



UNIVERSITÀ  
DI SIENA  
1240

QUADERNI DEL DIPARTIMENTO  
DI ECONOMIA POLITICA E STATISTICA

**Marcello Basili  
Carlo Zappia**

Ellsberg's Decision Rules and Keynes's Long-Term Expectations

**n. 777 – Febbraio 2018**



## **Abstract**

This paper presents an intuitive way to represent Keynes's notion of long-term expectations and its implications for decision-making, using the so-called  $\varepsilon$ -contamination approach. Further to a suggestion by Ellsberg, a coherent Keynesian expectational function for decisions under uncertainty is derived. The paper draws on the similarities between the analyses of Keynes and Ellsberg and contends that much of current decision theory under ambiguity follows in Keynes's footsteps.

*Keywords:* uncertainty, expectations, Keynes, consensus distribution, epsilon-contamination.

*JEL classification:* B26, D81

## 1. Introduction

On uncertainty issues, economics literature often refers to Keynes's notion of long-term expectations as a determinant of investment decisions. However, the debate on the modern relevance of Keynes's analysis of behaviour under uncertainty has never overcome the differences between mainstream economic theory and its radical critique put forward by Keynesian fundamentalists. The mainstream view – formulated as part of the Rational Expectations Hypothesis (REH) revolution in macroeconomics in the 1980s (see Lucas 1980; Begg 1982) – has been, and mostly still is, that Keynes lacked the mathematical tools to construct a coherent theory of how expectations are formed and revised. Keynesian fundamentalists have rejected this view, arguing that Keynes did not refer to probabilistic risk because he considered it impossible, even in mathematical terms, to evaluate the future outcomes of current decisions, and that this viewpoint was integral to Keynes's notion of uncertainty (see Lawson 1985; Dow 1995).

However, among Keynesian scholars a viewpoint has also been developed which places emphasis on the possibility to derive insights from Keynes's theory of probability which, in principle, are suitable for a formal representation. Indeed, while Lawson's (1985) viewpoint refers to a notion of uncertainty that evokes numerically indeterminate or non-comparable probability relations, Gerrard's (1994) analysis of what he calls the Keynesian Uncertainty Hypothesis (KUH) suggests a line of research focused on a generalised notion of rational degree of belief retaining analytical tractability (see also Runde 1994a and Carabelli 2002).

Following Gerrard's comparison between REH and KUH, the main theme of this article is to point to an interpretation of Keynes's theory of long-term expectations that is intimately related to his notion of uncertainty as different from mathematical risk, but investigates the formal properties of Keynesian decision-making under uncertainty. Gerrard (1994, p. 335) introduces a functional form to represent KUH, according to which “the propensity to act on an expectation depends on the credence of the expectation where credence reflects the agent's assessment of the adequacy of the available evidence”, that is, “an evaluation of the vagueness of the knowledge on which the expectation is based”. In fact, Keynes distinguished short-term expectations, based on which a certainty equivalent modelling strategy appears to be sensible because a probability distribution among possible outcomes is usually possessed by entrepreneurs, and long-term expectations, on which the previous modelling strategy loses ground. Keynes (1936, p. 148) argued that the need to form long-term expectations typically coincides with a situation in which an individual's decision “does not solely depend ... on the

most probable forecast we can make ... [but also] on the *confidence* with which we make this forecast – on how highly we rate the likelihood of our best forecast turning out quite wrong”. Moreover, changes in the available evidence may affect behaviour by operating on both expectations and confidence.

Under KUH, Gerrard summarises the idea of Keynesian long-term expectations for uncertainty contexts through what he calls a “behavioural function” representing decisions  $x$ , taken at time  $t$ , as:  $x(t) = X[s^e(T), \rho(T)]$ . On this decision rule, choices are dependent on both probability and confidence.<sup>1</sup> Confidence distinguishes a behavioural function  $X$  for KUH from the mainstream behavioural function  $X$  for REH, represented as  $x(t) = X[s^e(T)]$ . It is confidence that “determines the completeness of the information set and the shape and fuzziness of the probability distribution” pointed out by Keynes and left unaddressed by the mainstream (Gerrard 1994, p. 335). However, Gerrard did not offer a specific representation of this functional form generalising decisions taken under the REH in a Keynesian direction, and his methodological viewpoint was not developed further even in similar later investigations (Fontana and Gerrard 2004; Basili and Zappia 2009).

This paper posits a direct approach to the representation of Keynes’s notion of “rational” behaviour in a scenario characterised by incomplete knowledge. The paper concentrates on financial markets where uncertainty is reflected in individuals’ prior distributions about the expected value of an asset, and suggests that a decision rule accounting for long-term expectations can be modelled by using a class of so-called  $\varepsilon$ -contaminated probability priors, where the parameter  $\varepsilon$  is suggestive of the quality of information about the relevant odds.<sup>2</sup> That is, instead of completely committing to the probability distribution elicited from market behaviour, an economic agent assumes that the elicited probability distribution may not be the “true” prior, its lack of reliability being revealed by the parameter  $\varepsilon$ , the lower the confidence  $\rho$  the higher the contamination  $\varepsilon$ . We shall argue that the  $\varepsilon$ -contamination approach can easily encompass Gerrard’s behavioural function and provide the needed functional form for Keynesian uncertainty.

---

<sup>1</sup> In Gerrard’s representation,  $s$  is the state of the world over future times,  $s^e(T)$  is the rational expectation of future states of the world for horizon  $T$ , and  $\rho(T)$  is the confidence of  $s^e(T)$ . A distinction which Gerrard makes between credence and confidence will not be pursued in what follows.

<sup>2</sup> See, among others, Epstein and Wang (1994), Eichberger and Kelsey (1999), Kopylov (2009), Cerreia *et al.* (2013), Wolintzky (2016).

This approach to a representation of reasoning under uncertainty – and related criteria for decision-making – follows Ellsberg (1961) and his critique of Bayesian decision-making. As we shall see, after introducing his paradoxical urn examples, Ellsberg proposed a decision rule for agents taking into account what he called the ambiguity of the decision environment. Through the  $\varepsilon$ -contamination approach, it appears that Ellsberg’s representation of beliefs and related decision rules are well suited to represent the Keynesian behavioural function for long-term expectations, as Ellsberg himself hinted in his long unnoticed 1962 doctoral thesis (Ellsberg 2001). As a matter of fact, it can be argued that much of modern decision theory under uncertainty originated from Ellsberg’s suggestion to move beyond Bayesian decision theory and constitutes nothing but a development of Keynes’s model of “rational” behaviour.

The paper is organised as follows. Section 2 provides a brief introduction to Keynes’s theory of long-term expectations, with specific regard to financial markets. Section 3 illustrates the mainstream view of the relationship between asset prices and probabilities, both in portfolio theory and in general equilibrium models. Recent attempts to examine this relationship when markets are incomplete due to uncertainty are presented. Section 4 discusses Ellsberg’s criticism of the Bayesian approach, while the link between his approach and the  $\varepsilon$ -contamination approach are illustrated in Section 5. Section 6 proposes a unified interpretation of Ellsberg’s and Keynes’s viewpoints. Section 7 offers a few concluding remarks.

## 2. Keynes on the state of long-term expectation

As is well known, in Chapter 12 of his *General Theory* (henceforth *GT*), Keynes claims that “the state of long-term expectation, upon which our decisions are based, does not solely depend ... on the most probable forecast we can make ... [but also] on the *confidence* with which we make this forecast – on how highly we rate the likelihood of our best forecast turning out quite wrong”. He then exemplifies his viewpoint as follows: “If we expect large changes but are very uncertain as to what precise form these changes will take, then our confidence will be weak” (Keynes 1936, p. 148).

It is worth noting from the outset that when Keynes stresses that uncertainty is a crucial aspect that influences investors’ forecasts he has in mind an analytical issue that “economists have not analysed ... carefully”. He clarifies that “by ‘very uncertain’ I do not mean the same thing as ‘very improbable’,” and refers to the *Treatise on Probability* (henceforth *TP*) – specifically its Chapter 6 on the weight of argument – for his elaborate argument (Keynes 1936,

p. 148 n.). Therefore, Keynes not only establishes a direct relationship between *TP* and *GT*, but also sets out his notion of uncertainty as the one that emerges from *TP* (Runde 1994a). Every consideration of decision-making under uncertainty in the Keynesian setting must then start from the probabilistic approach put forward in *TP* (Basili and Zappia 2009).

Keynes remarks that while in forming expectations in uncertain contexts “it would be foolish … to attach great weight to matters which are very uncertain”, in actual markets “practical men” need to act. And indeed, contrary to economists, they usually “pay the closest and most anxious attention” to the state of confidence, thus concentrating on the reliability of knowledge about what is usually a vague and distant future. The issue is one of “business psychology” requiring a “different level of abstraction” from most of *GT* (Keynes 1936, p. 149).

Crucially, what is and how does the state of confidence change? Keynes (1936, p. 149) observes that “our knowledge of the factors which will govern the yield of an investment some years hence is very slight and often negligible”, even in a very short period. Different from heroic times – times in which the investment “was partly a lottery, though with the ultimate result largely governed by whether the abilities and character of the managers were above or below the average” (Keynes 1936, p. 150) – when the separation between ownership and management prevails, and organised investment markets develop, “certain classes of investment are governed by the average expectation of those who deal on the Stock Exchange as revealed in the price of shares, rather than by the genuine expectations of the professional entrepreneur” (Keynes 1936, p. 151). There then emerges the issue of how financial markets work and in particular of the conventional attitude of market participants.

Keynes condenses the process that induces the formation of convention in the famous metaphor of financial markets as a newspaper beauty contest in which “the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one’s judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest” (Keynes 1936, p. 156). Since such a setting applies to markets as well, an investor does not have to anticipate what the fundamental value of a firm will be in the future, but rather should estimate other investors’ value, with investors – like competitors

in a beauty contest – devoting their “intelligence to anticipating what average opinion expects the average opinion to be” (Keynes 1936, p. 156).

For Keynes, the individual’s estimated value is different from “the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities” as usually assumed: rather, it depends on how the individual decodes it. In fact, when making an investment decision, the assumption is that “the existing market valuation, however arrived at, is uniquely *correct* in relation to our existing knowledge of the facts which will influence the yield of the investment, and that it will only change in proportion to changes in this knowledge; though, philosophically speaking it cannot be uniquely correct, since our existing knowledge does not provide a sufficient basis for a calculated mathematical expectation. In point of fact, all sorts of considerations enter into the market valuation which are in no way relevant to the prospective yield” (Keynes 1936, p. 152).<sup>3</sup>

Keynes believes that the increase in number of ordinary investors who do not have the technical competence to evaluate future returns in a correct way does not simply induce valuations that are possibly unrelated to fundamental values, but also produces erratic changes in the prices of assets resulting in strong fluctuations of the conventional evaluation of an asset. These ordinary investors – the crowd, to be distinguished from professional investors – assume that existing market evaluation, or the existing state of affairs, will continue indefinitely, but their lack of knowledge about the real evaluation of investments makes them prone to sudden changes in attitude, which produce large variability in asset prices. In fact, when the assumption of stable market evaluation appears “less plausible than usual”, it may well occur that “the market will be subject to waves of optimistic and pessimistic sentiment, which are unreasoning and yet in a sense legitimate where no solid basis exists for reasonable calculation” (Keynes 1936, p. 154). That is, variability will turn in deep fluctuations of asset prices due to the “flimsy foundations” on which valuations are based. In such a situation, while he should be concerned with making “superior long-term forecasts of the probable yield of an investment over its whole life”, the professional investor is instead largely concerned with “foreseeing changes in the

---

<sup>3</sup> The importance of this point is restated in the summary of *GT* he published in the *Quarterly Journal of Economics* in 1937. Here Keynes places emphasis on the fact “knowing that our own individual judgment is worthless, we endeavour to fall back on the judgment of the rest of the world which is perhaps better informed. That is, we endeavour to conform with the behaviour of the majority or the average. The psychology of a society of individuals each of whom is endeavouring to copy the others leads to what we may strictly term a conventional judgment” (Keynes 1937, p. 214).

conventional basis of valuation a short time ahead of the general public" (Keynes 1936, p. 154).<sup>4</sup>

In summary, Keynes considers the stock exchange as ruled by professional investors and speculators who are forced to anticipate the mass psychology of the market, that is, to consider, first, how much valuations in markets are based on conventional attitudes and, second, how this conventional basis of valuation may change. In this instance, the behaviour of professional investors and speculators is the result of two different components: "the average expectation of those who deal on the Stock Exchange as revealed in the price of shares" (Keynes 1936, p. 151) and the competence to anticipate "what average opinion expects the average opinion to be" (Keynes 1936, p. 156). In order to rationalise all this, from a theoretical viewpoint, the issue is how to incorporate these two different components in an expectation function, as suggested by Gerrard.

### **3. Foundations of portfolio selection and asset pricing**

Before coming to a possible representation of a behavioural function for markets in a Keynesian setting, it is important to recall why Keynes's intuition and prescription have not been pursued by standard portfolio and asset market theory. Indeed, the issue can be put as follows: why was expected utility maximisation – with its clearly un-Keynesian implication about a weighted mean of expected return and a unique and fully reliable probability distribution on future returns – considered a plausible representation of investors' rational behaviour? A brief inspection of the rationale under the initial application of rational decision-making to portfolio selection may help in clarifying this issue.

In his seminal research, Harry Markowitz, the founder of modern portfolio theory, suggested representing the process of portfolio selection as a two-stage procedure, with a first stage that "starts with observation and experience and ends with beliefs about the future performances of available securities", and a second one that "starts with relevant beliefs about future performances and ends with the choice of portfolio". Concentrating on the second stage,

---

<sup>4</sup> Among professional investors, Keynes (1936, p. 158) distinguishes between speculators and entrepreneurs. 'Speculation' refers to the activity of "forecasting the psychology of the market", while 'enterprise' refers to the activity of "forecasting the prospective yield of assets". We shall come back to this distinction in section 6, but note now that Keynes's aim was to theorise about both activities (Zappia 2016).

he advised as a “rule concerning choice of portfolio” one such that “the investor does (or should) maximise the discounted (or capitalised) value of future returns” (Markowitz 1952, p. 77). Indeed, “in so far it is applicable, the expected utility analysis provides a new viewpoint on the choice of criteria for the selection of portfolios” (Markowitz 1959, p. 210). Markowitz then worked in the tradition of von Neumann and Morgenstern’s expected utility, as if a rational agent may consider his/her beliefs about the performance of future as kinds of objective probabilities about risky prospects, irrespective of whether information was imprecise or vague as in Keynes. In fact, Markowitz remarked, this is an adherence to expected utility based on the new axiomatic foundation provided by Savage (1954), enlarging the domain of expected utility to subjective probabilities about uncertain events.<sup>5</sup>

As is well-known, concentrating on preferences consistent with the fundamental axioms of rational choice theory presented by Savage, Markowitz examined a problem with a set of assets represented by the expected value of their future dividend stream, with probability distributions representable through mean and variance (first and second central moments). He showed that, if an individual has a concave expected utility function, portfolio diversification is always preferred to each of them. According to Markowitz (1959, p. 218), “diversification between two equally good portfolios cannot produce a worse portfolio and generally will produce a better one. A concave utility function, therefore, is a conservative one, consistent with the purchase of insurance and the diversification of portfolio”.<sup>6</sup>

The implication of this way of modelling portfolio decisions, as a strict application of subjective expected utility maximisation, was then put forward to a theory of asset markets. In his Capital Asset Pricing Model, Sharpe (1964) assumed that investors engage in expected utility maximising behaviour, following the prescriptions of Markowitz’s portfolio theory, and characterised their trading activity through general equilibrium conditions. Sharpe argued that

---

<sup>5</sup> Markowitz, a doctoral student at the University of Chicago in the early 1950s under Leonard Savage and Jacob Marschak, remarked: “In recent years a justification has been presented that goes beyond the apparent plausibility of the expected utility maxim. The new axiomatic approach begins with basic principles, which seem beyond denial, then demonstrates that expected utility maxim follows from these principles. The axiomatic approach has revised interest and gained a large number of adherents for the two hundred years old expected utility maxim” (Markowitz 1959, p. 209).

<sup>6</sup> It was in a similar vein that Keynes’s insights were reduced to the analysis of liquidity preference as attitude toward risk by Tobin (1958) and Hicks (1962), who also assumed that a risk-averse investor maximises expected utility of an asset, described by a unique objective, and fully reliable probability distribution with respect to the mean and variance.

his asset pricing model “incorporates assumptions about investors’ utility functions, and assumed a market with a large number of participants, each of whom has access to the same set of information”, proposing a general equilibrium model under risk (Sharpe 1990, p. 314). Crucially, all investors are in agreement concerning expected return and covariance of assets: in fact, CAPM introduces the pivotal assumption in Markowitz’s theory that, as far as expectations are concerned, there is “*complete agreement*: given market clearing asset prices at  $t-1$ , investors agree on the joint distribution of the asset returns from  $t-1$  to  $t'$ ” (Fama and French 2004, p. 26).

From this point of view, asset prices reflect the discounted expected value of payoffs from the asset, adjusted for risk, with the discounted value depending on a unique additive probability distribution over future dividends. Since probabilities are revealed by prices, there can be no disagreement about individuals’ degrees of belief so that a representative agent model can be thought as offering a consistent representation of asset markets (Lucas 1978). Moreover, it turns out that the price of an asset at time  $t$  is the discounted value of the assets’ future payoffs, conditional on the information available to the representative investor at  $t$ . That is, asset prices satisfy the martingale property and its implication that the representative investor’s expectation at  $t$  of his/her expectation at  $t+1$  is equal to his/her expectation at  $t$  of future payoffs, known as the law of iterated expectations. Traditionally, therefore, a key feature of financial literature has been that high-order expectations – market participants’ beliefs about other market participants’ beliefs – do not have a role in the determination of assets prices: what average opinion expects the average opinion to be is redundant.<sup>7</sup>

From a general equilibrium perspective à la Arrow-Debreu, this result entails a particular asset pricing rule. Under this pricing rule, each asset’s value is calculated as the weighted average of the number of fundamental assets coinciding with so-called Arrow securities, originating a perfect replicating portfolio. In fact, under completeness and in a frictionless market, portfolios of Arrow securities can replicate any pattern of revenues across possible future states of the world. A perfect replicating portfolio is a linear combination of marketed security payoffs that is equal to the asset payoff (hedging strategy). Under the no

---

<sup>7</sup> The law of iterated expectations and a martingale sequence of random variables are fundamental properties of many theorems in applied statistics. Given two random variables  $x$  and  $z$ , the basic statement of the law of iterated expectation is as follows:  $E(x) = E(E(x|z))$ . Given a probability space  $(\Omega, \Sigma, p)$ , where  $\Omega$  is the set of the states of the world,  $\Sigma$  is a  $\sigma$ -algebra and  $p$  is a probability measure, a time process is a martingale if  $x_i = E\{x_{i+1}|\Sigma_i\}$  almost certainly, for  $i = 1, 2, \dots, n-1$ .

arbitrage assumption – a necessary condition for equilibrium, since if arbitrage opportunity exists exchange could be infinite – two assets with identical payoffs have the same price (this is also known as the “law of one price” with respect to expected returns). Such a price function is both unique and linear, because an asset price is defined as its formation cost, that is, by the linear combination of the securities in the replicating portfolio (Dybvig and Ross 1982). Any asset price can then be replicated by trading securities and the linear pricing rule is equivalent to the mathematical expectation of asset returns with respect to a unique risk-neutral probability distribution (Fundamental Theorem of Asset Pricing); that is, the unique state contingent price is a risk-neutral probability distribution or martingale. Crucially “the shadow prices at the optimum – the intertemporal marginal rates of substitution – are the same for all agents” (Jouini and Kallal 2001, p. 347).

But if there are reasons for these necessarily restrictive conditions not to hold, the whole set-up fails to deliver the unique pricing rule. As noted by Allen *et al.* (2006), if there is differential information among investors that justifies a role for the average expectations about payoffs, “the folding back of future outcomes to the present cannot easily be achieved” since, in general, average expectations fail to satisfy the law of iterated expectations. Following on Keynesian insights, “it is *not* the case that the average expectation today of the average expectation tomorrow of future payoffs is equal to the average expectation of future payoffs” (Allen *et al.* 2006, p. 720). For instance, if an investor thinks that prices convey not only private information but also public information, he/she may think that the public signal is a better predictor of average opinion than the private one, so that asset prices will overweight public information (Morris and Shin 2002).

This result is theoretically analogous to that observed when asset markets are incomplete or some sort of friction – such as transaction costs, taxes or the bid-ask spread – affects markets. In incomplete or friction markets, there is more than one risk-neutral probability distribution, i.e., more than one martingale, and, as a result, there exist several hedging portfolios with different costs: there exist super-martingale measures, sub-martingale measures, absolutely continuous martingale measures, and so on (Jouini and Kallal 2001, Araujo *et al.* 2012). As a consequence, a contingent claim does not have a unique price but an interval of prices – identifying arbitrage bounds – derived from all the risk-neutral probabilities (martingales) that agree to the no-arbitrage condition (multiple linear prices). Multiple martingales measures then arise when markets are not perfect or agents have incomplete knowledge, entailing a multiple pricing rule that is compatible with the no-arbitrage assumption. In such a condition, the replication strategy could not be the cheapest way to hedge

(super-replication strategy or a strategy that hedges an asset with a larger cost than its price), typically inducing inertia.

Multiple pricing rules for contingent claims can then be interpreted as “the different implicit shadow prices – the intertemporal marginal rates of substitution – for different potential agents” (Jouini and Kallal 2001, p. 347). In markets with frictions, the pricing rule is generally non-linear: for instance, with bid-ask spread an asset has different selling and buying prices, then its equilibrium price cannot be obtained by a unique, linear and positive pricing rule. Jouini and Kallal (1995, 1999) show that for a large class of market imperfections the pricing rule is sub-linear, i.e., sub-additive or concave,<sup>8</sup> and the pricing rule is the maximum in the set of underlying pricing rules, i.e., risk-neutral probability distributions or martingales that hold a no-arbitrage condition.<sup>9</sup>

Apart from the motivations that may justify these results (for this, see the following sections), it is worth noting that a sub-additive pricing rule can be considered as a precautionary rule according to which it is generally less expensive to purchase a portfolio of securities than to purchase each security separately (Araujo *et al.* 2018). Indeed, a sub-additive pricing rule equals the Choquet integral of returns with respect to a concave measure (as proved axiomatically in Chateauneuf *et al.* 1996; Castagnoli *et al.* 2002). These results stem from the literature representing probability priors through non-additive probabilities (capacities) that achieve the maximisation of expected utility through a Choquet (1954) integral of outcomes with respect to the non-additive probability, as in Choquet Expected Utility introduced by Schmeidler (1989).<sup>10</sup>

---

<sup>8</sup> Given two assets  $x$  and  $z$ , a pricing rule  $H$  is sublinear (sub-additive) if  $H(x + y) \leq H(x) + H(y)$  and  $H(\lambda x) = \lambda H(x)$ , for all  $\lambda \geq 0$ .

<sup>9</sup> It is well known (Hahn-Banach Theorem) that any sub-linear functional is the upper envelope of a set of linear functionals, that is the supremum among all the expected value in the set. The set of underlying linear pricing rules of a sub-linear pricing rule is the set of martingales measures of all traded assets’ normalised price processes when: markets are incomplete, there exists a bid-ask spread, there are different borrowing and lending rates and contingent claims to future consumption are efficient in a multi-period economy under uncertainty (Jouini and Kallal 2001).

<sup>10</sup> Formally, if  $\Omega = \{w_1, \dots, w_n\}$  is a non-empty set of states of the world and  $\Sigma = 2^\Omega$  the set of all events, a capacity (i.e. a non-necessarily additive probability measure) is a function  $v: \Sigma \rightarrow \mathbb{R}_+$  such that: (i)  $v(\emptyset) = 0$ , (ii)  $v(\Omega) = 1$ , and (iii) for all events  $A, B \in \Sigma$  such that  $B \subseteq A$ ,  $v(A) \geq v(B)$ . A capacity is said to be concave (subadditive) if: (iv)  $v(A \cup B) \leq v(A) + v(B) - v(A \cap B)$ . The Choquet Integral of an act  $x$  with respect to a capacity  $v$  is:

$$\int x d\nu = \int_0^\infty v(\{\omega \in \Omega \mid x(\omega) \geq t\}) dt + \int_{-\infty}^0 [v(\{\omega \in \Omega \mid x(\omega) \geq t\}) - 1] dt.$$

Multiple, or non-linear, pricing rules allow for the introduction of Keynesian uncertainty and agents' perceived state of confidence with respect to the “true” price of an asset in a standard model. In fact, Choquet Expected Utility makes it possible to represent different attitudes with respect to uncertainty: a sub-additive rule representing an uncertainty-averse agent, and a super-additive pricing rule representing an uncertainty-seeking agent.<sup>11</sup> As we shall see in the following sections, this way of dealing with incomplete or imperfect markets is substantially coherent with Keynes’s view of financial markets, especially in view of the suggested representation of long-term expectations by means of  $\varepsilon$ -contaminated decision models.

#### 4. Ellsberg’s rules

We have seen in the previous section that the mainstream view of financial markets and the determination of assets prices has its theoretical foundations in Savage’s axiomatisation of decision-making under uncertainty. From this viewpoint, even REH is nothing but a claim that degrees of belief meet some standard of correctness, since it supposes that individuals know the “true” probability distribution of future states of the markets, so that when maximising the expected utility of their choices they use the “correct” probability distribution, and not simply their personal, subjective one as in Savage’s framework. We also saw previously that Keynes would have objected to this representation of the financial markets, in terms of both how to apply probability calculus under uncertainty and how to consider what it is rational to do in view of the conventional attitude of financial investors.

When looking for alternative representations, it is necessary to focus on the fundamental criticism of Savage’s axiomatisation as regards its ability to actually take into account uncertainty in decision-making. Put forward in the early 1960s and usually referred to as the Ellsberg Paradox, this criticism inspired much of the non-additive probability literature introduced earlier, but is worth reviewing here for its original contribution. In his seminal paper,

---

<sup>11</sup> Cerreia-Vioglio *et al.* (2015, p. 731) have recently shown that “a pricing rule is sub-linear and Choquet if and only if the non-additive probability that represents it is concave. In this case, the set of consistent price systems coincides with the core of this non-additive probability”. The *core* of a concave capacity  $v$  is the non-empty closed and convex set of probability distributions that minorises it, i.e.,  $\text{core } (v) = \{\pi \in \Pi \mid \pi(A) \leq v(A), \forall A \in \Sigma\}$  and  $\pi(\Omega) = v(\Omega)$ .

while introducing his famous urn examples against Savage's Bayesian approach, Ellsberg (1961) objected to the normative plausibility of Savage's axioms for decisions under uncertainty. His urn examples are presented as "a class of choice-situations in which many otherwise reasonable people neither wish nor tend to conform to the Savage postulates, nor to the other axiom sets that have been devised". From a methodological viewpoint, the aim of Ellsberg (1961, p. 656) was to identify situations where a decision-maker feels that he/she faces 'uncertainties' that are not representable as 'risks', therefore objecting to the mainstream view that, behaviourally, risk and uncertainty are undistinguishable. Examining gambling choices related to urns containing balls of different colours, Ellsberg argued that many subjects violate the so called sure-thing principle (Savage's P2 axiom), but do not regret making their choices even upon reflection. The Ellsberg Paradox thus concerns decision-makers who deliberately violate the maximisation of Subjective Expected Utility (SEU).

Ellsberg's idea is that, being retained after 'thorough deliberation', the observed choices represent violations of normative value, a step forward from what Allais observed (1953) and Savage (1954, pp.103-104) rejected as 'mistakes' of descriptive value only. As a result, violations of this kind call for a criterion for decision-making alternative to the maximisation of SEU, to be applied when the decisional context has characteristics usually related to uncertainty. Specifically, Ellsberg (1961, p. 657) asserts that violators' behaviour does not depend on the relative desirability (utility) or likelihood (probability) of consequences, but rather stems from "a third dimension of the problem of choice: the nature of one's information concerning the relative likelihood of events", that he identifies as ambiguity, defined as a "quality depending on the amount, type, reliability and 'unanimity' of information, and giving rise to one's degree of 'confidence' in an estimate of relative likelihoods".

Ambiguity, then, is intended to encompass situations in which a decision-maker does not commit to a single, sharp probability prior, but can be thought of as if he/she has more than one prior probability distribution  $y \in Y$  in mind, none of which is considered fully reliable, even if, in the set of all possible distributions  $Y$ , a subset  $Y^0$  of them can be considered more reasonable than others, and one of them possibly the 'best guess'  $y^0 \in Y^0$ . In these instances, Ellsberg (1961, p. 667) suggested a decision criterion that is related to Wald's minimax, where  $Y$  is restricted to  $Y^0$ , but allows for a variable degree of confidence in the 'best guess' probability estimate  $y^0$ . He considered the behaviour of violators of the maximisation of SEU as that of an individual from whom it is impossible to infer a prior probability on payoff-relevant events, since "in effect, he 'distorts' his best estimates of likelihood, in the direction

of increased emphasis on the less favourable outcomes and to a degree depending on  $\rho$ , his confidence in his best estimate”.

Using Ellsberg’s terminology, in an ambiguous decision context, each action (act)  $x$ , with the associate vector of von Neumann Morgenstern’s utility payoffs (consequences)  $\mathbf{x}$ , is evaluated by means of the functional  $I(x) = [\rho y^0 + (1 - \rho)y_x^{min}](\mathbf{x})$ . That is, a convex combination with respect to the confidence parameter  $\rho \in [0,1]$  of the expected payoff corresponding to the best estimated probability distribution  $y^0$  and the minimum expected payoff, where  $y_x^{min}$  is the probability vector corresponding to  $\min_x$ , the minimum payoff associated to  $x$ , with respect to the set of reasonable probability distributions  $Y^0$ . Following Hodges and Lehmann (1952), Ellsberg called the decision rule that suggests choosing the action  $x$  with the highest  $I(x)$  the Restricted Bayes Criterion, later renamed Restricted Bayes/minimax. It is immediate to see that Ellsberg’s decision rule coincides with the maximization of Savage’s subjective expected utility  $y^0(\mathbf{x})$ , when  $\rho = 1$ , of which Gerrard’s “behavioural function” is but a special case when all the individuals agree on the rational expectation of future consequences, that is, the “true” probability distribution.<sup>12</sup>

It is unfortunate that Ellsberg’s further elaboration on this topic presented in his doctoral dissertation remained unpublished for long. For instance, in the thesis, Ellsberg (2001, p. 192) clarifies what he means by confidence through the illuminating example of a “decision-maker who relies upon a panel of experts to guide his official opinions, and who finds in a particular case that each consultant produces a different, definitive probability distribution”. In such a case, he finds it convenient ‘to think of some of the members of  $Y^0$  in the concrete form of probability distributions each written down on a separate piece of paper with the name of the forecaster attached. For simplicity, we may imagine a decision-maker who is compelled, in some sense, to base his own opinions, and hence his actions, upon this set of conflicting forecasts ... In the end, we can imagine this decision-maker evolving a particular distribution  $y^0$  over the relevant events, representing his own ‘best guess’ opinions on all these questions that may influence, directly or remotely, his judgments of the relative probabilities of those events ... [H]is occasional (or frequent) failure to *act upon*  $y^0$  exclusively reflects another sort of judgment, concerning the reliability, credibility or adequacy of his information, experience, advice, intuition taken as a whole, not about the relative support it may give to one hypothesis as opposed to another, but about its ability to lend support to any hypothesis – any set of

---

<sup>12</sup> Alternatively, complete ignorance, represented by  $\rho = 0$ , implies that Ellsberg’s decision rule coincides with Wald’s maxmin, that is, with taking the maximum of  $y_x^{min}(\mathbf{x})$ .

definitive opinions – at all”. Moreover, Ellsberg’s explanation for introducing his decision rule is suggestive of an obvious connection with Keynes’s analysis, one that he did not point out in the original 1961 paper and is thus usually disregarded. Ellsberg (2001, p. 11) remarks that Keynes “had related the possible difficulty of comparing the probabilities … to the sort of *conflict of evidence* we have included in the notion of ‘ambiguity’,” and that “low ‘weight’ [of argument] he relates primarily to scantiness of evidence of any sort”. Ellsberg suggests that he is moving forward in his footsteps: “Many writers, including Frank Knight and Lord Keynes, have insisted upon the feasibility and relevance of this sort of judgment, without indicating precisely how it might affect decision-making; we shall consider now a meaningful role” (2001, p. 193).

The significance of Ellsberg’s proposal for a representation of Keynesian uncertainty is enlightened by some recent analysis endorsing his intuition as part of the epsilon-contamination approach. In fact, the Ellsberg Paradox can be explained as the choices made by an individual whose degrees of belief are represented by the  $\varepsilon$ -contamination of the most reasonable probability distribution in the set  $Y^0$  of all possible probability distributions defined for the Ellsberg urns (Nishimura and Ozaki 2006; Chateauneuf *et al.* 2007). Such an individual would violate Savage’s axioms simply because he/she is sorting gambling choices by a mixture of the worst belief in the set  $Y^0$  with the probability measure representing his/her best guess  $y^0$ . As suggested in the recent literature on decision-making under ambiguity, it is natural to interpret this probability as the decision-maker’s ex-ante probabilistic belief (Kopylov 2009).

## 5. The Epsilon-Contamination approach

We have noted above that, after introducing a paradoxical result to criticise the Bayesian mainstream, Ellsberg addressed the issue of how to rationalise decisions under ambiguity. He proposed a decision rule that utilises the available information but at the same time provides a ‘safeguard’ in case this information is not correct and suggested a functional allowing for confidence in the available information to play a role. Indeed, Ellsberg’s analysis (2001) provides a Bayesian-like justification for the choices of the unrepentant violator of Savage’s axioms to be included in the rational realm. In the probabilistic framework he put forward, there is, on the one hand, the strictly-Bayesian decision-maker – who may keep in mind a whole set of “reasonably acceptable” probability distributions before acting, but eventually settles upon a single distribution – and, on the other hand, the generalised-Bayesian – who retains all

those probability distributions that do not definitely contradict his/her “vague” opinion, especially when relevant information is ambiguous. Ellsberg (2001) also noted that the formal structure guaranteeing the consistency of such probabilistic beliefs had been provided by authors such as Koopman (1940), Good (1950) and Smith (1961) who, while adhering to a Bayesian approach, did not impose a unique probability prior on individuals.

It is therefore appropriate to look at the statistical literature placing emphasis on partially ordered probability and non-unique priors for further elaboration of Ellsberg’s viewpoint. Indeed, as part of an empirical Bayes analysis of individual beliefs – better known as robust Bayesian analysis – interesting results under the so-called  $\varepsilon$ -contamination approach have been provided. The robust Bayesian viewpoint affirms that one of the main justifications for using Bayesian analysis is the idea that prior distributions can never be quantified or elicited exactly (i.e. without error), especially in a finite amount of time (Berger 1984). Crucially, while this assumption is typically sufficient for a rejection of Bayesian analysis from a frequentist-oriented statistical viewpoint, for a Bayesian statistician it “precludes the obvious Bayesian solution of writing down a single prior distribution and doing a Bayesian analysis”, but it does not imply a demise of the whole approach. As summarised by Berger (1984, p. 65) the ‘robust Bayesian viewpoint’ is that “one should strive for Bayesian behaviour which is satisfactory for all prior distributions which remain plausible after the prior elicitation process has been terminated”.

From this viewpoint, an attractive method of modelling uncertainty in the prior distribution is the use of an  $\varepsilon$ -contamination functional, that is, a specific method to identify among all possible classes of priors modelling prior uncertainty those  $y$  which are encompassed by the following form:  $y = (1 - \varepsilon)y^0 + \varepsilon q$ , where  $y^0$  is the elicited prior,  $q$  is a contamination or perturbation of  $y^0$ , and  $\varepsilon \in [0,1]$  reflects the amount of error in  $y^0$  that is considered possible.<sup>13</sup> This view suggests that prior uncertainty can be modelled by means of a mixture of those distributions that are in the neighbourhood of the single Bayesian, best guess prior. Then the  $\varepsilon$ -contamination emerges as a robust Bayesian method for quantifying, in terms of the class of possible distributions, how partial and incomplete the subjective information encompassed in a single prior distribution is. It offers an analytic way for characterising, among other

---

<sup>13</sup> This makes it possible to identify a class  $Y^0$  of prior distributions such that  $y^0 \in Y^0$ . Since it is in principle arguable that the higher the error admitted the fuzzier the probabilities to work with, robust Bayesian analysis has concentrated on characterisations and results that are robust over the set of distributions involved (Berger 1984).

technical aspects, the upper and lower limits to the prior – the elements which are ‘reasonable’ to consider with respect to the best guess –and, most of all, an easily elicitable prior information (Berger and Berliner 1984).

From our analysis so far, the fact that partial prior knowledge – or worry about prior misspecification – is modelled with a class of  $Y^0$  of priors makes the class of probabilities studied within the robust Bayesian approach a natural candidate for representing Keynes’s long-term expectation approach. Moreover, the analogy with Ellsberg’s analysis is apparent: following on from Ellsberg’s suggestion that one can imagine that the prior probability used by deliberate violators of Savage’s axiom can be seen as a ‘distortion’ of their best estimate, the  $\varepsilon$ -contamination procedure suggests how to model the distorted probability, in view of the perceived ambiguity of the decision context. This is of course true once it is noted that Ellsberg’s confidence  $\rho$  equals  $1 - \varepsilon$ , the higher the confidence the lower the error.<sup>14</sup>

## 6. Keynes and the Epsilon-Contamination approach

As we have seen in Section 2, Keynes considers investors’ long-term expectations about an asset as the combination of the asset price and an estimate of other investors’ asset value, and this asset prices evaluation depends on the ‘confidence’ with which it is made. Although this intuition did not find room in mainstream asset market theory, a series of recent developments can be considered as admitting its relevance and trying to address it.

Keynes believes that in financial markets, two different attitudes – possibly corresponding to different classes of investors – coexist: speculators who are concerned “not with making superior long-term forecasts of the probable yield of an investment over its whole life, but with foreseeing changes in the conventional basis of valuation a short time ahead of the general public” and skilled individuals who act on the base of “the best genuine long-term

---

<sup>14</sup> A representation with  $\varepsilon$ -contamination complies with Choquet Expected Utility and a capacity  $v$  corresponds to the  $\varepsilon$ -contamination of a probability  $\pi \in \Pi$ , given a set  $\Omega$  of states of the world and the set  $\Sigma$  of all subsets, if for all  $A \subseteq \Sigma$ :  $v(A) = (1 - \varepsilon)\pi(A)$  if  $A \neq \Sigma$ , and  $v(A) = 1$  if  $A = \Sigma$ ).

Nishimura and Ozaki (2006), Chateauneuf *et al.* (2007), and Kopylov (2009) give different sets of axioms by which investors’ preferences with regard to uncertain acts (prospects or lotteries) can be represented by Choquet Expected Utility with the  $\varepsilon$ -contamination of confidence.

expectations” (Keynes 1936, pp. 153-154). From this perspective, long-term expectations can be considered as the result of the combination, through the degree of confidence, of the probability distribution that represents the expectation of future payoffs,  $y^0$ , and the probability distribution that characterises an investor’s most reliable evaluation of the asset,  $y^{mr}$ . That is, an investor’s long-term expectation can be thought of as the parametric combination, with respect to the error that is deemed possible, of the prior representing the actual convention – which we have seen earlier can be derived through the sub-additive pricing rule with respect to a capacity – and the probability distribution that characterises an alternative possible common evaluation of the asset.<sup>15</sup>

If, as in Ellsberg, the parameter  $\rho \in [0,1]$  expresses the individual confidence in market evaluation – with  $1 - \rho = \varepsilon$  representing the  $\varepsilon$ -contamination of confidence – each investor’s long-term expectation about an asset  $x$  can formally be denoted by  $I'(x) = [\rho y^0 + (1 - \rho)y^{mr}](x)$ , where the probability distribution  $y^0$  is the one elicited from the pricing rule that assesses the expectation of future payoffs of the asset, and  $y^{mr}$  is the probability distribution that characterises an investor’s alternative evaluation of the asset, in the core of the sub-additive capacity (that is, as recalled in footnote 11, the set of probability distributions associated to the capacity). Specifically,  $y^{mr}$  represents the investor’s expectation of the average opinion, that is, what he/she considers the “true” consensus distribution. As a result, in analogy with what Ellsberg suggested in his critique, and was later elaborated through the  $\varepsilon$ -contamination approach, uncertainty can be represented by a specific probability distribution in the set of all consistent and efficient evaluations of the asset, and the  $\varepsilon$ -contamination interpretation of investors’ expectation makes it possible to describe both the indeterminacy/imprecision of a priori knowledge and the behavioural effect of its awareness.

It is straightforward then to consider Gerrard’s behavioural function as a particular case of the previous decision rule. Indeed, Gerrard’s functional  $x(t) \equiv X[s^e(T), \rho(T)]$  can be represented by the capacity  $v$  obtained by the contamination of  $y^0$ , such that for any prospect and  $\rho = (1 - \varepsilon) \in [0,1]$ , if  $s^e(T)$  is based on a fully credible probability prior, then  $\rho(T) = 1$  (i.e.,  $\varepsilon = 0$ ) and  $x(t) \equiv X[s^e(T)] = y^0(x)$ , as in the REH representation. On the contrary, if

---

<sup>15</sup> As suggested by Keynes (1936, p. 155): “This battle of wits to anticipate the basis of conventional valuation a few months hence, rather than the prospective yield of an investment over a long term of years, requires no gulls amongst the public to feed the maws of the professionals; it can be played by professionals amongst themselves. Nor is it necessary that anyone should keep his simple faith in the conventional basis of valuation having any genuine long-term validity”.

$s^e(T)$  is only partially credible, that is, if  $\varepsilon > 0$ , then  $x(t) \equiv X[s^e(T), \rho(T)] = [\rho y^0 + (1 - \rho)y^{mr}](x) = I'(x)$ .

In the main, Keynes's problem of assessing conventional judgment can be represented as the problem of how to aggregate the probability distributions of all investors, that is, how to find proper pooling methods apt both to discover other opinions – as represented by the asset value – and to judge how well informed they are. In the set of consistent probability distributions, i.e., the core of the sub-additive capacity, the investor could extract a distribution that in his/her opinion better represents the asset value, or what he/she considers the consensus distribution, that is his/her expectation of average opinion. In short, Keynes's issue is how to form and elicit a consensus distribution.<sup>16</sup>

In this setting, one can also take a closer look at how certain professional investors may entertain high order expectations. Differently from others – those investors who elicited their most reliable probability distribution among their own set of consistent probability distributions – certain professional investors have better competence and could be seen as trying to estimate the average long-term expectation of other investors in the market. These professional investors may be thought of as eliciting the conventional judgement by considering the common set (technically, the intersection) among all consistent probability distributions of agents. The estimate they try to envisage consists of considering the weighed mean of the set of investors' probability distributions, which can be represented as a Steiner point distribution  $y^{stp}$ . The Steiner point is the weighted probability distribution in the

---

<sup>16</sup> If investors' opinions are not all independent and equally likely, each investor has to cope with ambiguity and stochastically dependent evaluations. As a consequence, each investor could calibrate the aggregation of investors' opinions through his/her confidence or degree of belief by pooling methods based on Dempster's rule of combination or theory of evidence, combination rules based on possibility distributions and fuzzy measures, or aggregation based on multiple priors or capacity (DeMiguel *et al.* 2009; Huang 2010; Basili and Chateauneuf 2011, 2015; Basili and Pratelli 2015). This way of dealing with consensus distributions differs from the more conventional Bayesian axiomatic approach (DeGroot and Montera 1991).

intersection among all the investors' cores and is unique.<sup>17</sup> By so doing, a speculator's long-term expectation can be denoted by  $I''(x) = [\rho y^0 + (1 - \rho)y^{stp}](x)$ , with  $\rho \in [0,1]$ .<sup>18</sup>

Finally, once the consensus distribution or conventional judgment is estimated, professional investors have to anticipate its change when new information is revealed, that is, when a relevant non-null event occurs. In this framework, foreseeing changes in the conventional basis of valuation is equivalent to evaluating a conditional probability distribution, i.e., updating the Steiner point given an event with strict positive probability. If the Steiner point satisfies certain general properties, i.e., dynamic consistency to certainty (Araujo *et al.* 2016), a professional investor can directly update the Steiner point by the so-called Full Bayes Rule (Chateauneuf *et al.* 2011) and the updated probability distribution will be his/her estimated new conviction judgement.<sup>19</sup>

## 7. Concluding remarks

In this paper, we have argued that there is room for a formal representation of decision-making under uncertainty, and that the critical thread of modern decision theory which starts with Ellsberg's intuition to develop a non-additive representation of individual's expectation and asset market pricing is indeed elaborating this in Keynes's footsteps. This makes it possible to generalise Ellsberg's decision rule as a Keynesian decision rule catching the intuitive meaning of a behavioural function differing in a critical way from the mainstream one. The  $\varepsilon$ -contamination representation of long-term expectation gives a functional form to Keynes's contrast between actual beliefs and reasonable beliefs in *TP*, and makes apparent the continuity with the notion of probability used in Chapter 12 of *GT*. Our contention has been that the investor's expectation function, and the related decision rule we have introduced, can indeed

---

<sup>17</sup> The Steiner point can be considered a sort of weighed centre point (of a set of all probability distributions of agents) that takes into account the different market power of agents. In fact a dealer or a speculator moves more shares than a non-professional investor and for this reason we use the Steiner point instead of the barycentre of a set of probabilities (Basili *et al.* 2017). It should be also noted that the Steiner Point is equal to the Shapley Value, in a finite set.

<sup>18</sup> Basili *et al.* (2017) introduce an axiomatic model to represent this kind of professional investor behaviour in a Keynesian perspective.

<sup>19</sup> Under the Full Bayes Rule the updated Steiner point will represent the update core, that is the intersection of all investors' updated consistent probability distributions.

be considered as functional apt to give specific formal content to Gerrard's KUH form. In fact, along with the expectations about future outcomes weighted by the confidence attached to them, as Gerrard assumes, the  $\varepsilon$ -contamination representation we suggest also includes the expectations elicited through a consensus distribution, addressing Keynes's concern for the conventional attitude of individuals that market prices may convey.

The previous interpretation of financial assets pricing may appear overstretched in its attempt to put Keynes's idea of long-term expectation in the context of the current criticism of the mainstream view put forward in an array of studies emerging from the non-additive probability approach. Nonetheless, an accurate reading of *TP* and its relationship with *GT* suggests that this is not so, even with regard to the axiomatic bases of subjective probabilities. Specifically, it has been shown that, in order to try to give formal meaning to his contention that epistemic probability may not be numerical as requested by the frequentist interpretation he confronted, Keynes worked with interval-valued probability estimates (Runde 1994b, Basili and Zappia 2009), and that in Chapter 15 of his *TP* he used a Boolean framework providing a mathematical structure for a kind of non-additive decision theory approach (Brady and Arthmar 2012). And notwithstanding claims to the contrary, there is also considerable continuity between Keynes's earlier views on probability and parts of his later economic writings on the nature and effects of uncertainty (Zappia 2015). It is the formal content of this continuity that this paper has tried to place emphasis on.

## REFERENCES

Allais, M. (1953). Le Comportement de l'Homme Rationnel Devant le Risque: Critique des Postulats et Axioms de l'Ecole Americaine. *Econometrica*, 21: 503-46.

Allen, F., Morris, S., & Shin, H.S. (2006). Beauty Contests and Iterated Expectations in Asset Markets. *The Review of Financial Studies*, 19, 719-752.

Araujo, A., Chateauneuf, A., & Faro, J.H. (2012). Pricing rules and Arrow-Debreu ambiguous valuation. *Economic Theory*, 49: 1-35.

Araujo, A., Chateauneuf, A., & Faro, H. (2018). Financial Market Structures Revealed by Pricing Rules: Efficient Complete Markets Are Prevalent. *Journal of Economic Theory*, 173: 257-288.

Araujo, A., Chateauneuf, A., Faro, H., & Holanda, B. (2016). Updating Pricing Rules. Working Paper 2016.

Brady, M. & Arthmar, R. (2012). Keynes, Boole and the Interval Approach to Probability. *History of Economic Ideas*, 20: 65-84.

Basili, M. & Chateauneuf, A. (2011). Extreme Events and Entropy: A Multiple Quantile Utility Model. *International Journal of Approximate Reasoning* 2011; 52: 1095-1102.

Basili, M. & Chateauneuf, A. (2015). Aggregation of Coherent Experts' Opinion: A Tractable Extreme-Outcomes Consistent Rule. In G. Rogova & P. Scott, eds., *Fusion Methodologies in Crisis Management: Higher Level Fusion and Decision Making*, New York: Springer-Verlag.

Basili, M. & Pratelli, L. (2015). Aggregation of Not Necessarily Independent Opinions. *Structural Safety*, 52: 144-149.

Basili, M., A. Chateauneuf, & Scianna, G. (2017). Coherent Representation of Keynes's Long-Term Expectation. Working Paper, University of Siena.

Basili, M. & Zappia, C. (2009). Keynes's 'Non-numerical' Probabilities and Non-additive Measures. *Journal of Economic Psychology*, 30: 419–430.

Begg, D. (1982). *The Rational Expectations Revolution in Macroeconomics: Theories and Evidence*. Oxford: Philip Allan.

Berger, J. (1984) The Robust Bayesian Viewpoint. In J. Kadane, ed., *Robust Bayesian Analysis*. Amsterdam: North-Holland.

Berger, J. & M. Berliner. Robust Bayes and Empirical Bayes Analysis with  $\varepsilon$ -Contaminated Priors. *The Annals of Statistics*, 14: 461-486.

Carabelli, A. M. (1988). *On Keynes Method*. London: MacMillan.

Carabelli, A. M. (2002). Speculation and reasonableness: A Non-Bayesian theory of rationality. In S. Dow & J. Hillard (Eds.), *Keynes, Uncertainty and the Global Economy. Beyond Keynes* (Vol. II). Aldershot: Edward Elgar.

Castagnoli, E., Maccheroni, F. & Marinacci, M. (2002). Insurance Premia Consistent with the Market, Insurance. *Mathematics and Economics*, 31: 267-284.

Cerreia-Vioglio S., Maccheroni F., Marinacci M., & Montrucchio, L. 2013. Ambiguity and robust statistics. *Journal of Economic Theory*, 148: 974-1049.

Chateauneuf, A., Eichberger, J. & Grant, S. (2007). Choice Under Uncertainty with the Best and Worst in Mind: Neo-additive Capacities. *Journal of Economic Theory*, 137: 538-567.

Chateauneuf, A., Gajdos, T. & Jaffray, J.Y. (2011). Regular Updating. *Theory and Decision*, 71: 111-128.

Chateauneuf, A., Kast, R. & Lapied, A. (1996). Choquet Pricing for Financial Markets with Frictions. *Mathematical Finance*, 6, 323-330.

Choquet, G. (1954). Theorie des capacités. *Annales de l'Institut Fourier*, 5: 131-233.

DeGroot, M. & Montera, J. (1991). Optimal Linear Opinion Pools. *Management Science*, 37: 546-558.

DeMiguel, V., Garlappi, L. & Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies*, 22: 1915-1953.

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

Dybvig, P. & Ross, S. (1982). Portfolio efficient sets. *Econometrica*, 50: 1525-1546.

Eichberger, J. & Kelsey, D. (1999). E-capacities and the Ellsberg Paradox. *Theory and Decision*, 46: 107-140.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics*, 75: 643-69.

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

Epstein, L. & Wang, T. (1994). Intertemporal Asset Pricing and Knightian Uncertainty. *Econometrica*, 62: 283-382.

Fama, E. & French, K. (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives*, 18: 25-46.

Fontana, G. & Gerrard, B. (2004). A Post-Keynesian Theory of Decision Making under Uncertainty. *Journal of Economic Psychology*, 25: 619-637.

Gerrard, B. (1994). Beyond Rational Expectations: A Constructive Interpretation of Keynes's Analysis of Behaviour Under Uncertainty. *Economic Journal*, 104: 327-337.

Good, J. (1950). *Probability and the Weighting of Evidence*. London: Griffin.

Hicks, J. (1962). Liquidity. *Economic Journal*, 72: 787-802.

Hodges, J. & Lehmann, E. (1952). The Use of Previous Experience in Reaching Statistical Decisions. *Annals of Mathematical Statistics*, 23: 396-407.

Huang, X. (2010), *Portfolio Analysis. From Probabilistic to Credibilistic and Uncertain Approaches*. Berlin: Springer-Verlag.

Jouini, E. & Kallal, H. (1995). Martingales and arbitrage in securities markets with transaction costs. *Journal of Economic Theory*, 66: 178-197.

Jouini, E. & Kallal, H. (1999). Viability and equilibrium in securities markets with frictions. *Mathematical Finance*, 9: 275-292.

Jouini, E. & Kallal, H. (2001). Efficient Trading Strategies in the Presence of Market Frictions. *The Review of Financial Studies*, 14: 343-369.

Keynes, J.M. (1921). *A Treatise on Probability. The Collected Writings of John Maynard Keynes, Vol. VII*. London: Macmillan, 1973.

Keynes, J.M. (1936). *The General Theory of Employment, Interest and Money. The Collected Writings of John Maynard Keynes, Vol. VIII*. London: Macmillan, 1973.

Keynes, J.M. (1937). The General Theory of Employment. *Quarterly Journal of Economics*, 51: 209-223.

Koopman, B.O. (1940). The axioms and algebra of intuitive probability, *Annals of Mathematics*, 41, 269-92.

Kopylov, I. (2009). Choice Deferral and Ambiguity Aversion. *Theoretical Economics*, 4: 199-225.

Lawson, T. (1985). Uncertainty in Economic Analysis. *Economic Journal*, 95: 909-927.

Lucas, R. (1978). Asset prices in an exchange economy. *Econometrica*, 46: 1429-1445.

Lucas, R. (1980). Methods and Problems in Business Cycle Theory. *Journal of Money, Credit and Banking*, 12, 696-715.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, 7: 77-91.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investment*. New York: John Wiley and Sons.

Morris, S. & Shin, H.S. (2002). Social Value of Public Information. *American Economic Review*, 92: 1521-1534.

Nishimura, K. & Ozaki, H. (2006). An Axiomatic Approach to Epsilon-Contamination. *Economic Theory*, 27: 333-340.

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

Runde, J. (1994b), Keynes after Ramsey: in defence of *A Treatise on Probability*. *Studies in the History and Philosophy of Science*, 25: 97-121.

Sargent, T. (1983). An Economist's Foreword. In P. Whittle, (ed.), *Prediction and Regulation by Linear Least-Square Methods*. Minneapolis: University of Minnesota Press.

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

Schmeidler, D. (1989). Subjective Probability and Expected Utility without Additivity. *Econometrica*, 57: 571-587.

Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Condition of Risk. *The Journal of Finance*, 19: 425-442.

Sharpe, W. (1991). Capital Asset Prices with and without Negative Holdings. *Journal of Finance*, 46: 489-509.

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

Tobin, J. (1958). Liquidiy Preference as Behavior Towards Risk. *Review of Economic Studies*, 25: 65-86.

Wolintzky, A. (2016). Mechanism design with maxmin agents: theory and an application to bilateral trade. *Theoretical Economics*, 11: 971-1004.

Zappia, C. (2015). Keynes on probability and decision: Evidence from the correspondence with Hugh Townshend. *History of Economic Ideas*, 23, 145-164.

Zappia, C. (2016). Whither Keynesian probability? Impolite techniques for decision-making, *European Journal of the History of Economic Thought*, 23: 835-862.