QUADERNI

Università degli Studi di Siena
DIPARTIMENTO DI ECONOMIA POLITICA

MARCELLO BASILI
CARLO ZAPPIA

Ambiguity and uncertainty in Ellsberg and Shackle

n. 460 - Settembre 2005

Abstract - This paper argues that Ellsberg's and Shackle's frameworks for discussing the limits of the (subjective) probabilistic approach to decision theory are not as different as they may appear. To stress the common elements in their theories Keynes's *Treatise on Probability* provides an essential starting point. Keynes's rejection of well-defined probability functions, and of maximisation as a guide to human conduct, is shown to imply a reconsideration of what probability theory can encompass that is in the same vein of Ellsberg's and Shackle's concern in the years of the consolidation of Savage's new probabilistic mainstream. The parallel between Keynes and the two decision theorists is drawn by means of a particular assessment of Shackle's theory of decision, namely, it is interpreted in the light of Ellsberg's doctoral dissertation. In this thesis, published only as late as 2001, Ellsberg developed the details and devised the philosophical background of his criticism of Savage as first put forward in the famed 1961 QJE article. The paper discusses the grounds on which the ambiguity surrounding the decision maker in Ellsberg's urn experiment can be deemed analogous to the uncertainty faced by Shackle's entrepreneur taking "unique decisions." The paper argues also that the insights at the basis of the work of both Shackle and Ellsberg, as well as the criteria for decision under uncertainty they put forward, are relevant to understand the development of modern decision theory.

JEL Classification: B21, D21
Keywords: uncertainty, weight of argument, non-additive probability

Acknowledgements
A previous version of this paper was presented at the Annual Meeting of the Associazione Italiana per la Storia dell'Economia Politica, Siena, June 2005, and at the Annual Meeting of the History of Economics Society, Tacoma, WA, June 2005. Comments by Lawrence Boland, Daniel Ellsberg, Alberto Feduzi and Roberto Romani are gratefully acknowledged. Corresponding author:
zappia@unisi.it

Marcello Basili, Department of Economics, University of Siena
Carlo Zappia, Department of Economics, University of Siena

1. Introduction

George L. S. Shackle and Daniel Ellsberg represent two main positions among the critics of the forthcoming (at their time) mainstream in modern decision theory as represented by Savage's *Foundations of Statistics* (1954). Although both opposed maximisation of (subjective) expected utility as a criterion for choice under uncertainty, at first their theoretical enterprises appear to have few points in common.

Ellsberg (1961) introduced the notion of "ambiguity" to analyse actual decision-makers acting in a context that, though confounding and denoted by information perceived as scanty, can be represented through an exhaustive list of the possible states of the world. The paradoxical results of his thought experiment about choices from urns containing coloured balls in unknown proportions were confirmed by early experimental evidence (Becker and Brownson 1964). But, after a brief round of discussion (Raiffa 1961 and Brewer 1963), Ellsberg's results were put aside, "simply because researchers at the time were helpless to address them" (Machina 2001, xxxix). It is likely that the fact that Ellsberg never became an academic – because of his involvement first with Rand Corporation as strategic analyst and then with the U. S. Government as consultant during the Vietnam War – contributed to the decline of interest in his paradox. This neglect has been unfortunate, however, because Ellsberg's dissertation submitted to the Economics Department of Harvard University in 1962 is a major achievement. It provides such a thorough discussion of decision criteria for solving the paradox that, had it been published in the 1960s, it would have surely contributed to the progress of research attempting to represent the behaviour brought to light by the paradox.[1] Only since the late 1980s the paradox has become a focus of interest to decision theorists, as demonstrated by an exponential growth in the literature (surveyed for the first time in Kelsey and Quiggin 1992).

As regards Shackle (1949 and 1955), he insisted that the notion of uncertainty could not be reduced either to aleatory probability or subjective probability. In doing so he gave birth to a strand of thought emphasizing the role of "fundamental uncertainty" in decision-making, a perspective especially taken up by Keynesians and Austrians (Lachmann 1976 and Davidson 1983). Shackle rejected the use of probability measures in decision theory on the basis that the context of crucial entrepreneurial decisions is characterized by the fact that the list of possible states of the world known to the entrepreneur is not exhaustive. Shackle's

[1] Ellsberg's thesis was published only as late as 2001 (Ellsberg 2001).

1. Introduction

George L. S. Shackle and Daniel Ellsberg represent two main positions among the critics of the forthcoming (at their time) mainstream in modern decision theory as represented by Savage's *Foundations of Statistics* (1954). Although both opposed maximisation of (subjective) expected utility as a criterion for choice under uncertainty, at first their theoretical enterprises appear to have few points in common.

Ellsberg (1961) introduced the notion of "ambiguity" to refer to situations in which, due to lack of information, there is uncertainty about probabilities on events. Ellsberg aimed to work in the footsteps of Knight (1921) and his notion of "unmeasurable uncertainty," but started from the analysis of actual decisions to be taken in a context that, though confounding and denoted by information perceived as scanty, can be represented through an exhaustive list of the possible states of the world. The paradoxical results of his thought experiment about choices from urns containing coloured balls in unknown proportions were confirmed by early experimental evidence (Becker and Brownson 1964). But, after a brief round of discussion (Raiffa 1961, Brewer 1963, Ellsberg 1963, and Brewer and Fellner 1965), Ellsberg's results were put aside, "simply because researchers at the time were helpless to address them" (Machina 2001, xxxix).[1] The experimental evidence was to became huge in the following years (Kahneman and Tversky 1979, and Camerer and Weber 1992), but only since the late 1980s the paradox has become a focus of interest to decision theorists, as demonstrated by an exponential growth in the literature (surveyed for the first time in Kelsey and Quiggin 1992).

As regards Shackle (1949 and 1955), he insisted that the notion of uncertainty could not be reduced either to aleatory probability or subjective probability. In doing so he gave birth to a strand of thought emphasizing the role of "fundamental uncertainty" in decision-making, a perspective especially taken up by Keynesians and Austrians (Lachmann 1976 and Davidson 1983). Shackle rejected the use of probability measures in decision theory on the basis that the context of crucial entrepreneurial decisions is characterized by the fact that the list of possible states of the world known to the entrepreneur is not exhaustive. Shackle's

[1] It is likely that the fact that Ellsberg never became an academic – because of his involvement first with Rand Corporation as strategic analyst and then with the U. S. Government as consultant during the Vietnam War – contributed to the decline of interest in his paradox. This neglect has been unfortunate, however, because Ellsberg's dissertation submitted to the Economics Department of Harvard University in 1962 is a major achievement. It provides such a thorough discussion of decision criteria for solving the paradox that, had it been published in the 1960s, it would have surely contributed to the progress of research attempting to represent the behaviour brought to light by the paradox. Ellsberg's thesis was published only as late as 2001 (Ellsberg 2001).

extreme position is testified by his remark that this list cannot be complete even in principle, because the entrepreneur's decisions effectively shape the future environment. The interest raised by Shackle's theory among decision theorists in the 1950s (in particular, see Arrow 1951, and the essays collected in Bowman 1958 and in *Metroeconomica* 1959 special issue) faded away in the early 1960s. Shackle's 1961 volume *Decision, Order and Time in Human Affairs*, where he restated the notion of potential surprise as an alternative to probability measures and discussed for the first time subjective probability theory, signalled the inability of its author to comprehend the critics' arguments and produced the effect of distancing decision theorists from his approach.

This paper argues that Ellsberg's and Shackle's frameworks for discussing the limits of the (subjective) probabilistic approach to decision theory are not as different as they may appear. To stress the common elements in their theories Keynes's *Treatise on Probability* provides an essential starting point. Keynes's rejection of well-defined probability functions and of maximisation as a guide to human conduct is shown to imply a reconsideration of what probability theory can encompass. Both Ellsberg and Shackle had similar concerns in the years of the consolidation of Savage's new probabilistic mainstream. The parallel between Keynes and the two decision theorists will be drawn by means of a particular assessment of Shackle's theory of decision, namely, it will be interpreted in the light of Ellsberg's doctoral dissertation. In this thesis, Ellsberg developed the details and devised the philosophical background of his criticism of Savage as first put forward in the famed 1961 QJE article. Our general point is that the insights at the basis of the work of both Shackle and Ellsberg, as well as the criteria for decision under uncertainty they put forward, are relevant to understand the development of modern decision theory.[2]

Section 2 discusses Keynes's criticism of the frequentist approach to probability theory. The goal is to highlight the unconventional points made by Keynes well before the consolidation of the mainstream. In this section we stress that Keynes's view of the potential incomparability of alternative probability assessments hints at the crucial notion of weighted probabilities. Section 3 reviews the main elements lying at the basis of Shackle's rejection of the probabilistic approach and presents the criterion for choice under uncertainty he proposed. Section 4 deals with Ellsberg's discussion of Shackle's criterion and analyses Ellsberg's proposal for a theory alternative to subjective expected utility. In section 5 the approaches of

[2] Our paper can be intended as a step towards an historical reconstruction of the development of modern decision theory in the 1950s, that is, at the time of the making of Savage's subjective probability "revolution." See also Zappia and Basili (2005).

Ellsberg and Shackle are compared. The grounds on which the ambiguity surrounding the decision maker in Ellsberg's urn experiment can be deemed analogous to the uncertainty faced by Shackle's entrepreneur taking unique decisions are pointed out. Section 6 summarises the argument.

2. The Keynesian origins[3]

To put it schematically, there are two main traditions in the history of probability theory. First, and this is the conventional view, probability is intended as aleatory probability. Second, since Jacob Bernoulli and especially after Keynes's criticism of the classical approach, there developed a current assessing probability as referring to epistemic probability as well.[4] Both Shackle and Ellsberg discussed probability in the tradition of epistemic probability. To seek similarities between Shackle's and Ellsberg's approaches two general issues stand out as most important: first, the analogy between their views on probability and Keynes's view, and, second, the application of probability to conduct in accordance with Keynes's perspective as put forward in the *Treatise on Probability*.

Bernoulli's notion of probability as presented in his *Ars Conjectandi* was essentially epistemic. He defined probability as a part of certainty and considered it as a measure of human knowledge, while games of chance were used as mere examples.[5] Probabilities are calculated from arguments: "probabilities are appraised from the number together with the

---

[3] The scattered elements of Keynes's philosophy of probability discussed in this section emphasize, in the main, those aspects which are at the basis of the so-called new Keynesian fundamentalism (Lawson 1985, Carabelli 1988 and O'Donnell 1989). However our understanding of Keynes' rejection of well-defined probability functions, and of maximisation as a guide to human conduct, points to a theoretical option that is different from the one favoured by most Keynesian interpreters. In fact, our interpretation does not imply the rejection of probability theory tout-court, but discusses a reformulation of what probability theory can encompass, with specific regard to non-additive probability measures. More recently Carabelli (2002) has conceded that there is a strong link between certain new developments in modern decision theory and Keynes's probability theory, with specific regard to the application to financial markets.

[4] See in particular Hacking 1975, Ch. 6. We use Glen Shaffer's definition (1978, p. 312-313): "Aleatory probabilities find their role in the paradigm of chance; they are number assigned to various possible outcomes of a chance event. The aleatory probability (or chance) of each outcome is thought to measure its propensity to occur … an epistemic probability on the other hand describes our knowledge. It is a number that represents, albeit with a usually ludicrous affectation of precision, the degree of which we are certain of something, or alternatively, the degree to which we believe it or the degree to which our evidence supports it."

[5] "Probability is a degree of certainty and differs from it as a part from a whole" (Bernoulli 1713, p. 211)

weight of the arguments which in any way prove or indicate that thing is, will be, or has been. By weight, moreover, I mean the force of proof" (Bernoulli 1713, p. 214). To Bernoulli arguments are contingent, and contingency is related to human knowledge.[6]

Like Bernoulli, in the *Treatise on Probability* Keynes interpreted probability as different from chance or frequency: "the identification of probability with statistical frequency is a very grave departure from the established use of the words; for it clearly excludes a great number of judgments which are generally believed to deal with probability" (1921, p. 109). As is well-known, in Keynes's view (1921, p. 4) probability is a logical relation between two sets of propositions:[7] "the terms certain and probable describe the various degrees of rational belief about a proposition which different amount of knowledge authorise us to entertain …The theory of probability is logical, therefore, because is concerned with the degree of belief which is rational to entertain in given conditions" Moreover, "a definition of probability is not possible unless it consents us to define degrees of the probability-relation by reference to degrees of rational belief."[8]

Keynes (1921, p. 12) distinguished between direct and indirect knowledge. The former is "that part of our rational belief which we know directly;" the latter is "the part which we know by argument." Keynes explained: "our knowledge of propositions seems to be obtained in two ways: directly by contemplating the objects of acquaintance and indirectly, *by argument*, through perceiving the probability-relation of the proposition, about which we seek knowledge, to other propositions." It ensues that Keynes is following Bernoulli's path when he argues that indirect knowledge is based on argument and involves a degree of probability

---

[6] Bernoulli distinguished between pure and mixed arguments: "I call those arguments pure which prove a thing in certain cases in such a way that they prove nothing positively in other cases. I call those arguments mixed which prove the thing in some cases in such a way that they prove the contrary in the remaining cases" (Bernoulli 1713, p. 218). As commented by Shafer (1978, p. 329), this distinction is remarkable since the possibility of non-additivity lies in the definition of pure arguments: as a matter of fact "the evidence of a mixed argument points to both the questions it addresses, whereas the evidence of a pure argument points only to the positive side." Hacking (1975, p. 144) notes: "until 1713 it was not in the least determined that that the addition law for probability would be accepted. Bernoulli was the last master to contemplate non-additive probabilities."

[7] Keynes (1921, p. 317) stressed that "a careful examination of all cases in which various writers claim to detect the presence of 'objective chance' confirms the view that 'subjective chance', which is concerned with knowledge and ignorance, is fundamental, and the so-called 'objective chance', however important it may turn out to be from the practical or scientific point of view, is really a special kind of 'subjective chance' and a derivative type of the latter."

[8] As a result, Keynes defined the subject matter of the theory of probability as follows: "Between two sets of propositions, therefore, there exists a relation, in virtue of which, if we know the first, we can attach to the latter some degree of rational belief. This relation is the subject-matter of the logic of probability" (Keynes 1921, p. 7). And the notion of probability is closely related to the knowledge of the secondary proposition "which assert the existence of probability-relations in the fundamental logical sense" (1921, p. 12).

lower than certainty: "the knowledge of a secondary proposition leads only to a rational belief of the appropriate degree in the primary proposition. The knowledge in this latter case I have called knowledge about the primary proposition or conclusion of the argument, as distinct from knowledge of it" (1921, p. 15). The *Treatise* is an attempt to derive knowledge from probability arguments. Measurement of probabilities should consider both the magnitude of the probability of an argument and the weight of the argument. In Keynes (1921, p. 21), measurement of probability means comparison of the arguments, for such a comparison is "theoretically possible, whether or not we are actually competent in every case to make."[9]

Keynes (1921, p. 34) argued then that unknown probabilities can occur only due to lack of skill in arguing from given evidence, and not to lack of evidence. This is because "the weakness of our reasoning power prevents our knowing what this degree [of probability] is. At the best, in such cases, we only know *vaguely* with what degree of probability the premises invest the conclusion." It follows that a probability is a magnitude between impossibility and certainty, that is a number in the interval from 0 to 1; so that when one argues that one probability is "greater" than another, "this precisely means that the degree of our rational belief in the first case lies *between* certainty and the degree of the rational belief in the second case" (Keynes 1921, p. 37).

At the same time, Keynes was well aware that the probabilities of two quite different arguments can be impossible to compare. Probabilities can be compared if they belong to the same series, that is, if they "belong to a single set of magnitude measurable in term of a common unit" (35). In contrast, probabilities are impossible to compare if they belong to two different arguments and one of them is not (weakly) included in the other. As Keynes (1921, p. 38) put it: "some probabilities are not comparable in respect to more or less, because there exists more than one path, so to speak, between proof and disproof, between certainty and impossibility; and neither of two probabilities, which lies on independent paths, bears to the other and to certainty the relation of 'between' which is necessary for quantitative comparison." What Keynes is telling us about comparability is evident from the figure in chapter 3 of the *Treatise* (reproduced below in footnote 10). This is a diagram that on the horizontal axis has the scale of probability ranging from 0 to 1: every point on this axis can be compared to the other points, because there exists a numerical representation of our degree of

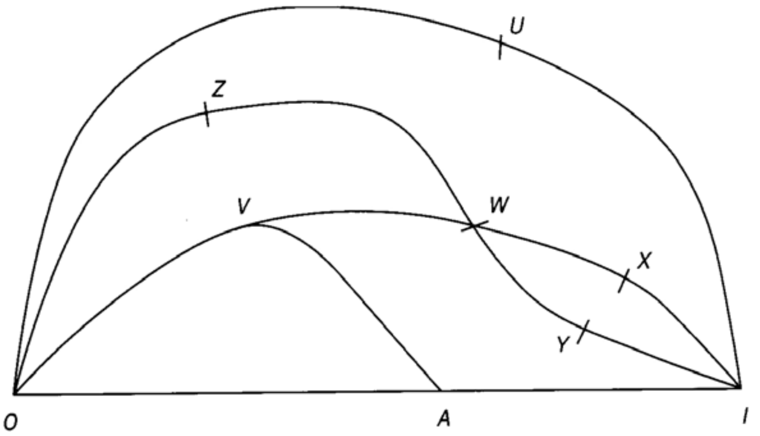
[9] Keynes (1921, p. 33) distinguished between four alternatives: "in some cases there is no probability at all; or probabilities do not all belong to a single set of magnitudes measurable in terms of a common unit; or these measures always exist, but in many cases are, and *must remain*, unknown; or probabilities do belong to such a set and their measures are capable of being determined by us, although we are not always able so to determine them in practice."

belief about a logical proposition. But in the plane created by the diagram there are also other different paths, starting with 0 and ending with 1, which do not lie on the straight line between the extremes. Each point on every non-linear path identifies what Keynes calls a "non-numerical probability," but Keynes leaves undetermined the interpretation of the vertical dimension.[10] Keynes did not provide a mathematical structure for his probability values, but suggested that probabilities lying on the same path can be compared among themselves. However he contended that comparisons are possible only among points on the same path (or among paths that have points in common).

In the perspective of the modern theory of decision, the Keynesian paths are nothing but distorted probabilities, that is, contractions or expansions of prior linear probabilities. Up until the approaches of Rank Dependent Expected Utility (Quiggin 1982, and Yaari 1987) and Choquet Expect Utility (Gilboa and Schmeidler 1989), a consistent theory of probability distortions was not provided. But it is significant that the issue was paramount in decision theory at the time of the making of Savage's subjective interpretation of probabilities. Since the inception of experimental economics it was evident (see in particular Edwards 1954) that plenty of experimental evidence challenged the hypothesis that decision-makers act as if they were endowed with an additive probability distribution. This experimental evidence constitutes the main starting point of Kahneman and Tversky's work (1979), which discusses the usefulness of a weighting function within the framework of their descriptive model of decision making under risk.[11] As will be shown later, with respect to this issue both Shackle and Ellsberg follow in Keynes's footsteps. In fact, both Shackle and Ellsberg discussed, though under different headings, instances of distorted or weighted probabilities: this reveals

---

[10] Keynes's diagramme (1921, p. 42).



[11] Tversky and Wakker (1995) show that prospect theory can be axiomatised by means of a non-additive measure and hence turned into a model of decision making under uncertainty, labelled cumulative prospect theory after Tversky and Kahneman (1992).

that their rejection of Savage's approach has a common origin in Keynes's probability theory.[12]

Apart from the magnitude of its degree of probability, the second relevant aspect in the measurement of an argument is, in Keynes's view, its weight. Keynes (1921, p. 77) maintained that the weight of an argument is correlated, but independent of its degree of probability. On the one hand, "the magnitude of probability of an argument depends upon a balance between what may be termed the favourable and the unfavourable evidence; a new piece of evidence which leaves the balance unchanged, also leaves the probability of the argument unchanged." On the other hand, there exists another characteristic of the arguments by which they can be compared, and this is termed the weight, that is, "a balance, not between the favourable and the unfavourable evidence, but between the absolute amounts of relevant knowledge and of relevant ignorance respectively." Keynes (1921, p. 77) clarified the relationship between magnitude and weight of an argument as follows: "the magnitude of the probability of the argument may either decrease or increase, according as the new knowledge strengthens the unfavourable or favourable evidence; but something seems to have increased in either case … an accession of new evidence increases the weight of an argument. New evidence will sometimes decrease the probability of an argument, but it will always increase its weight." Moreover in the case of weight the comparison between arguments with respect to less and more is possible only if one is a subset of the other.

There is a huge literature about the notion of weight of argument in Keynes, basically because Keynes did not define weight in a consistent way. For long, the leading interpretation was to relate the weight of argument to the notion of second order probability distribution (Gardenfors and Sahlin, 1982). The point of this interpretation is that, granted that weight of argument is only a probability distribution over the probability distribution on the set of

---

12 Interestingly enough, in talking about vagueness and numerical measurement of probabilities Keynes (1921, 176) maintained that "it is evident that the cases in which exact numerical measurement is possible are a very limited class….many probabilities, which are incapable of numerical measurement, can be placed nevertheless *between* numerical limits." In this way Keynes introduces the notion of interval of probability as a representation of ambiguity. Well before both von Neumann and Morgenstern and Savage, Koopman (1940) had introduced an axiomatic treatment of intervals of probability into probability theory. The notion of upper and lower probabilities was then taken up by Good (1950) – who argued that the necessity for intervals arises from the fact that individuals' initial judgements are imperfect – and remained part of a lively, though minor, tradition of thought opposing the restrictive interpretation of subjectivism proposed by Savage (in particular see Smith 1961, Shafer 1976 and Walley 1991). As will be discussed in section 4 below, Ellsberg (2001) referred to Keynes, Koopman and Good to argue that a subjectivist tradition emphasising vagueness and indeterminacy in probability judgements existed before Savage's systematisation.

events, it is always possible to compare two arguments simply by applying the rule of reduction of compound lotteries (Anand, 1991).

But this interpretation betrays Keynes's idea of probability in a crucial sense. The weight of argument, though expressing a degree of confidence in a probability assessment, is quite a different thing from a probability of a second order. Keynes (1921, p. 345) argued that the weight of an argument is "the degree of completeness of the information upon which a probability is based," and it ranges between 0 and 1. He also proposed a precise way to calculate it, however, when he stated that a weight is the "balance between the absolute amounts of the relevant knowledge and of the relevant ignorance respectively" (Keynes 1921, p. 77).[13]

In the language of modern decision theory, given an event A, the relevant ignorance can be defined as $\psi = 1-(v(A)+v(A^C))$, that is, as the difference between complete knowledge, normalised to the unity, and the probability of the occurrence of the event plus the probability of its complement (negation of the event). The weight of argument can be represented by $\omega = 1-\psi$, that is, by the difference between the absolute amount of relevant knowledge and the absolute amount of relevant ignorance. It is worth noting that if v is an additive probability measure, the relevant ignorance is zero and the weight of argument is 1. But if v is a non-additive probability measure (convex capacity), the relevant ignorance is different from zero and the weight of argument belongs to the interval zero-one.[14]

Following this interpretation of Keynes's thought, the significance of the weight of argument emerges only when the decision-maker is not endowed with a unique additive probability measure and does not behave as an expected utility maximizer. It is worth observing that Keynes's inability to compare magnitude and weight of different arguments depends on the lack of a coherent theory through which the decision-maker's beliefs can be

[13] Runde (1990) was the first commentator to call attention to the fact that there are two different definitions of weight in the *Treatise*. He emphasised the importance of the definition of evidential weight as the degree of completeness of information on which a probability assessment is based, instead of the mere absolute amount of evidence implicit in the second order probability interpretation.

[14] Dow and Verlang (1992) introduced the notion of relative ignorance we refer to in order to characterise uncertainty aversion with non-additive measures, drawing on Schmeidler's (1989) hint that the convexity of the capacity v indicates the decision maker's confidence in the probability assessment. Kelsey (1994) discussed a representation of the weight of evidence (that do not collapse into second order probability) in terms of a range of possible probabilities and relates it to Gilboa and Schmeidler's maximin expected utility. Vercelli (1999a and 1999b) put emphasis on the link between the Keynesian weight and non-additive probability theory.

transformed into weighted probabilities.[15] As is implicitly assumed by Keynes, the necessity to introduce weighted probability is related to the degree of ambiguity the decision-maker has to take into account in a given situation. It will be argued in the following sections that both Shackle and Ellsberg focused on the missing link between priors and weighted probabilities. They shared this line of research with scholars like Ward Edwards and William Fellner, who discussed the soundness of Savage's subjective approach in the late 1950s and early 1960s. But, as noted above, it was only recently that an axiomatic theory incorporating a rule to substitute decision weights for probabilities was provided: this task was accomplished through the rank-ordering of consequences assumed in both Rank Dependent Expected Utility and Choquet Expected Utility.[16] As will be shown in the next sections, even with respect to this second aspect of the weighing evaluation, Shackle's and Ellsberg's analyses are squarely in the Keynesian tradition, and not surprisingly point to a solution similar to Keynes's. In fact, they weigh the agent's evaluation of an ambiguous act by a parameter that represents her relative ignorance.

There is another aspect of Keynes's theory is instrumental to our argument. Keynes devoted chapter 26 of his *Treatise* to the application of probability to conduct. His problem was the interpretation of "goodness" of choice when "it is not rational for us to believe that the probable is true." Referring to the selection of an appropriate rule of choice, Keynes (1921, p. 343) recalled that "normal ethical theory at the present day makes two assumptions: first, that degrees of goodness are numerically measurable and arithmetically additive, and second, that degrees of probability also are numerically measurable." As a result, ethical theory decides among alternative acts on the basis of their mathematical expectations, which he presented as "a technical expression originally derived from the scientific study of gambling and games of chance, and stands for the product of the possible gain with the probability of attaining it." Keynes disagrees with a generalized application of mathematical expectation for two main reasons: first, because the assumption that "quantities of goodness are duly subject to the law of arithmetic, appears to me to be open to a certain amount of doubt," and, second, because to assume that "degrees of probability are wholly subject to the

15 Formally weighed functions are obtained from beliefs by a non-decreasing function $\lambda : [0,1] \rightarrow [0,1]$ with $\lambda(0) = 0, \lambda(1) = 1$.

16 All these recent developments presuppose that the decision weight of events depends on the rank-ordering of their consequences. That is, consequences are ranked from the best to the worst or vice-versa and decision-weights, that add up to one, are derived as the marginal contribution to the cumulative distribution function. The distinguishing characteristic of these models is that the transformed probability of an outcome depends on the rank of that outcome in the induced preference ordering on the set of outcomes (Karni and Schmeidler, 1991).

law of arithmetic, runs directly counter to the view which has been advocated in part I… [because it] ignores what I have termed the weights of arguments, namely the amount of evidence upon which each probability is founded" (Keynes 1921, p. 344). By explicitly referring to the probability paths introduced in the diagram of chapter 3 mentioned above, Keynes held that "mathematical expectations, of goods or advantage, are not always numerically measurable, and hence … even if a meaning can be given to the sum of a series of non-numerical mathematical expectations, not every pair of such sums are numerically comparable in respect of more and less." In other words, "it is not always possible by a mere process of arithmetic to determine which of the alternative ought be chosen" (Keynes 1921, p. 344-345).

In the light of our interpretation of the relationship between weight of argument and additivity, it follows that since mathematical expectation neglects the weight of argument this is an adequate criterion for choice only if the probabilities are additive, that is, only if the weight of argument is one. In all other cases it cannot be applied. As will be shown in the following sections, Shackle and Ellsberg rejected the rule of mathematical expectation because the decision-makers they assume face ambiguous situations and take the weight of (absolute) relevant ignorance into their evaluations, that is, they evaluate a weighed average of all possible consequences. Moreover, both Shackle and Ellsberg apply an evaluation method that explicitly takes into account the agent's attitude towards ambiguity. Thus, the rationale underlying this last aspect of Keynes's criticism of aleatory probability indicates that Shackle and Ellsberg's rejection of mathematical expectation has a Keynesian origin.

3. Shackle's non-probabilistic approach

Most economic decisions, Shackle argued (1949, p. 6), are crucial, unique "experiments," namely situations where "the person concerned cannot exclude from his mind the possibility that the very act of performing the experiment may destroy forever the circumstances in which it was performed." The fact that these decisions are non-replicable precludes the possibility of applying probabilities, both numerical probabilities according to the frequency probability theory and, later, subjective probabilities.

Shackle's argument had a clear Knightian flavour. The use of probability calculus to analyse decisions under uncertainty was inappropriate, in his view, simply because the conditions required for its application do not exist in many relevant economic contexts. In actual situations individuals do not have a complete knowledge of the structure of the world. Individual choices are made between alternatives which are subjective representations of alternative future sequels to actions, and not between future sequels themselves. In Shackle's words, "choice is among imagined experiences," a view which implies that the individual is not given an exhaustive list of the alternatives between which choice should be made.

In particular, Shackle argued that individuals are not capable of enumerating all possible contingencies, or states of the world. This is the main analytical point at the basis of Shackle's theory.[17] On this basis, Shackle developed a formal theory intended to capture both the mental processes and the non-repetitive, and often irreversible, nature of actual economic decisions. Shackle's argument was initially intended to oppose the objective frequency-ratio interpretation of probability, which he regarded as the mainstream view in the late 1940s. In 1961, however, he maintained that the same argument could apply to the subjective interpretation of probability (Shackle 1961).

To dispense with probability, Shackle put forward the concept of potential surprise. First, he distinguished between distributional uncertainty variables, which can be used if "the list [of suggested answers to a question] is *complete without a residual hypothesis*," and non-distributional uncertainty variables, which must be used when "the list in order to attain formal completeness must be rounded off with a residual hypothesis" (Shackle 1961, pp. 49-50). In order to describe the "mental state of uncertainty" of the decision-maker, Shackle maintained, "the inclusion of a residual hypothesis in his list of suggested answers is his acknowledgement that he has no basis for considering his existing list of particularised hypotheses to be comprehensive." As a result, what was needed was "a measure of acceptance, of a hypothesis proposed in answer to some question, that shall be independent of the degrees of acceptance simultaneously accorded to rival hypothesis," a measure such that "the individual can give to new rival hypotheses, which did not at first occur to him, some degree, and even the highest degree, of acceptance without reducing the degrees of acceptance accorded to any of those already present in his mind" (Shackle 1949-50, p. 70).

Shackle's potential surprise function then amounts to a substitute for probability distributions, to be used under uncertainty. The distinction between distributional and non-

[17] To economists working in heterodox traditions of thought, like the post-Keynesians and the Austrians, this point has become an indispensable analytical reference in their effort to represent decisions under genuine uncertainty (for instance see Lachmann 1976, Davidson 1983 and Dow 1995). In particular, this is the crucial argument upon which the distinction between "rational ignorance" and "radical ignorance" has been drawn in the Austrian tradition (Langlois 1994 and Vaughn 1994). By this distinction, the Austrians intend to distinguish Savage's approach from the Knightian tradition. On this point see Zappia (1998).

distributional variables shows that Shackle's theory is essentially non-additive, a point he was well aware from the beginning (see Shackle 1949-50). This point was discussed by some of his critics in the 1950s, who made explicit reference to the possibility of avoiding to dispense with probability, but left largely unaddressed by Shackle.[18]

Shackle's next step was to apply his (non-additive) measure to decision making. He analyzed a decision-maker, typically an entrepreneur, who had to choose among alternative "sequels" to actions on the basis of two elements: the possible gains and losses embedded in a sequel, called face-values, and a valuation of the "possibility" of the gains and losses, called potential surprise. The latter element can be considered as a degree of disbelief, or implausibility of the hypothesis that supports the sequel; it ranges from 0 (absence of disbelief or zero potential surprise) to a maximum value expressing impossibility (absolute disbelief or maximum potential surprise).[19] When the decision maker chooses among alternative sequels, she re-considers the face-values of each sequel by their degree of potential surprise.

Finally, Shackle defined a function φ, called "degree of stimulus" (or "ascendancy function") whose arguments are the face-values and the associated potential surprises implicit in a sequel. Given a degree of stimulus, it is possible to determine a prospect of the possible outcomes of a sequel weighed by degrees of reliability. On this ranking of outcomes, Shackle superimposed "the particular potential surprise curve which [the decision maker] assigns to some particular project." It is worth noting that the value of the degree of stimulus is bounded from below to 0, when the potential surprise is at its maximum. But it is also bounded from above, when the potential surprise is 0, and there is no loss of generality to assume that upper bound is normalised to 1.

---

[18] For instance, in his discussion of Shackle's insistence on the ineffectiveness of expected utility, Edwards (1958, 49) argued as follows: "Shackle's refusal to add subjective improbabilities … seems to me to be the most important and desirable feature of his system. A great deal of experimental evidence that bears on the additivity of subjective probability is now available and it argues against the additive property so strongly that I do not see how it is possible any longer to defend that property. Fortunately, it may be possible to develop a utility-subjective probability model that is mathematically satisfactory and that does not require subjective probabilities to add to one or anything else." Edwards's optimism as to the alternatives to subjective expected utility had to wait for many years. A similar point was made by Weckstein (1959, p. 110) and dismissed by Shackle (1961, p. 107-108).

[19] The aim of measuring the degree of belief in a certain event by means of its opposite, the degree of disbelief, or potential surprise, is instrumental to the construction of a non-additive index. The emergence of a new unanticipated event does not necessarily reduce the degree of disbelief previously assigned to other events, as it would be if (the opposite of) this degree is measured by a probability. "By disbelief I do not now mean the absence of perfect certainty, but the positive recognition of some disabling circumstances … and there is, in general, no limit to the number of mutually exclusive hypotheses to all of which simultaneously a person can, without logical contradiction, attach zero potential surprise" (Shackle 1952, pp. 30-31).

Following this procedure Shackle (1953, p. 46) determined "the highest bids which the particular project in question can make for the decision maker's attention and interest. One of these bids is the most powerful suggestion of success and the other the most powerful suggestion of disaster that the conception of project conveys." These extreme values represent the limits of all possible outcomes of any feasible sequel, after the outcomes and their potential surprise are valued in terms of the attitude the decision-maker shows towards the uncertain situation. Finally, the criterion for choice Shackle proposed amounts to a rule of thumb by which the decision maker takes into account both "the best possible" and the "worst possible" outcomes of each sequel, respectively called "focus-gain" and "focus-loss."

It must be noted that the ascendancy function has a role that is analogous to the Keynesian weight of argument, since different "degrees of stimulus" select different focus values for the same potential surprise function. Or, to put it differently, the potential surprise function is the (non-additive) substitute for the (additive) probability function, but its significance in the decision process is mediated by the ascendancy function.[20]

This way of formulating a criterion for decisions can be phrased in the language of modern decision theory. Shackle's decision-maker first orders prospect revenues of an act (sequel) on the basis of both their value and their degrees of possibility, and then she re-evaluates them by her specific (that is, relative to the project at hand) attitude towards the uncertain environment. At this point, she takes into account the best and the worst outcomes involved in the feasible act, instead of calculating an expected value of the act.[21] Eventually acts are ranked in terms of the best/worst pairs.

Shackle was well aware that he was proposing a heresy. "Those accustomed to think in terms of the actuarial calculation of the result of a divisible experiment," Shackle (1953, pp. 42-43) admitted, "will find this point exceedingly hard to appreciate."[22] Nonetheless Shackle was clear that he was suggesting a procedure alternative to von Neumann and

---

20 Among the followers of Shackle, Levi (1966 and 1972) is possibly the only one to insist on this point.

21 Shackle (1953, p. 47) summarises: "because the project is a non-divisible non-seriable experiment, his [the entrepreneur] various hypotheses as to its outcome are mutually exclusive and therefore there is here no logical basis for the additive procedure by which a 'mathematical expectation' is assigned to a divisible experiment."

22 Shackle's argument was weakened by the fact that he did not discuss the cognitive bases of the advocated procedure of selecting only two values for each sequel of action, apart from scattered consideration like the following one: "Suppose, for the sake of the argument, that the decision-maker does fix his attention momentarily on one of the interior hypotheses, the idea of a small gain, by hypothesis, is perfectly easy to imagine as the outcome of the project. But his attention will not rest on that hypothesis. His mind will be immediately challenged by the thought that much larger gains are equally easy to imagine: why then stop at the relatively small one?" (Shackle 1953, 42).

Morgenstern's maximin criterion, which Wald (1950) had transformed into a decision rule for decisions under complete uncertainty. Shackle (1953, p. 43) first observed that "the fundamental hypothesis of von Neumann and Morgenstern's *Theory of Games* is that a decision-maker will choose that action (strategy) whose worst possible outcome is the least bad amongst the respective worst possible outcomes of all the actions open to him." Then he added: when players do not have the "comprehensive, exact and certain knowledge" supposed by von Neumann and Morgenstern's setup, "would the von Neumann and Morgenstern hypothesis be more plausible than that the decision-maker will choose that action whose best possible outcome is the best amongst the respective best possible outcomes of all the actions open to him?" A maximax criterion, Shackle maintained, should be considered equally plausible at least. But, he then concluded:

"there is surely a third [criterion] which is more plausible, and of more general analytical power, than either of the two former, namely, that he [the decision-maker] will take into account both the 'best possible' and the 'worst possible' outcome of each course of action and make these *pairs* of outcomes the basis of his decision."

In the early 1950s the discussion about criteria for decision making was lively. In particular, the decision-theoretic view of statistics advanced by Wald had an obvious interpretation in terms of decision-making under complete ignorance. Wald (1950) discussed the problem of statistical inference as the problem of selecting a strategy, or decision function, from several available strategies when all that is known about the state of nature is that it is one of a given set of states of nature, that is, as a two-person zero-sum game in which the decision maker plays against Nature. The maximin strategy was shown by Wald to be a best response against Nature's minimax strategy, that is, against the least favourable a priori distribution Nature can employ. Wald's criterion is extremely conservative even in a context of complete ignorance, though ultra-conservatism may sometimes make good sense.[23] Several other criteria were proposed and discussed, and a list of properties of rationality and consistency were set forth as a set of axioms to be obeyed by a "rational criterion" (Milnor 1954 and Chernoff 1954). Among the criteria the most well-known was Hurwicz's. Hurwicz (1951) introduced a parameter intended to obviate the conservative (or pessimistic) procedure of concentrating only on states having the worst consequences. Hurwicz's decision rule selects the minimum and the maximum payoff to each given action *x*, and then associates to

[23] It is worth noting that the literature on environmental problems, where decisions are typically irreversible and potentially high losses related to catastrophic events are to be considered, makes explicit reference to Wald-related criteria (for instance, see Perrings 2003).

each action the following index: α max(x) + (1-α) min(x). Of any two actions, the one with the maximum index would be preferred.

In Hurwicz's view α expresses the degree of optimism the decision-maker is endorsed. If α=0 Hurwicz's criterion is Wald's maximin criterion, while if α=1 it is the maximax criterion.[24] After labelling Wald's maximin criterion as "ultraconservative," Luce and Raiffa (1957, p. 282) presented the rationale of Hurwicz's amendment in the following way: "Why not look at the best state, or at a weighted combination of the best and worst?" The resemblance with Shackle's apparently odd statement on von Neumann-Morgenstern's criterion cannot but strike the reader.

4. Ellsberg on ambiguity, uncertainty and Shackle

This section deals mainly with Ellsberg's doctoral thesis, which, although submitted to the Harvard Department of Economics in 1962, remained unpublished until very recently (Ellsberg, 2001). The thesis elaborates on Ellsberg's famous critique of Savage's subjective expected utility theory. In particular the thesis supplies both a philosophical background to the critique and a thorough discussion of criteria alternative to the maximization of subjective expected utility. The insights provided in the thesis are so deep that had Ellsberg decided to pursue academic research on these themes his influence on the development of modern decision theory would have been even wider than usually recognized.

Ellsberg's thesis started with a thorough analysis of the theories of those economists, mathematicians, and statisticians who had emphasized vagueness and imprecision of subjective assessments before his criticism of Savage. In the 1961 article Ellsberg had defined ambiguity as a situation differing from both risk and ignorance in Knight's sense. The article, which opens with the Knightian distinction between risk as measurable uncertainty and uncertainty as unmeasurable uncertainty, aims to counter scepticism about the behavioural significance of the distinction. In order to do so Ellsberg concentrated on a few decisional urn problems in which the information about the contents of the urn can be described with accuracy, but induces, in cases contemplating also urns with unknown proportion of balls of different colours, deliberate violations of Savage's sure thing principle. Ellsberg (1961, p.

[24] Because of its derivation from Wald, Hurwicz's decision rule can be termed α-maximin. Though discussed in the 1950s as one of the possible criteria for decision-making under complete ignorance, Hurwicz's paper was not published and circulated only as discussion paper. The criterion is best known as Arrow-Hurwicz criterion because it was published only in Arrow and Hurwicz 1972.

657) claimed that the nature of the individual's information concerning the likelihood of events was a relevant dimension of the decision problem, and proposed to called it the ambiguity of information, "a quality depending on the amount, type, reliability and 'unanimity' of information" expressing the individual's "degree of confidence in an estimate of relative likelihoods." After discussing the urn problems, however, Ellsberg noted "the ambiguity surrounding the outcome of a proposed *innovation*, a departure from current strategy, may be much more noticeable," thus hinting at uncertainty proper.[25]

In the much deeper analysis of his thesis, on the other hand, Ellsberg (2001, Ch. 1) quoted at length Keynes's *Treatise on Probability*, mostly because Keynes had inspired Koopman's and Good's use of interval probabilities. Ellsberg showed that, before the new subjectivist mainstream consolidated, there was a lively tradition of thought using a probabilistic view of decision-making. This tradition was more multifarious than usually recognized, insofar as it included economists like Knight, Keynes, and Shackle, philosopher mathematicians like Koopman, Hurwicz and Good, and statisticians like Wald, Hodges, and Lehmann. These authors shared, on different grounds, the same contention about the meaningfulness of assuming that subjective beliefs could be always represented by point probabilities. They formulated criteria for choice alternative to maximization of expected utility, as re-affirmed by von Neumann and Morgenstern and Savage in the tradition of Daniel Bernoulli and classical probability.

Ellsberg's aim (2001, p. 12) is explicitly posited:

"Actually, the theories of Koopman and the more recent analyses of I. J. Good seem to me to go far towards rendering the concepts of relative 'sureness' and 'vagueness' less vague in themselves. I hope in the present study to add to that development; and beyond that, to develop meaningful and useful hypotheses on appropriate decision criteria for basing action, reasonably and systematically, upon 'vague' and 'unsure' options."

[25] Ellsberg (1961, 653) noted that Knight, in discussing an urn situation analogous to the one he himself was proposing (in particular the two-colour urn example), did not oppose the intuition implicit in the so-called principle of insufficient reason. That is to say that people missing statistical information on the composition of the urn should act, according to Knight, on the supposition that the chances of balls of different colours were equal. In Ellsberg's view this assumption would contradict the significance of Knight's own distinction between measurable and unmeasurable uncertainty. Keynes (1921, 53-54) too had discussed the two-colour urn example, but, unlike Knight, with the intent to give a critical account of the limits of the principle of insufficient reason, in accordance with the idea that in situations where information is scanty or perceived to be vague the traditional approach to probability is not adequate. Ellsberg did not quote Keynes's *Treatise on Probability* in the 1961 article, because he did not read it until working on the definite draft of the thesis (Ellsberg, personal communication, June 2005). On this point see also Feduzi 2005.

Shackle is a major reference in the thesis. He is criticized[26] for his extreme position, namely, for rejecting a representation of uncertainty in terms of numerical probabilities not only in ambiguous situations but in any situation of crucial importance to the decision-maker. However, argues Ellsberg, "when ambiguity is extreme, by any of his indices: relevant information sparse, or obviously unreliable and contradictory; wide differences in the expressed expectations of different individuals; low confidence in available estimates," Shackle's "somber reflections" seem "too ominously relevant to the very circumstances upon which this study focuses to be dismissed" (Ellsberg 2001, pp. 16-17).[27]

A point is worth noting. Granted that Ellsberg (2001, p. 2) started his analysis by defining decision-making under ambiguity as a special case of decision-making under uncertainty, he contended that the arguments he was raising against the use of subjective expected utility in case of ambiguity, as well as the solutions he was proposing, were valid a fortiori to deal with the "broader problem" of uncertainty.

Ellsberg's thesis contains a more accurate discussion of the examples known from the QJE article and a full treatment of the philosophical background of his stance. More relevantly from our viewpoint, there is a comparison between alternative criteria for explaining the observed departures from Savage's axioms, and this comparison includes decision rules not discussed in the article. In particular Ellsberg considers both maximin expected utility and the more general α-maximin version put forward by Hurwicz. As in the 1961 article, Ellsberg aimed to find out a decision rule alternative to subjective expected utility that could account for the behaviour of unrepentant violators of Savage's axioms. In 1961 Ellsberg's solution had been a weighted average of the expectation of the most reliable ("best guess") probability distribution and the maximin solution. The idea came from Hodges and Lehmann (1952), who had taken as starting point a set of plausible probability distributions as priors, instead of a unique probability, and had discussed how the decision-maker could use them. A decision-maker who fails to rule out of a set of distributions any one as unacceptable, nonetheless may regard one of them as the most reliable.[28] Accordingly the

---

[26] As in the 1961 article, Ellsberg quotes Shackle's example of coin-tossing in cricket test matches. Here Shackle, in order to emphasize his point, unintelligibly refuses to take into account statistical information on the aleatory mechanism. Shackle's insistence that, under certain conditions, all unique events are crucial and therefore frequency-ratios are irrelevant contributed to his marginalization.

[27] This consideration of Ellsberg reflects the attitude of many critics of Shackle, who at the time of *Expectations in Economics* showed great interest in Shackle's insights but subsequently interest faded in the face of the cumbersome formal structure. On this point see Zappia and Basili 2005.

[28] Ellsberg (2001, p. 182) summarises the rationale under the use of multiple probabilities as follows: "under most circumstances of decision-making you can *eliminate*, from the set Y of all

decision rule adopted by Ellsberg (1961, p. 664) was to associate the following index with each act *x*: ρ E(*x*) + (1-ρ) min(*x*).[29] The parameter $\rho \in [0,1]$ expresses reliability, that varies between absence of confidence and absolute confidence, and that weighs the additive distribution serving as a best estimate as well as all other possible probability distributions assumed to be reasonable by a decision-maker under ambiguity.

In the 1962 thesis Ellsberg, in accordance with Hodges and Lehmann, retains the idea of a set of distributions over the states of the world, but applies the Hurwicz criterion to the restricted set of possible distributions. The new index, called "restricted Bayes/Hurwicz criterion," is: ρ E(*x*) + (1-ρ) [α max(x) + (1-α) min(*x*)]. Ellsberg presented a taxonomy of the possible criteria in terms of the two parameters ρ and α, that is in terms of the degrees of confidence and optimism, respectively, characterising the epistemic state of the decision-maker.

Two points are worth making before moving on to a comparison between Ellsberg and Shackle. First, in discussing Hurwicz Ellsberg (2001, p. 163, 190) notes the analogies of Hurwicz's rule with Shackle's analysis. Ellsberg infers that Shackle's decision rule amounts to a generalization of Hurwicz's criterion: "the Hurwicz rule corresponds to a special case in which the indifference sets in Shackle's gambler indifference map are parallel straight lines, with slope dependent on α." In Ellsberg's thesis, thus, it is argued for the first time that Shackle's criterion for ranking acts is nothing but a generalized version of Hurwicz's criterion. As noted by Levi (2001, xxiii) in his introduction to Ellsberg 2001, it is regrettable that Luce and Raiffa did not note this in their influential survey of decision making under uncertainty (Luce and Raiffa, 1957, Chap. 13).

The second point concerns the significance of the decision rules Ellsberg privileged. To a large extent Ellsberg referred to a model of ambiguity aversion currently known as the maximin expected utility model (MEU) after Gardenfors and Sahlin (1982). This model, by virtue of the axiomatisation provided by Gilboa and Schmeidler (1989), is possibly the most referred to in the current literature in decision-making under uncertainty. Moreover, as will be

possible probability distributions over the relevant states of the world, certain distributions as *unacceptable* for representing your opinion … But … there may remain a sizeable subset Y° of distributions … that still seem 'reasonable acceptable' … that do not contradict your ('vague') opinions [and that] may yet be large, particularly when relevant information is perceived as scanty, unreliable, contradictory, *ambiguous*."

[29] Where E(*x*) is the expected payoff to the act corresponding to the best guess distribution, min(*x*) is the minimum expected payoff to the act as the probability distribution ranges in the set of non-unacceptable distributions, and *ρ* represent the degree of confidence in the best guess distribution. Following Hodges and Lehmann, Ellsberg called the criterion "restricted Bayes." The adjective restricted hints of course to the set of plausible priors.

seen in the next section, recent developments (see for instance Girardato, Maccheroni and Marinacci 2003) put to work the intuition under Hurwicz (and Shackle) criterion in their axiomatisation of the so-called α-MEU criterion.

5. Shackle's and Ellsberg's frameworks compared

As seen in the previous sections, awareness of missing information and partial knowledge is crucial to an understanding of the implications of Shackle's approach to decision-making. Shackle's entrepreneur is supposed to invest in a context in which the list of possible future states of the world is not necessarily complete. As a result, Shackle's entrepreneur does not maximise expected utility, and uses a peculiar criterion for ranking acts.

Shackle's decision problem shares an essential feature with the decision problem from which Ellsberg paradox emerges. In our view, both decision-making problems refer to an epistemic state that can be represented by the notion of ambiguity, a notion which belongs to the region between the two extremes of complete ignorance and risk.[30] At first the two problems seem quite different since to Ellsberg the question is one of ambiguous probabilities with a complete list of all possible events, whereas Shackle aims to discuss situations in which the list of possible events in not exhaustive. But this difference has more to do with the degree than with the quality of uncertainty faced by decision-makers. In fact, in both decision-making problems the individual faces two kinds of possible events, that is unambiguous and ambiguous events. The decision-maker has a clear perception of the unambiguous events and is able to attach a unique probability distribution to them; this means that her choice involves risk only. But she has also a vague knowledge of ambiguous events and she considers only her own representation of an uncertain situation as an approximate description of the "true" scenario. In this second case she is unable to attach a unique probability to each state of the

[30] Putting the degree of knowledge on an axe, we can establish the following order: certainty, risk, ambiguity, and complete ignorance. It ensues that characterizing these points by means of probability would imply the following definitions: (a) certainty: the probability of a state A in the state space is one; (b) risk: there exists a single additive and fully reliable probability distribution, either objective or subjectively determined, on the state space; (c) ambiguity: there exist more than one additive probability distribution, or an interval of probability for each state is assumed to be possible. Both a capacity (non-additive probability) on the state space and a fuzzy measure are alternative representations of the epistemic state of the decision-maker; (d) complete ignorance: either all states of the world have the same probability of occurring, on the basis of the principle of insufficient reason, or there is no reliable probability distribution, on the basis of the rejection of the principle of insufficient reason.

world because either she does not know them or she is unable to distinguish clearly among them. She can be unable to list each state included in an event, but she can also be unable to attach the "true" probability to each state. As a consequence, the decision-maker's ambiguity deriving from incomplete knowledge can be represented by measures that are alternative to a single additive probability, like an interval of probabilities, a set of probabilities, a capacity, or a fuzzy measure.

In the early 1950s there was a notable interest in Shackle's theory exactly because he had made the point just considered explicit.[31] In facing situations that cannot be represented by a single additive probability distribution over their outcomes, the decision-maker has to evaluate alternative courses of action by an alternative measure, which Shackle called potential surprise. This measure entails the use of a procedure of choice based on non-additive expected value.[32] Both Shackle and Ellsberg assume that the decision-maker has partial knowledge about the future states of the world, and faces ambiguity in the sense of our definition, that is, she is aware of the unreliability of a single, additive probability distribution. What is relevant, as a result, is the representation of the "small world" in which the decision-maker has to act. Once assessed from this viewpoint Shackle's and the Ellsberg's scenarios are analogous; in fact both are misrepresentations of the hypothetical "grand world" in which all future states of the world are (potentially) completely described.

The distinction between small and grand worlds is credited to Savage. The grand world is a complete list of the states that are of concern to an individual. The small world is a construction derived from a certain partition of the grand world into events, a partition which

---

31 See for instance Arrow 1951. Arrow (1951, p. 404) gave sizeable relevance to Shackle's analysis because it was the single formalised theory among those put forward by a group of authors who, following in Knight and Keynes's footsteps, "do use other than probability statements in their description of behaviour under uncertainty." Zappia and Basili (2005) assess the intellectual environment in which Shackle's theory was discussed in the years between *Expectations in Economics* (Shackle 1949) and the last restatement of its main theses in *Decision, Order and Time in Human Affairs* (Shackle 1961).

32 It is worth noting that even in Shackle's scenario an investment decision can be considered as a bet. But Shackle rejected the main consequence of de Finetti's notion of probability as a bet, that is, the corollary that if someone is rejecting a bet she is accepting its complementary bet. Shackle argued as follows: "when the uncertainty in a person's mind arises from the plurality of the answers which suggest themselves to him for some one question, these answers are rivals mutually excluding each other. To believe in one of these answers is therefore to disbelieve in the others. By contrast it is not true that to disbelieve in one answer is to believe in the others" (Shackle1961, 74). This is a necessary consequence of the assumption of non-additivity. As is usually recognised in current decision theory, under ambiguity the decision-maker may reject both sides of the bet, and her epistemic attitude can be represented by either capacities or non-additive measures.

constitutes the states of the small-world.[33] Savage (1954, p. 9) maintained that an individual usually confines her attention to a relatively simple situation in almost all decisions; this amounts to say that, in practice, the individual is always concerned with a small world, which is "derived from a larger [one] by neglecting some distinctions between states, not by ignoring some [grand-world] states outright." By taking a small world as the proper context of her decisions, the decision-maker obtains an approximate description of the states of the world and their consequences. In principle, the individual can gradually come to consider an even more refined and detailed small world, until she arrives at the grand world where everything is taken into account. However, Savage's point is that it is "utterly ridiculous" to pretend that "one envisages every conceivable policy for the government of his whole life (at least from now on) in its most minute details, in the light of a vast number of unknown states of the world" (Savage 1954, p. 16).

Savage (1954, pp. 82-84) claimed that subjective expected theory should be applied only to small worlds. In fact, it is only in small worlds that all possibilities can be exhaustively enumerated in advance, and all implications of all possibilities explored in detail; hence the possibilities can be labelled and placed in their proper positions as in the famous omelette example. Savage stressed the "practical necessity of confining attention to, or isolating, relatively simple situations in almost all applications of the theory of decision developed."[34] But he admitted that the representation provided by a small world "is completely satisfactory only if it is actually a *microcosm*, that is, only if it leads to a probability measure and a utility well articulated with those of the grand world" (Savage 1954, p. 88). Leaving utility aside, in Savage's view there is no certainty that the probability of an event in the small world equals the probability of the corresponding collection of states in the grand world. If probabilities are different at the two degrees of refinement, the probabilities attached in the small world are "correct" only if they equal those calculated in the grand world. Thus the implicit assumption is that the individual should ultimately be able to deal with the grand world, since the condition which assures equality in probability "seems incapable of verification without taking the grand world much too seriously" (Savage 1954, p. 90). There is a decisive implication here: if the theory is to be consistently applied, the decision-maker should be able to enumerate exhaustively all possibilities in advance, and to

[33] A "small world … is determined not only by the definition of a state, but also by the definition of small-world consequences. A *small-world consequence* is a grand-world act" (Savage 1954, p. 85).

[34] Savage's rejection of the critiques of his theory arguing that "real people frequently and flagrantly behave in disaccord with utility theory," is mainly based on the distinction between the grand and the small world (Savage 1954, pp. 100-101).

explore all consequences in detail, though she works exclusively in a practical setting called the small world; thereby it is as if she had a sort of "divine" knowledge of the outside world. As a result, in situations in which outcomes and states are not clearly given in the description of the problem, it is clear neither what the normative implications of Savage's sure-thing principle are, nor why Savage's expected utility approach should inform actual behaviour.

This representation of the decision-theoretic context in terms of small and grand world has a counterpart in modern decision theory. In Gilboa and Schmeidler (1994), in particular, the space of real-valued functions on the original space (the small world in which the decision-maker acts) is embedded in a larger state space (the grand world) in an integral-preserving way. In this way the set of states of the small world that is relevant for the decision-maker may be a mis-specified representation because either certain possible states that are accounted for in the larger space are missing or she is unable to split an event into its constitutive states (singleton) in the small world. If "omitted stated of the world were introduced into the model explicitly the non-additivity would disappear" (Gilboa and Schmeidler 1994, p. 13).

It is time to sum up our argument on the analogy between Shackle and Ellsberg. Both authors talk of decision-making problems located in Savage's small world. But Shackle and Ellsberg's small worlds are not microcosm, that is, the decision-maker is unable to split an event in the small world into the states composing the event in the grand world. To be precise, in both scenarios the decision-maker is unable to enumerate the single mutually exclusive possible future states of the world. In other words, the decision-maker has only a rough representation (partition) of the set of states of the world. That is, she acts on the basis of a partition that is a not fine enough and, as a result, she cannot apply Savage's rule for the optimal choice.[35]

This interpretation of Shackle's theory is not customary especially in Keynesian circles, which emphasize that the unanticipated events Shackle referred to need to be outside the list of states (inaccurately) assumed to be exhaustive in probabilistic decision-making (for instance, see Davidson 1983 and Dow 1995).[36] However, our point is that the "novelty"

---

[35] Savage himself observes that if the decision-maker is unable to define a microcosm she can not apply the maximisation of subjective expected utility

[36] In his recent assessment of Shackle's probability theory, Runde (2000) maintains that Shackle's objections to the assumption about the absence of the residual hypothesis remain unaddressed in orthodox theory. Runde refers to Gilboa and Schmeidler's non-additive probability in a footnote, but contends that it addresses situations, like the Ellsberg paradox, "where the problem is one of ambiguous probabilities rather than that of providing an exhaustive list of states." We have tried to show that this distinction is no longer accurate.

characterising choices in uncertain situations can be more fruitful interpreted as an inability to discern the elementary states composing the event, an inability that may be solved, given Shackle's approach, through the passing of time. Shackle (1956, p. 115) himself hinted at this interpretation in a paper on the "logic of surprise:"

"An actual experience will cause surprise when an imperfect image of it has been formed in advance; and the imperfectness can consist either in some wrong characteristics or details being specified in place of the right ones, or in blanks having being left in the picture, that is, it can consist in characteristics or details which in fact belong to the experience having being omitted altogether; or again we can say, in the experience having a 'dimension' beyond the list of those which the individual can specify in advance."

The point we are emphasizing is analogous to the one made by Arrow and Hurwicz in their 1972 paper, often quoted in the literature on "complete ignorance" as the main building block for theories of decision-making alternative to Savage's (for instance, see Barrett and Pattanaik 1993).[37] Arrow and Hurwicz (1972, p. 1) provided "a possible characterisation of the concept of complete ignorance," which they defined as, first, "a situation in which there is no a priori information available which gives any state of nature a distinguished position", and, second, as a situation "not presupposing a fixed list of states of nature." In situations of complete ignorance, Arrow and Hurwicz (1972, p. 2) argued, the subjective probability framework can only assign equal probabilities to all states of nature, in accordance with Laplace's principle of insufficient reason. Shackle's formulation, on the other hand, "permits to interpret complete ignorance as meaning that all states of nature have zero potential surprise. Then dividing a state of nature into two would have no effect on the action chosen, if the reward to an action is the same under either substate." Arrow and Hurwicz concluded that if a description of the world needs making finer than previously anticipated (a likely situation if complete ignorance is assumed), this can be done easily in Shackle's framework, but not in the subjectivists.'[38] As seen in the previous sections, Shackle viewed this capability of accounting for new, unanticipated events as the main reason for dispensing with probability measures and the use of potential surprise in its place.

[37] As recalled above, this is the published version of Hurwicz 1951a, including the basic proof provided by Arrow (1953) that any criterion satisfying certain basic properties under complete ignorance must be of the Hurwicz(-Shackle) type. On this point see also the discussion in this section.

[38] In the subjectivist approach, if the event to which choices are conditioned is not in the list of the possible states the decision maker is forced to modify the probability weights originally attributed to the states, because these weights add up to 1. As noted in the previous section, Shackle insisted that a correct representation of the epistemic state of an individual should make room for a "residual hypothesis" - that is, the realisation of an unanticipated state - without reducing the weights attributed to the original list of states. However, this can be achieved also in a probabilistic approach, if the probability measure on the original list of states is superadditive.

After dealing with the main similarity between Ellsberg's notion of ambiguity and Shackle's notion of fundamental uncertainty, two other similarities remain to be examined. Both similarities were highlighted by Ellsberg in his doctoral thesis, and are discussed at length by Levi (1972 and 1980).

The first important aspect that makes the two decision-problems analogous emerges from a comparison between the valuation functions of the hypothetical gains and losses (the "degree of stimulus," in Shackle's set-up), and the "reliability" that characterizes the decision-maker's information in Ellsberg's paradox. As seen above, the degree of stimulus can be interpreted as a function representing the decision-maker's degree of confidence in her representation of the (small) world.[39] As already illustrated, Shackle and Ellsberg share a method for evaluating the value of an ambiguous act that is different from Savage. Moreover, Shackle and Ellsberg's approaches are analogous as to the conclusion they reach. After recognizing the existence of ambiguity, they assume that the decision-maker will take into account her own attitude with respect to ambiguity, that is, she will weighs her beliefs with a judgment about the quality of her knowledge. The measure they use is nothing but an alternative formulation of Keynes's weight of argument, as Ellsberg himself recognised (1961, p. and 2001, p.). But, contrary to Ellsberg's opinion, this holds true not only if the weight of argument is interpreted as a measure of the amount of evidence relevant to a certain hypothesis, but also if it is interpreted as the degree of completeness of the evidence. As argued in section 3, if probabilities are assumed to be non-additive, even using the second definition the weight ranges from 0 to 1, and represents the reliability of the information at hand.

The second main aspect is related to the criterion for choice Shackle and Ellsberg proposed. In the main both authors hinted at the same procedure Hurwicz adopted, with the relevant difference that Ellsberg used Hurwicz criterion while Shackle proposed a decision rule which turned out to be of the Hurwicz type. In their 1972 paper, they clarified from the outset that under these hypotheses the criterion they provided differed from that formulated within the "now more standard subjective probability framework," and that their arguments and conclusions "are much closer to Shackle's (1949) than those of Ramsey (1931), de Finetti (1937), and Savage (1954)." Arrow and Hurwicz were the first authors (after Ellsberg in his unpublished thesis) to make the point that Shackle's criterion was more accurate than

39 This point is implicitly made by Arrow and Hurwicz (1972, p.2) when they argue that "the standardized focus-gain and focus-loss become simply the maximum and minimum payoff to a given action, and the final decision among possible actions is made on the basis of the gambler indifference map, which in this case of complete ignorance, simply order these pairs."

previously admitted. Proceeding along the lines of Chernoff (1954) and Milnor (1954), Arrow and Hurwicz (1972) postulated, first, a set of "desirable" properties of a rational criterion of choice. These are mostly properties regarding the relationships between sets of actions, identified as decision problems; these sets permit the identification of the minimum and the maximum value associated with each decision problem under any particular state of nature. Second, they proved that the necessary and sufficient conditions guaranteeing that an optimality criterion possesses the postulated properties are related only to the minimum and the maxim value. That is, any criterion consistent with those desirable properties ranks actions solely on the basis of their best and worst possible outcomes, and not on the basis of averages. Finally, Arrow and Hurwicz (1972, p. 2) commented on Shackle's theory that if the focus values of Shackle are interpreted as the minimum and maximum payoff to a given action, and the final decision among possible actions is made on the basis of the gambler indifference map, "we demonstrate that a plausible set of desirable properties for a rational criterion of choice under complete ignorance in fact leads to this special case of Shackle's theory."

It is appropriate to outline the importance of the Arrow-Hurwicz criterion for modern decision theory. As noted earlier the Arrow-Hurwicz criterion has originated a literature on decision making under complete ignorance. Savage's subjective expected utility theory applies to situations in which the individual acts as if he had a single reliable probability distribution over the states of the world (elicited from choices). Conversely, Arrow and Hurwicz introduced the case in which the individual has no idea at all about the likelihood of states. Gilboa and Schmeidler (1989) have proposed an alternative to subjective expected utility to encompass the case between these two extremes, that is the case in which the individual has opinions about the likelihood of different states, but she is not able to assign exact probabilities to them. According to their theory the decision maker has a convex set of subjective probability, which expresses the range of probabilities she considers possible. Since the subjective probability is not unique there is a set of expected utilities for each action. Gilboa and Schmeidler then propose the following criterion: an action *a* is preferred to *b* if and only if the minimum possible value of the expected utility of *a* is greater than the minimum expected value of *b*. They have called this criterion maximin expected utility (MEU). If the set of probabilities consists only of a single probability distribution, maximin expected utility coincides with subjective expected utility. If, on the other hand, it consists of all possible probability distributions it coincides with Wald's maximin.

There is a close relationship between maximin expected utility and non-additive probability theory which needs emphasising. Gilboa and Schmeidler (1994) showed that there is an isomorphism between a non-additive probability measure and a convex set of additive probability measure. The interest in non-additive measure stays in the fact that the behaviour of decision-makers exhibiting a pessimistic attitude towards uncertain situations and can be rationalised through the use of a sub-additive probability measure. As noted earlier, in both the contexts discussed by Ellsberg and Shackle the decision-maker's beliefs can be represented by non-additive probabilities (or capacities).[40] Since the capacity is a non-additive measure, the integration of a real-valued function with respect to it is impossible in the Lebesgue sense. Schmeidler (1989) has provided the axiomatic structure to show that the proper integral for a capacity is the Choquet integral. The Choquet integral with respect to a capacity is a generalisation of the Lebesgue integral, which requires that states of the world have been ranked from the most to the least favourable ones, or vice versa, with respect to their consequences. As a result, the Choquet integral is a generalisation of the mathematical expectation usually used in expected utility models with respect to a capacity.[41]

Choquet Expected Utility (CEU) accommodates both Shackle's and Ellsberg's evidences. Moreover if the convex capacity that represents the decision-maker's beliefs has a non-empty core of additive measures, than the CEU with respect to the capacity equals the MEU of the set of additive measures, that is, the SEU with respect to the less favourable probability distribution in the core. In particular if a *simple capacity* is considered to model ambiguity it is evident that a non-additive measure is only a contraction of a probability distribution, except that for the universe.[42] The utility of an act is only the weighted average utility, or the Choquet integral with respect to a capacity is the difference between minima of regular integrals over sets of additive measures (Gilboa and Schmeidler 1994).

More recently a slightly different measure has been proposed under the heading of α-MEU approach. Girardato, Maccheroni and Marinacci (2005) assume that the decision-maker's ambiguity is expressed by a set of additive probability distributions (multiple priors)

40 A convex capacity is monotone measure that is normalised to 1 on the full set and 0 on the null set, like probability, but, unlike probability, the sum of the capacities of two subsets may be strictly less than the capacity of the union of these sets. The convexity of the capacity is the property suggested by the Ellsberg Paradox.

41 The Choquet integral of an act *f* with respect to the capacity μ is

$$\int f d\mu = \int_0^{\infty} \mu(\{w | f(w) \geq t\})dt + \int_{-\infty}^{0} [\mu(\{w | f(w) \geq t\}) - \mu(\Omega)]dt$$

42 The capacity υ is called simple whenever $\upsilon(A) = \begin{cases} 1 if A = \Omega \\ \gamma.\pi(A) if A \subseteq \Omega \end{cases}$ where $\gamma \in [0,1]$ and π is an additive probability distribution.

and the parameter α represents her ambiguity attitude.[43] The α-MEU preference model is a generalization of the Hurwicz's maximin functional in which the parameter α is constant.[44] As a consequence, the Shackle's and the Ellsberg's approach can be considered as variations of Hurwicz's maximin expected utility criterion (α-MEU).

6. Conclusion

This paper has discussed the relationship between Ellsberg and Shackle in terms of their representation of the environment for actual decision-making. Shackle (1955) rejected probabilistic reasoning per se, and accordingly formulated a non-probabilistic theory of decision-making under uncertainty. Ellsberg (1961), who stressed the failure of Savage's sure-thing axiom in ambiguous contexts, sought criteria of decision-making alternative to subjective expected utility. In his doctoral thesis Ellsberg (2001) noted that the decision rule that accommodates for decisions taken by violators of Savage's axiom (a rule put forward by Hurwicz) was a special case of Shackle's decision rule.

We have argued that the similarities between Ellsberg and Shackle are more relevant than Ellsberg himself recognised. First, we have emphasised that both authors' analyses have a common origin in Keynes's theory of probability and, more specifically, in a peculiar interpretation of Keynes's weight of argument. Second, we have interpreted Shackle's fundamental uncertainty and Ellsberg's ambiguity as two instances of the same representation of the decision problem, namely one in which the uncertainty characterising the environment is expressed in terms of a "small world" interpretation.

Our historical reconstruction is meant to have theoretical implications. Basically, our claim is that Keynes, Shackle and Ellsberg can decisively enrich decision analysis by virtue of

---

43 The α-MEU preferences, $\alpha \in [0,1]$ are represented by the functional

$$I(f) = \alpha \max_{\pi \in \Pi} E_{\pi}(u \circ f) + (1-\alpha) \min_{\pi \in \Pi} E_{\pi}(u \circ f)$$

where Π is a set of additive probability distributions.

44 The α-MEU approach establishes a connection between the opposite valuations (the best and the worst) of each act, given a set of relevant probability distributions. The functional shows a sandwiching property, because it is placed between the worst and the best scenario evaluation of the decision maker. Straightforwardly, the assumed constant function α is an index of ambiguity aversion: the lower (higher) it is, the smaller (larger) the pessimism in the decision maker's evaluation is. If α is equal to zero, then a standard CEU functional is obtained. In a similar vein Basili, Chateauneuf and Fontini (2005) proposes a functional in which the decision maker is supposed to split outcomes between familiar and un-familiar ones.

certain recent developments in non-additive probability theory, which make it possible to incorporate their idiosyncratic views in a substantial body of theory.

REFERENCES

Anand, P. (1991), "The Nature of Rational Choice and *The Foundations of Statistics*," *Oxford Economic Papers*, 43: 199-216.

Arrow, K. and Hurwicz, L. (1972), "Decision making under ignorance," in C. F. Carter and J. L. Ford (eds.), *Uncertainty and Expectations in Economics. Essays in Honour of G. L. S. Shackle*. Oxford: Basil Blackwell.

Arrow, K. J. (1951), "Alternative approaches to the theory of choice in risk-taking situations," *Econometrica*, 19: 404-437.

Arrow, K. J. (1953), "Hurwicz's optimality criterion for decision making under ignorance," *Technical Report* 6, Stanford University.

Barrett, C. R. and Pattanaik, P. K. (1993), "Decision making under complete uncertainty," in S. Sen (ed.), *Uncertainty and Risk in Economic Life*. Aldershot: Edward Elgar.

Basili, M., Chateauneuf, A. and Fontini, F. (2005), "Choices Under Risk and Ambiguity with Familiar and Unfamiliar Outcomes," *Theory and Decision*, forthcoming.

Becker, S. and Brownson, F. (1964), "What price ambiguity? Or the role of ambiguity in decision making," *Journal of Political Economy*, 72: 63-73.

Bernoulli, J. (1713), *Ars Conjectandi*. Basle.

Bowman, M. J., ed. (1958), *Expectations, Uncertainty and Business Behaviour*. New York: Social Science Research Council.

Brewer, K. (1963), "Decision under uncertainty: comment," *Quarterly Journal of Economics* 77: 159-161.

Brewer, K. and Fellner, W. (1965), "The slanting of subjective probabilities: agreements on some essentials," *Quarterly Journal of Economics* 79: 657-663.

Carabelli, A. (1988), *On Keynes's Method*. London: Macmillan.

Carabelli, A. (2002), "Speculation and reasonableness: a non-Bayesian theory of probability," in Dow, S. and Hillard, J. (eds.), *Keynes, Uncertainty and the Global Economy: Beyond Keynes*, vol. II. Cheltenham: Elgar.

Chernoff, H. (1954), "Rational selection of decision functions," *Econometrica*, 22: 422-443.

Davidson, P. (1983), "Rational expectations: a fallacious foundation for studying crucial decision making," *Journal of Post Keynesian Economics*, 5: 182-198.

Dow, S. (1995), "Uncertainty about uncertainty," in Dow, S. and Hillard, J., eds. *Keynes, Knowledge and Uncertainty*. Aldershot: Edward Elgar.

Edwards, W. (1954), "The theory of decision making," *Psychological Bullettin*, 5: 380-417.

Edwards, W. (1958), "Note on Potential Surprise and Nonadditive Subjective Probability," in M. J. Bowman (ed.), *Expectations, Uncertainty and Business Behaviour*. New York: Social Science Research Council.

Ellsberg, D. (1961), "Risk, ambiguity, and the Savage axioms," *Quarterly Journal of Economics*, 75: 643-669.

Ellsberg, D. (1963), "Risk, ambiguity, and the Savage axioms: reply," *Quarterly Journal of Economics*, 77: 336-342.

Ellsberg, D. (2001), *Risk, Ambiguity and Decision*. New York: Routledge.

Feduzi, A. (2005), "Re-examining Keynes's conception of the weight of arguments and the Ellsberg paradox," mimeo, paper presented at the 2nd Annual Meeting of StorEP, Siena, June 2005.

Fellner, W. (1961), "Distortion of subjective probabilities as a reaction to uncertainty," *Quarterly Journal of Economics* 75: 670-689.

Gardenfors, P. and Sahlin, N. E. (1982), "Unreliable probabilities, risk taking and decision-making," *Synthese*, 53: 361-86.

Georgescu-Roegen, N. (1958), “The nature of expectation and uncertainty,” in M. J. Bowman (ed.), *Expectations, Uncertainty and Business Behaviour*. New York: Social Science Research Council.
Ghirardato, P., Maccheroni, F. and Marinacci, M. (2005), “Differential ambiguity and ambiguity attitude,” *Journal of Economic Theory*, forthcoming.
Gilboa, I. and Schmeidler, D. (1989), “Maxmin expected utility with a non-unique prior,” *Journal of Mathematical Economics*, 18: 141-153.
Gilboa, I. and Schmeidler, D. (1994), “Additive representations of non-additive measures and the Choquet integral,” Annals of Operations Research, 52: 43-65.
Good, I. J. (1950), *Probability and the Weighing of Evidence*. London: Charles Griffin.
Hacking, I. (1975), *The Emergence of Probability*, Cambridge: Cambridge University Press.
Hodges, J. L. and Lehmann, E. L. (1952), “The Uses of Previous Experience in Reaching Statistical Decision,” *Annals of Mathematical Statistics*, 23: 396-407.
Hurwicz, L. (1951a), “Optimality criteria for decision making under ignorance,” *Cowles Commission Discussion Paper* 370, Chicago.
Hurwicz, L. (1951b), “Some specification problems and application to econometric models” (abstract), *Econometrica*: 19: 343-344.
Kahnemann, D. and Tversky, A. (1979), “Prospect theory: An analysis of decision under risk,” Econometrica, 47: 263-291.
Kelsey, D. (1993), “Choice under partial uncertainty,” International Economic Review, 34: 297-308.
Kelsey, D. (1994), “Maxmin Expected Utility and Weight of Evidence,” *Oxford Economic Papers*, 46: 425-44.
Kelsey, D. and Quiggin, J. (1992), “Theories of choice under ignorance and uncertainty,” *Journal of Economic Surveys*, 6: 133-153.
Keynes, J. M. (1921), *A Treatise on Probability*. *The Collected Writings of John Maynard Keynes*, Vol. VIII. London: Macmillan, 1973.
Knight, F.H. (1921), *Risk, Uncertainty and Profit*. Chicago: University of Chicago Press, 1985.
Koopman, B. O. (1940), “The Axioms and algebra of intuitive probability,” *Annals of Mathematics*,. 41: 269-92.
Lachmann, L. (1976), ‘From Mises to Shackle: an essay on Austrian economics and the kaleidic society’, Journal of Economic Literature, 14: 54-62.
Langlois, R. N. (1994), “Risk and uncertainty,” in *The Elgar Companion to Austrian Economics*, edited by P. Boettke. Cheltenham: Edward Elgar.
Lawson, T. (1985), “Uncertainty and economic analysis,” *Economic Journal*, 95: 909-927.
Levi, I, (1966), “On potential surprise,” *Ratio*, 8: 107-129.
Levi, I. (1972), “Potential surprise in the context of inquiry,” in C. F. Carter and J. L. Ford (eds.), *Uncertainty and Expectations in Economics. Essays in Honour of G. L. S. Shackle*. Oxford: Basil Blackwell.
Levi, I. (1980) ‘The Enterprise of Knowledge’, Cambridge, Mass: MIT Press.
Levi, I. (2001), “Introduction,” in Ellsberg 2001.
Luce, D. and Raiffa, H. (1957), *Games and Decisions*. New York: Wiley.
Machina, M. (2001), “Bibliografical note,” in Ellsberg 2001.
Milnor, J. (1954), “Games against Nature,” in R. M. Thrall, C. H. Coombs and R. L. Davis (eds.), *Decision Processes*. New York: Wiley.
Mukerji, S. (1997), “Understanding the non-additive probability decision model,” *Economic Theory*, 9: 23-46.
O’Donnell, R. M. (1989), *Keynes: Philosophy, Economics and Politics*, London: Macmillan.

Quiggin, J. (1982), "A theory of anticipated utility," *Journal of Economic Behaviour and Organization*, 3: 323-343.

Raiffa, H. (1961), "Risk, ambiguity, and the Savage axioms: comment," *Quarterly Journal of Economics*, 75: 690-694.

Runde, J. H. (1990), "Keynesian uncertainty and the weight of argument," *Economics and Philosophy*, 6: 275-92.

Runde, J. H. (1994), "Keynesian uncertainty and liquidity preference," *Cambridge Journal of Economics*,18: 129-44.

Runde, J. H. (2000), "Shackle on probability," in P. E. Earl and S. F. Frowen, eds., *Economics as an Art of Thought. Essays in memory of G. L. S. Shackle*. London: Routledge.

Runde, J. H. (2001), "Chances and choices: notes on probability and belief in economic theory," in Maki, U. (ed.), *The Economic World View: Studies in the Ontology of Economics*. Cambridge: Cambridge University Press.

Savage, L. J. (1954), *The Foundation of Statistics*. New York: Wiley, 1972, 2nd ed.

Schmeidler, D. (1989), "Subjective probability and expected utility without additivity," *Econometrica*, 57: 571-587.

Shackle, G. L. S. (1949), *Expectations in Economics*. Cambridge: Cambridge University Press.

Shackle, G. L. S. (1949-50), "A non-additive measure of uncertainty," *Review of Economic Studies*: 17: 70-74.

Shackle, G. L. S. (1952), "On the meaning and measure of uncertainty: I," *Metroeconomica*, reprinted in Shackle 1955.

Shackle, G. L. S. (1953a), "On the meaning and measure of uncertainty: II," *Metroeconomica*, reprinted in Shackle 1955.

Shackle, G. L. S. (1953b), "The logic of surprise," *Economica*, reprinted in Shackle 1955.

Shackle, G. L. S. (1955), *Uncertainty in Economics and Other Reflections*. Cambridge: Cambridge University Press.

Shackle, G. L. S. (1961), *Decision, Order and Time in Human Affairs*. Cambridge: Cambridge University Press.

Shackle, G. L. S. (1972), *Epistemics and Economics*. Cambridge: Cambridge University Press.

Shackle, G. L. S. (1976), "Time and choice," *Proceedings of the British Academy*, 62: 309-329.

Shafer, G. (1976), *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.

Shafer, G. (1978), "Non-additive probabilities in the work of Bernoulli and Lambert," *Archive in the History of Exact Science*, 19: 309-370.

Smith, C. A. B. (1961), "Consistency in statistical inference," *Journal of the Royal Statistical Society*, 23: 1-25.

Tversky, A. and Kahneman, D. (1992), "Advances in prospect theory: cumulative representation of uncertainty," *Journal of Risk and Uncertainty*, 5: 297-323.

Tversky, A. and Wakker, P. (1995), "Risk attitudes and decision weights," *Econometrica*, 63: 1255-1280.

Vaughn, K. (1994), *Austrian Economics in America. The Migration of a Tradition*. New York: Cambridge University Press.

Vercelli, A. (1999a), "Peso dell'argomento e decisioni economiche," in Marzetti Dall'Aste Brandolini, S. e Scazzieri, R. (a cura di), *La Probabilità in Keynes: Premesse e Influenze*. Bologna: Clueb.

Vercelli, A. (1999b), "The recent advances in decision theory under uncertainty: a non-technical introduction," in Luini, L. (ed.), *Uncertain Decisions: Bridging Theory and Experiments*. Dordrecht: Kluwer.

von Neumann, J. and Morgenstern O. (1944), *Theory of Games and Economic Behaviour*. Princeton: Princeton University Press.

Wald, A. (1950), *Statistical Decision Functions*. Wiley, New York.

Walley, P. (1991), *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall.

Weckstein, R. (1959), "Probable knowledge and singular acts," *Metroeconomica*, 11: 104-118.

Yaari, M. E. (1987), "The dual theory of choice under risk," *Econometrica*, 55: 95-115.

Zappia, C. (1998), "Radical subjectivism and Austrian economics," in R. Koppl and G. Mongiovi (eds.), *Subjectivism and Economic Analysis*. London: Routledge.

Zappia, C. and Basili, M. (2005), "Shackle versus Savage: non-probabilistic alternatives to subjective probability theory in the 1950s," *Quaderni del Dipartimento di Economia Politica*, 452, University of Siena.