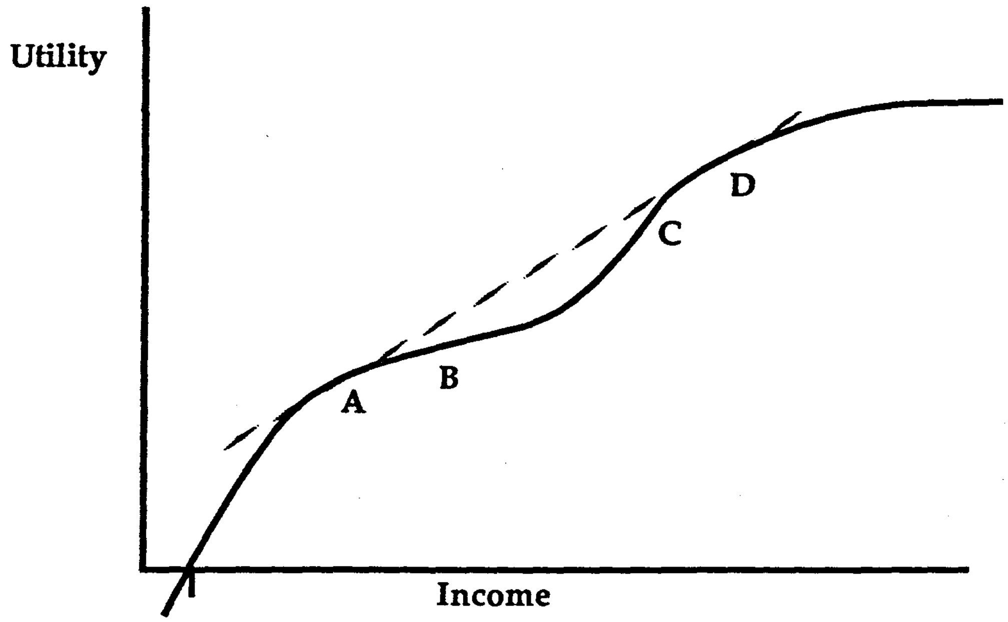


Figure 2.1 Friedman-Savage utility function

A variation on this theme, expressed in terms of indivisibilities in expenditure, is developed by Ng (1965). He argues that, even if the underlying utility function is concave, the existence of large indivisible items of expenditure, such as the purchase of a university education, will lead people to buy lottery tickets. It is obviously open to question how many items of expenditure are in fact indivisible in the sense required by Ng. Most expensive goods have cheaper substitutes - for example correspondence courses in place of full-time university training, or motor-bicycles in place of cars. For each individual type of good, broadly defined, it is possible to define a marginal utility of expenditure. If it is the case, for a wide range of goods, that this marginal utility is increasing, then it seems reasonable to say that the marginal utility of wealth is increasing. As an aside, it may be noted that Friedman and Savage justify their own proposed utility function by arguing that sufficiently large increments in wealth will allow individuals to move, say, from the working to the middle classes. This argument has an obvious affinity with that offered by Ng.

A slightly different approach to the problem, based on imperfections in capital markets, has been taken by Flemming (1969), Hakansson (1970), Kim (1973). The basic point which is raised by these writers is that the utility function over income levels is dependent not only on the utility of consumption, but also on the capital market opportunities open to people. For example, given a perfect, fully informed, capital market and a wide range of uncorrelated risks, everyone would behave as if they were risk neutral, since they could costlessly spread risk on the capital market. This insight may be developed in a number of ways. Kim's approach is perhaps the most interesting. He assumes that the underlying utility function is linear, but that, because of finance market imperfections, the effective interest rate for borrowing increases as wealth falls, while the effective rate for lending increases as wealth rises. Thus, one will be willing both to gamble on lotteries with large prizes and to insure against large losses.

None of these approaches appears adequate as explanations of gambling in a countries, such as Australia, where a large proportion of people buys lottery tickets despite the existence of well-developed credit markets. Even allowing for considerable costs associated with imperfections in credit markets, it is unlikely that these could justify the purchase of lottery tickets with an expected return of about 60 cents for each dollar in outlays.

Returning to the basic Friedman-Savage approach, Yaari (1965) observed that the non-concavity of U implied that the set of bets which would be accepted in preference to a given certain gain was not convex, in contrast to the normal properties of demand for goods. He undertook a number of tests and concluded that acceptance sets were indeed convex. Since the gambling behavior involved was inconsistent with EU theory, Yaari concluded that some form of probability weighting was involved.

A number of other objections to the Friedman-Savage approach have been made by Machina (1982). The most important is that the observed gambling behavior of individuals does not appear to change radically in response to changes in their initial wealth. However, the utility function in Figure 3.1 suggests that behavior will be sensitive to changes in initial wealth, with only those individuals near the inflexion point displaying propensities to both gamble and insure.

There is an even more fundamental problem with the Friedman - Savage model, which does not appear to have been observed previously. This point is developed from the argument against the plausibility of risk-preference given in Section 2.2. A broadly similar argument applies to people with both concave and convex segments in their utility functions. Mutually beneficial gambles will always be available to people whose present wealth lies in a convex segment of the curve. Even in a concave segment of the curve, people will tend to gamble until they reach a point at which they are risk-averse in the sense that they would refuse all risky gambles with expected outcome zero or less. This condition may be expressed by saying that, at a stable outcome x_0, the utility function "looks concave" in the sense that, for any x_1, x_2 such that $x_0 = \lambda x_1 + (1-\lambda)x_2$,

$$(3.1.1)\ U(x_0) > \lambda U(x_1) + (1-\lambda)\, U(x_2).$$

Friedman and Savage indicate at least partial awareness of this point but they do not appear to realise that their model implies gambling behavior radically different from that observed in practice. Whereas most people only buy one or a few lottery tickets at a time, the model implies that people should gamble continuously until their wealth falls below point A, or they win a prize which takes wealth above point D, at which point they should stop altogether. The change in utility arising from one unit of expenditure on a fair lottery may be written, in terms of changes from the initial wealth level, as (1/n U(n)) - U(-1)) . Since this value will be approximately constant for small changes in the initial wealth level, it is apparent that, if it is worthwhile to buy one lottery ticket, it will be worthwhile to continue buying tickets until initial wealth has changed enough to substantially alter the values U(-1) and U(n). This feature of gambling behavior in the Friedman-Savage model implies extreme sensitivity to initial wealth. An increase in wealth will be spent exclusively on lottery tickets until a point, such as C, is reached. Conversely, a reduction in wealth will affect only expenditure on lottery tickets until this is reduced to zero.

Similar arguments apply in the case, examined by Ng (1965), where a person with a concave utility function is faced with consumption indivisibilities. Once again, lottery tickets should be purchased until the marginal utility of wealth falls below the expected utility of the prize, so gambling behavior should be highly sensitive to initial wealth.

All of the approaches cited above have focused on the value of the outcomes. This is a natural consequence of the use of the EU framework, but it seems far more reasonable to suppose that participation in lotteries has to do with attitudes to probabilities and, in particular, with the placing of a high weight on extremely favorable, low probability events. Adam Smith (1776, pp 164-165) stated:

> "That the chance of gain is naturally over-valued, we may learn from the universal success of lotteries. The world neither ever saw, nor ever will see, a perfectly fair lottery; or one in which the whole gain compensated the whole loss; because the undertaker could make nothing by it. In the state lotteries, the tickets are not worth the price which is paid by the original subscribers, and yet commonly sell in the market for twenty, thirty and sometimes forty per cent advance. The vain hope of getting some of the great prizes is the sole cause of this demand. The soberest people scarce look upon it as folly to pay a small sum for the chance of gaining ten or twenty thousand pounds; though they know that even that small sum is perhaps twenty or thirty per cent more than the chance is worth. In a lottery in which no prize exceeded twenty pounds, though in other respects it approached nearer to a perfectly fair one than the common state lotteries, there would not be the same demand for tickets. In order to have a better chance for some of the great prizes, some people purchase several tickets, and others, small shares in a still greater number. There is not, however, a more certain proposition in mathematicks, than that the more tickets you adventure upon, the more likely you are to be a loser. Adventure upon all the tickets in the lottery, and you lose for certain; and the greater the number of your tickets the nearer you approach to this certainty"

This passage encapsulates a number of objections which have been made to the EU explanations of gambling and a number of requirements for a succesful explanation. In particular, it is necessary that the theory should explain preference for lotteries with a few large prizes, that it should not suggest that people are unaware that the game in which they are participating is not (actuarially) fair, and that it should explain the purchase of one or a few tickets.

EU theory, even with the abandonment of the plausible postulate of risk-aversion, cannot account for the observed behavior of a large segment of the population in undertaking both gambling and insurance. What is needed is a model of choice under uncertainty which takes explicit account of phenomena such as probability weighting.

3.2 Insurance

In the previous section, the purchase of insurance was treated as evidence of risk-aversion, and hence as suggesting that gambling could not be explained in terms of EU theory. A number of other aspects of insurance decisions have been suggested as contrary to the predictions of EU theory, though, in my view, the evidence here is weaker, and it is usually necessary to rely on a mixture of market and experimental evidence. In particular, because of the complexity of insurance decisions, it rarely possible to obtain clear-cut results from observations of market behavior alone. Apparent deviations of market behavior from the predictions of EU theory can, however, serve as a guide to the design of appropriate experimental studies.

Kahneman and Tversky (1979, p269) describe an experiment involving **probabilistic insurance** cast in the following terms.

> "Suppose you consider the possibility of insuring some property against damage eg fire or theft. After examining the risks and the premium you find that you have no clear preference between the options of purchasing insurance or leaving the property uninsured. It is then called to your attention that the insurance company offers a new program called probabilistic insurance. In this program, you pay half of the regular premium. In the case of damage, there is a fifty per cent chance that you pay the other half of the premium and the insurance company covers all the losses; and there is a 50 per cent chance that you get back your insurance payment and suffer all the losses. For example, if an accident occurs on an odd day of the month [1], you pay the

1 Pedants will wish to note that rather more than half the days in the year are odd-numbered.

other half of the premium and your losses are covered; but if the accident occurs on an even day of the month, your insurance payment is refunded and your losses are not covered. Recall that the premium for full coverage is such that you find this insurance barely worth its cost. Under these circumstances would you purchase probabilistic insurance?"

When this problem was presented to a group of 95 Stanford University students, 80 per cent of theme said they would not purchase probabilistic insurance. Kahneman and Tversky (1979, p270) point out that this modal choice is inconsistent with EU theory because "if at asset position w one is just willing to pay a premium y to insure against a probability p of losing x, them one should definitely to pay a smaller premium ry to reduce the probability of losing x from p to (1-r)p." This is basically because, given concavity, $U(w) - U(w-ry) < r(U(w) - U(w-y))$. However, as long as y is small, this difference will be of the order o(y), and is likely to be outweighed by the fact that insurance purchases involve certain fixed transactions costs. To put it simply, if full insurance is barely worth the bother, probabilistic insurance is not worth the bother at all.

The refusal to accept probabilistic insurance constitutes evidence against EU theory, but in view of the foregoing argument, not very strong evidence. Moreover, there is the problem that the concept is unfamiliar and hence likely to be misunderstood. Observations on actual market behavior or on experimental problems which are more closely modelled on real life are likely to be of somewhat more interest.

Perhaps the most surprising claim to come out of studies of insurance decisions is the claim that people prefer insurance against high probability, low loss events to insurance against low probability, high loss events when the expected losses and premiums are equal. This claim was advanced by Slovic et al (1977), and followed the earlier results of Edwards (1962). It was supported by an empirical study of flood and earthquake insurance decisions (Kunreuther et al 1978). This specific claim has been elaborated into a more general claim that people are **risk-takers in the domain of losses.** A detailed experimental study of the problem was undertaken by Schoemaker and Kunreuther (1979).

They presented two groups of subjects with binary choices of the form

A: a 6 out of 10 chance of losing $100

B: a 1 out of 100 chance of losing $6000,

where each of the two prospects had the same (negative) expectation. In each case of this kind, EU risk averters should choose the safe alternative A. By contrast, the Slovic et al theory would suggest that the risky option B should be preferred. Among the first group of subjects, consisting of students a majority chose the risky alternative B, the average over three problems being 50% for B with 40% for A and 10% undecided. By contrast, in the second group, which consisted of randomly selected clients of an insurance agency, the preferences were reversed. On average, on the same set of three problems, 63% preferred the safe alternative.

It is not immediately clear which of these contradictory results is to be given the most weight. On the one hand, since the 'clients' are people who have actually bought insurance, they may not be a fair sample of the population (indeed, a rigorous application of the hypothesis of risk-taking in the domain of losses would suggest that this behavior pattern is incompatible with insurance). On the other hand, since few of the students actually had $6000 to lose (and presumably, for some, even losses of the order of $100 would imply a major financial adjustment) it is difficult to know whether the suggested outcomes would fall within the domain of their utility functions.

An important qualification of experimental results tending to support the 'risk-aversion in the domain of losses' claim is that experimental settings may encourage a gaming atmosphere in which the major objective is to maximise the probability of winning, that is, of finishing ahead. This criticism has been made specifically in relation to the Slovic et al (1977) results by Schoemaker and Kunreuther (1979).

A more solidly confirmed piece of evidence on insurance behavior is the widespread reluctance to accept contracts involving excess clauses, deductibles etc. when full insurance is available, even at actuarially unfair prices (Eisner and Strotz 1961, Pashigian, Shkade and Menefee 1966). In this case, it is straightforward to show that partial insurance should be preferred, since people should approach risk-neutrality as the potential loss approaches zero. It is not clear, however, that 'full insurance' is really complete. For example, medical insurance may pay for hospital bills but not for lost wages and incidental costs, and it almost never compensates for pain and suffering.Thus, it may be rational in EU terms to seek full medical insurance or even to over-insure against medical bills if this is possible.

A final piece of evidence relates to the bundling of insurance options. EU theory predicts that the price people are willing to pay for insurance against any of three mutually exclusive risk should be equal to the sum of the prices they are willing to pay for any one. By contrast, Slovic et al (1977) found it was easier to sell a package consisting of insurance against a high-probability low loss event and a low-probability high-loss event, than to sell either separately. The type of event insured here appears to be crucial.

In summary, while many of the apparent violations of EU theory in empirical choices can be explained (eg by reference to information problems and context effects), they are sufficiently widespread to suggest that an alternative theory capable of explaining them would be useful. At the very least, it would be desirable to have a theory of rational choice of which EU was a subset, and which embodied plausible representations of some of the non-EU behavior patterns which have been suggested. In this way, instead of setting up problems in which violations of EU might or might not be observed, it would be possible to estimate a consistent set of preferences and test the restrictions implied by EU theory.

3.3 Experimental violations of EU theory - the Allais paradox

Experimental tests have the advantage that subjects can be presented with a wide range of decision problems couched in terms of probabilities, thus permitting easy analysis of the results. The main problems are, first, that the problems do not normally deal with real money (see, however, the work of Binswanger), and second, that it is difficult for experimenters to avoid imposing their own preferences. A wide range of experimental violations of EU theory have been observed. The earliest and most famous is the well-known **Allais paradox**. The subject is first asked to choose between

A_1 $1m with certainty

and

B_1 $1m with probability 0.89, $5m with probability 0.1, zero with probability 0.01,

and then between

A_2 $1m with probability 0.11 and

B_2 $5m with probability 0.1

Most subjects choose A_1 and B_2, a decision which is inconsistent with any EU utility function. Allais' problem was chosen as a counter-example to the controversial independence axiom, although, as has been shown by Segal (1984), it can be consistent with this axiom if other EU axioms are dropped. It is intuitively apparent that the critical feature of the paradox is the 0.01 possibility of receiving nothing in prospect B_1. Because this outcome is, in some sense, overweighted, choices inconsistent with EU theory are generated.

Although the classic presentation of the Allais paradox involves large sums of money, similar choices have been elicited in less extreme situations. MacCrimmon and Larsson (1979) reported a series of experiments in which attempts were made to test the prevalence of Allais-type violations of the EU axioms.

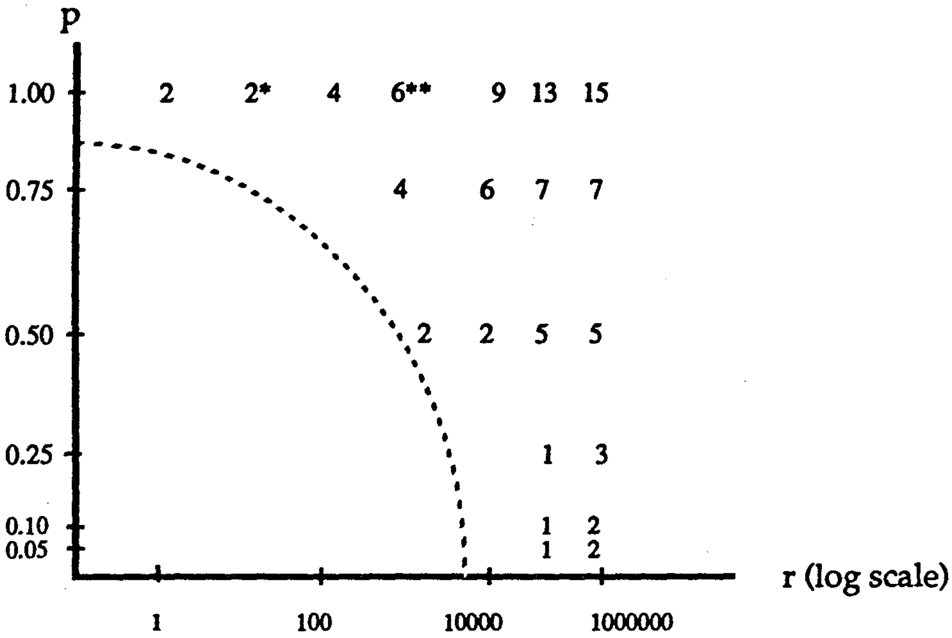
MacCrimmon and Larsson considered a range of problems with three payoff parameters r,s and t, with $r > s > t$, and a probability parameter p, with $0.11 < p < 1$. The choices may then be cast in the form

A {(s, t); (p,1-p)}

B {(r, s, t); (0.1,p-0.11, 1.01-p)},

and the Allais-type behavior is elicited by examining the choice with two different values of p. Thus, the standard Allais paradox has r = $5m, s = $1m, t = 0, $p_1 = 1$, $p_2 = 0.11$. This problem is referred to as the common consequence problem, since changes in p affect only consequences which are common to the prospects A and B, and thus should not, according to the independence axiom, alter the preference ranking between them.

MacCrimmon and Larsson examined values of s ranging from $1,000,000 to $10,000 and values of p_1 and p_2 ranging from 1 to 0.11, while holding t = 0 and r = 5s. On average about one-third of their subjects exhibited Allais-type violations, with the number of violations tending to decline with the size of the payoff, s, and therefore also with r = 5s. Their results are summarised in Figure 3.2.



*One subject chose A on one presentation of this set, and B on the other

** Two subjects chose A on one presentation of this set and B on the other

Figure 3.2: The number of subjects, out of 19, selecting the alternative (r,p) over the alternative (5r, 4/5p).

Segal (1986) defined the Generalised Allais Paradox (GAP) as follows: ' Let A,B, C, D be lotteries such that C and D stochastically dominate A and B resp, and $F_D - F_C \equiv F_B - F_A$. Assume moreover, that B differs from A by a simple compensated spread, and let x* be such that ,for $x \leq x^*$, $F_B(x) \geq F_A(x)$ and $x \geq x^*$, $F_B(x) \leq F_A(x)$. If, for $x \geq x^*$, $F_C(x) = F_A(x)$ (and $F_D(x) = F_B(x)$), then C is preferred to D. The standard Allais paradox arises with $A = A_2$, $B=B_2$, $C=A_1$, $D=A_2$. The question of whether GAP is a reasonable description of preferences is discussed in section 5.5.

3.4 Other experimental violations

Although the Allais problem is the best-known, a wide range of experimental violations of the EU approach, have been observed including ambiguity effects (Ellsberg 1961)the **common ratio effect** and **reflection effect** (Kahneman and Tversky 1979, Machina 1983), the overweighting of small outlying probabilities, and the phenomenon of **preference reversal** (Lichtenstein and Slovic 1971). Ellsberg's analysis of ambiguity has already been discussed. The other violations will be described briefly (see MacCrimmon and Larsson 1979 for extensive empirical work and Machina 1983 for a survey).

The following example of the common ratio effect is given by Kahneman and Tversky. Subjects were asked to choose between a 0.90 chance of winning $3000 and 0.45 chance of winning $6000, and then between a 0.002 chance of winning $3000 and a 0.001 chance of winning $6000. The ratio of the probabilities is the same in both cases, so according to EU theory the same outcome should be chosen each time. Moreover, risk-aversion implies that the prospect with the $3000 outcome should be preferred in both cases. In fact, 86 per cent of people made this choice in the first problem, but only 27 per cent in the second problem. MacCrimmon and Larsson generalised this problem to cover problems of choice pairs $A = \{(0, r);(1-p, p)\}$, $B = (0,ar;1-bp,bp)$, where p varies between choice pairs. Thus in the example cited above, $r = \$3000, a = 2, b = 0.5, p_1 = 0.9, p_2 = 0.002$. MacCrimmon and Larsson fixed a = 5, b = 0.8 and allowed r to vary from $1m to $1 while p took on the values 1.00, 0.75, 0.25, 0.1, and 0.05. They also examined negative values for r ranging from -$1000 to -$1, with p taking on the values 1.00, 0.80, 0.20, and 0.04.

In general, the proportion choosing alternative A declined with p and with the absolute value of r. Overall, about 65 per cent of the subjects exhibited consistent violations of the independence axioms. There are, then, two possible interpretations of the common ratio effect. The first is the simple observation that people presented with common ratio choices will not, in general, make choices consistent with EU theory. The second is the MacCrimmon-Larsson interpretation, suggesting that the proportion choosing the higher probability will decrease with p and with the size of the payoff r. the extent to which the data justifies the MacCrimmon-Larsson interpretation, and possible generalisations of it, is discussed in Chapter 5.

A closely related phenomenon is the reflection effect observed by Kahneman and Tversky (1979). This involved the presentation of a series of pairs of choices. In each pair, the first choice consisted of two prospects of the form $\{(w_1, w_2);(p_1, 1-p_1)\}$. The second pair consisted of the same two choices except that the w_i were replaced by $-w_i$, that is, gains were replaced by losses. In a number of cases observed by Kahneman and Tversky the modal choices were inconsistent with EU risk-aversion. The really striking feature of their results was the fact that, in each case, the modal choice pattern was reversed. Unfortunately, Kahneman and Tversky do not state what proportion of their subjects actually displayed the reflection effect. It should be noted that it is possible for the modal pattern to be reversed even while a majority of subjects chooose consistently. However, in a number of cases, the majority adopting the modal pattern was so great that it is clear that a majority of subjects must have displayed a reflection effect.

It may also be noted that the results obtained by Kahneman and Tversky do not support any general claim of risk-seeking in the domain of losses. Rather there is a tendency to prefer the lower probability event in a comparison of two high probability losses and the higher probability event in a comparison of two low probability losses.

The phenomenon of preference reversal was first observed by Lichtenstein and Slovic (1971). They asked individuals to choose between appropriately selected pairs of lotteries. Having done so they were asked to state the lowest price they would be willing to accept in return for their right to participate in the lottery. Many respondents set a lower price on the preferred lottery. These results were replicated by Grether and Plott (1979), Pommerehne, Schneider and Zweifel (1982) and Reilly (1982).

This result appears, at first sight, to be inconsistent with any sort of consistent preference ordering. In fact, however, as has been shown by Karni and Safra (1985), the way in which the "reservation price" was elicited means that this price is not equal to a certainty equivalent. The participants were informed that, after they had set their prices, a random sum would be chosen from a uniform distribution over some range (eg [$0, $10] with increments of $0.01). If the sum of money selected in this way exceeded the price set by the participant, then the money was paid and the lottery forgone. Otherwise the lottery went ahead, and the participant received the indicated price. This procedure was presumably adopted to prevent participants mis-stating their evaluations for strategic reasons. However, it has the effect of estimating the certainty equivalent by means of a probabilistic reservation price. This will equal the certainty equivalent under EU theory, but otherwise it need not. Thus, the claim by Grether and Plott (1979) that preference reversal represents a violation not merely of EU theory but of basic notions of transitivity, cannot be upheld.

A major feature of most of the violations which have been observed is the apparent overweighting of low-probability outlying events. This has also been observed directly in many psychological experiments aimed at eliciting subjective probabilities used in decisions, such as those of Edwards (1954). This has been a major point of departure for the alternatives to EU theory which are examined in subsequent chapters.

3.5 Studies of probability weighting

While economists working in the EU tradition have generally either ignored probability preferences or treated them as a noise factor to be minimised while attempting to determine utility functions, a number of psychologists have attempted to estimate probability attitudes. The most intensive period for activity in this field was the late 1940s and 1950s ; that is, immediately after the development of EU theory. Unfortunately, the evidence is not as conclusive as might be hoped and activity since about 1960 has been sporadic, with few attempts to make a systematic study over a wide range of probabilities and outcomes.

Preston and Baratta (1948) estimated probability weights by comparing hypothetical 'bids' for risky gambles with their mathematucal expectation. Thus, if someone bid $305 for a prospect of the form {(0,500);(0.25, 0.75)}, they would be assumed to have a probability weight for the winning event of 0.61. Since this procedure takes no account of the possibility of declining marginal utility, it is likely to lead to systematic under-estimation of probability weights. Despite this, bets with a small probability of winning were systemantically overvalued, relative to their objecctive probability. The estimated weight for bets with higher probabilities was well below the objective probability with the crossover occurring around 0.2. A comparable result was found in the study of racetrack betting by Griffiths (1949), where the crossover point was 0.16. Once again, the design of the study was such that utility and probability effects could not be distinguished easily.

Mosteller and Nogee (1951) undertook an extensive series of trials in which real gambles for small sums were offered. This procedure is less likely to be confounded by the effects of declining marginal utility. Mosteller and Nogee had two groups of subjects - Harvard undergraduate students and National Guards. The students generally followed a strategy close to expected value maximisation with a slight tendency to conservatism. On average, their estimated weight was below the objective probability for all odds. The National Guards, on the other hand, gave results similar to those of Preston and Baratta, except that the crossover point was about 0.5. This may reflect the absence of risk-aversion effects.

The most extensive experiments were those undertaken by Edwards (1953, 1954a,1954b) and discussed further in Edwards (1962). In all of these studies, Edwards followed the same basic procedure. A set of eight gambles, each with two possible outcomes, was devised in which the probability of winning rose steadily from 1/8 to 8/8. In some, but not all, cases the expected value was constant. Subjects were presented with all possible pairs of gambles, in a random order, and asked to choose between them. For each gamble Edwards reported the proportion of times it was chosen. This proportion ranged from 0 (if it was never chosen) to 0.25 (if it was chosen whenever it was available), while the average was 0.125. Other aspects of the situation were varied systematically. For example, payoffs were alternatively hypothetical amounts, worthless chips, real money and imagined gains in a military strategy game. Expected values were variable in some cases and fixed in others, and were chosen to be positive, negative or zero.

A number of features of Edwards' design must be commented on. First, it is unfortunate that no probabilities smaller than 1/8 were examined. From the other experiments reported here, it is apparent that probability weighting effects are likely to be strongest where the probability of wining (or losing) is below 0.1. Second, because subjects knew they were to be faced with a large number of gambles it is difficult to know whether they evaluated each gamble 'as if' it was an isolated source of risk, or whether they followed a strategy designed to yield a preferred 'portfolio'. It would be quite difficult to describe precisely the decision context in which these choices were made.

Edwards observed three major patterns in his data. These were preference for winning probabilities of 1/2, aversion for winning values of 6/8 and preference for low-probability high loss bets when expected values were negative. The first two patterns applied for positive and zero EV bets. For negative EV bets there was a strong aversion to probabilities of 1/2 and a strong preference for losing probabilities of 6/8. The results for three real-money sessions (with different subjects) are given in Table 3.1. The first two patterns show up clearly enough on all the positive and zero EV sessions. However, the negative EV results are much less clear-cut. The preference for low probability high loss gambles is strong in Experiment 1, weaker in Experiment 2 and non-existent in Experiment 3.

	PEV			NEV			ZEV		
	1	2	3	1	2	3	1	2	3
Probability *									
1/8	0.1116	0.1674	0.1347	0.1957	0.1585	0.1177	0.1399	0.1897	0.1518
2/8	0.1265	0.1339	0.1356	0.1729	0.1629	0.1250	0.1592	0.1696	0.1607
3/8	0.1376	0.1607	0.0909	0.1652	0.1327	0.1688	0.1562	0.1473	0.1144
4/8	0.1830	0.1942	0.2175	0.1116	0.0982	0.0706	0.2009	0.1763	0.2045
5/8	0.1406	0.1207	0.1282	0.1049	0.1161	0.1331	0.1339	0.1161	0.1315
6/8	0.0900	0.0603	0.0747	0.1265	0.1719	0.1672	0.0677	0.0692	0.0820
7/8	0.0975	0.0893	0.0950	0.0871	0.1183	0.1315	0.0804	0.0670	0.0552
8/8	0.1131	0.0915	0.1234	0.0372	0.0424	0.0860	0.0618	0.0647	0.0998

Table 3.1 Results of three experiments on probability preferences

There is some support for the notion of a preference for 50-50 bets from Coombs and Pruitt's (1960) study of 'variance preference'. Coombs and Pruitt attribute this preference to the supposed 'fairness' of these bets (compare the popular view that two-up[2] is 'the fairest game in the world'). However, their results from an explicit study of probability preferences presented in the same paper do not bear this out. Instead, a great majority of subjects preferred either a very high probability of winning or a very low probability. The first of these patterns would be consistent with EU and the second with overweighting of low-probability events.

In general, Edwards' results seem to give fairly strong support to the notion of overweighting of small probabilities, at least for positive and zero EV bets. This support is strengthened by consideration of responses to bets with differing EV (Edwards 1954a). Acceptance of these bets showed a general tendency to rise with EV. In all cases,

2 A traditional Australian coin-tossing game in which all bets are 50-50

however, bets in which the more preferred outcome occurred with a probability of 1/8 were selected more frequently than EV maximisation would suggest, and those in which the less preferred outcome had a probability of 1/8 were selected less frequently.

In summary, the psychological evidence tends to support the view that probability weighting is important and that low probability events in two-outcome gambles tend to be overweighted. However, the value of the evidence is reduced by small population samples, general absence of independent replication and the lack of an adequate theoretical framework. More experimental evidence on the basic issues is required.

3.6 Construction of utility functions

The final category of violations of EU theory arises from attempts to estimate utility functions. This evidence is particularly important because it arises from attempts to apply EU theory rather than, as in the case of the Allais paradox, from attempts to refute it. Thus, there has been a conscious attempt to avoid postulating extreme or unrealistic circumstances.

The first approach used to estimate the utility function, pioneered by von Neumann and Morgenstern themselves, was based on the 'standard reference contract' . Here the outcomes x_1 and x_2 of a risky prospect are held constant. For any x_3, such that $x_1 \, P \, x_3 \, P \, x_2$, the questioner elicits the probability p such that x_3 is the certainty equivalent of $\{(x_1,x_2); (p,1-p)\}$. The utility function is then constructed on the basis that

(3.6.1) $U(x_3) = pU(x_1)+(1-p)U(x_2)$,

given some arbitrary initial choice such as $U(x_1)=0$, $U(x_2)=1$. This method has generally performed poorly (Dillon 1971).It is apparent that it depends crucially on the assumption that preferences are 'linear in the probabilities'. If any form of probability weighting is involved, this method can lead to hopelessly inconsistent results.

As a result of these difficulties, a number of writers have advocated the use of fixed probabilities with variable outcomes (Davidson Suppes and Siegel 1957, Anderson, Dillon and Hardaker 1977). One approach of this kind, known as the equally likely risky outcomes method (ELRO), is to fix p and x_1 and then, for each x_2, elicit $x_3= (x_1,x_2;p,1-p)$. It is possible to apply equation (3.6.1) recursively to generate values of the utility function for a set of outcomes. An appropriate functional form may then be fitted to the resulting curve. This procedure has typically been more satisfactory than the standard reference contract approach, in that it yields internally consistent estimates of the utility function. However, changes in the probability p lead to changes in the estimated utility function (Karmarkar 1974). Once again, this is consistent with some sort of probability weighting.

The most successful method of eliciting utility functions has been the "equally likely certainty equivalent" (ELCE) approach of Anderson, Dillon and Hardaker (1977) based on fixed 50-50 odds. In this approach the utility function is estimated on the basis of choices between prospects all of which are of the form $\{x_1,x_2 ; 1/2, 1/2\}$. This method has encountered fewer problems than any of the others described above. In summary, the evidence from attempts to construct utility functions suggests the existence of some form of probability weighting which is sufficiently important to make the derivation of the expectation of utility unreliable as a guide to preferences whenever prospects involving different probabilities are compared. The evidence also suggests that probability weighting is most significant for small probabilities and least significant when both outcomes are equi-probable. However, an alternative possible explanation for the success of the ELCE approach is that probability weighting for the same probability pair (p, 1-p) may be different when the preferred outcome arises with probability p than when it arises with probability 1-p.

3.7 The response of EU theorists

All scientific research programs encounter anomalies and apparent refutations of their predictions. The nature of the response to such setbacks is an important test of a program's continuing viability. in particular, a progressive program is likely to respond with new hypotheses which generate unexpected new predictions, while a degenerating program is likely to respond with content decreasing hypotheses which effectively immunise the program against refutation. The main response of EU theorists to the problems described above has been to argue that the responses made by individuals in these circumstances are, in some sense, "mistakes" (see, for example Savage 1954). This can be interpreted in two ways.

First, the EU theory can be treated as a normative theory of rational choice, rather than a positive theory of behavior.(Raiffa 1961) This is essentially an admission of defeat in a scientific sense. Moreover, the EU axioms are a questionable basis for such a normative theory. Although normative propositions cannot be stated with certainty, it would seem reasonable to suggest that most normative theories would want to proscribe some of the preference patterns admitted by EU theory, such as global risk-preference . As has been shown in section 2.2, global risk-preference engenders behavior patterns which would generally be regarded as self-destructive in the extreme. On the other hand,given the fact that examples such as the Allais paradox were generated precisely because some people did not accept the normative force of the EU independence axiom, it does not seem likely that a retreat from positive prediction to normative presciption will greatly increase the appeal of the theory.

Some direct evidence on the normative appeal of the EU axioms has been obtained. MacCrimmon and Larsson (1979) presented subjects with a range of 20 proposed axioms and decision rules, some consistent with EU and others inconsistent and asked the subjects to rate them on acceptability from 1 to 10. The EU axiom of transitivity received the strongest assent, with an average rating of 8.8, while dominance was third with 8.1 and completeness seventh with 6.9. By contrast, various forms of the independence axiom received fairly low levels of support, less than those given to proposed rules which directly contradicted the axiom. As an aside, it may be noted that the lowest average rating for any rule was 5.0, suggesting that any plausibly phrased rule will receive wide assent, even if it is not obeyed in practice.

A more interesting interpretation of the "mistakes" view involves an attempt to maintain the claims of EU theory to provide a predictive framework by denying the value of apparently contradictory evidence. A crucial point here is the argument that problems in the experimental setting lead subjects to make choices they would not make in serious real-world settings. A number of problems of this kind may be considered.

First, there is the simple observation that, since the experimental settings usually do not involve real money, subjects have no incentive to make careful choices. Second, there is the fact that some experiments may be regarded in the light of games in which what matters most is whether one wins (finishes ahead) or loses (finishes behind). This is a possible explanation for some of the results showing a sharp change in behavior between the domain of losses and the domain of gains. Third, there is the observed phenomenon of context effects; it is possible to elicit different choices concerning the same prospects if they are presented in different ways (Kahneman and Tversky 1979, Schoemaker and Kunreuther 1979). Presumably, people would be more careful to look out for such context effects in real-life choices than in experiments.

These argument have been bolstered up by claims that, when presented with the appropriate arguments from EU theory, subjects would revise their choices in line with the axioms. This appears to be true in some cases, such as violations of transitivity. Tversky (1969) devised experiments in which a number of subjects made intransitive choices. When the intransitivity was pointed out nearly all subjects revised their choices. Indeed, some refused to believe they could have made such choices.

However, in the case of the more controversial independence axiom, subjects who initially conform to the axiom are just as likely to revise their choices (when faced with anti-EU arguments) as are those who initially break the axioms (MacCrimmon and Larsson 1979). A similar experimental outcome, along with an entertaining hypothetical dialogue between 'Professor A' and 'Professor S', is given by Slovic and Tversky (1974).

It does not seem possible to arrive at a final resolution of these issues yet. A crucial point is that continued adherence to the EU theory will depend heavily on the extent to which alternative theories can account account for the anomalies described above while maintaining the powerful predictive scope of the EU approach. In this thesis, it is argued that alternative theories have now been developed which permit the accomplishment of these tasks, and hence that the EU approach has been superseded. However, in a broader perspective, it must be noted that the alternatives advocated here are in the same spirit as that which motivated the development of EU theory and are generalisations of this theory. Their development constitutes a forward step for the broad research program of analysing choice under uncertainty in terms of cardinal utility.