



Asian Journal of Mathematics & Statistics

ISSN 1994-5418

An Ordinal Utility of Moments for Decisions under Uncertainty

¹Riccardo Cesari and ²Carlo D'Adda

¹Università di Bologna, Dip. Matematiche, viale Filopanti, 5, 40126 Bologna, Italy

²Università di Bologna, Dip. Scienze Economiche, Strada Maggiore,
45, 40125 Bologna, Italy

Abstract: During the last 50 years, a considerable effort has been devoted to connecting the expected utility approach to a utility function directly expressed in terms of moments of the probability distribution of an uncertain prospect. We follow the alternative route of providing, for the first time, the theoretical, autonomous foundation of an ordinal utility function of moments, representing rational choices under uncertainty, free of any independence axiom and compatible with many paradoxes of choice and behavior, well documented in the literature.

Key words: Moments, preferences, mean-variance utility, expected utility, indifference prices, behavioral paradoxes

INTRODUCTION

Tobin (1958), in setting, 50 years ago, the microeconomic foundations of the Keynesian liquidity preference theory in the light of the Markowitz (1952, 1959) portfolio approach, has shown an important equivalence between the Von Neumann and Morgenstern (1944) expected utility (VNM) and a preference function in mean and standard deviation.

His extension of this result to all two-parameter distributions was subsequently proved incorrect by Samuelson (1967), Borch (1969) and Feldstein (1969) so that a mean-variance analysis has been justified for a long time only under the restrictive assumption of quadratic VNM utility or Gaussian distribution.

Two research areas have therefore attracted considerable interest: one focused on suitable distributional assumptions and one devoted to the most appealing form of the utility function.

Regarding the former, the admissible distributions have been identified by the elliptical class (Chamberlain, 1983; Owen and Rabinowitch, 1983), which is closed under linear transformations of random variables. Elliptical distributions are also known as scale-location parameter distributions (as in Tobin's conjecture) or linear distributions (Sinn, 1983; Meyer, 1987; Levy, 1989). They include the symmetric stable distributions analyzed by Fama (1971).

In addition, a large amount of work has been done to justify mean-variance (EV) analysis just as a second-order approximation of the expected utility model, avoiding the absurd assumption of quadratic utility (Hicks, 1962; Pratt, 1964; Arrow, 1965) and its property of increasing risk aversion and decreasing asset demand as wealth increases, making risky assets inferior goods.

Early studies concerning EV as an approximation of expected utility were carried out by Samuelson (1967, 1970), Tsiang (1972, 1974) and Rubinstein (1973), also including higher-order moments in a generalized approach.

Corresponding Author: Riccardo Cesari, Università di Bologna, Dip. Matematiche,
viale Filopanti, 5, 40126 Bologna, Italy

Levy and Markowitz (1979), Kroll *et al.* (1984), Reid and Tew (1986) and Markowitz (1987) assess the effectiveness of the EV approximation, confirming Markowitz's intuition that mean-variance is, in practice, as efficient as expected utility in selecting optimal portfolios.

Even in more theoretical works, such as Hakansson (1972), Baron (1977) and Bigelow (1993), the aim was to investigate the consistency between mean-variance (or moment-) utility and the VNM axioms or utility function, providing restrictions to a simultaneous validity of both approaches.

However, any attempt to make expected utility and moment utility equivalent does not seem a truly compelling task.

Many authors (Borch, 1969, 1973; Levy, 1989) are, in fact, well aware of the different set-up of the two approaches, particularly when the arguments of moment utility are not confined to mean and variance but include all relevant moments of the probability distributions.

Moreover, during the last century, a number of works, from Knight (1921) to Keynes (1921), from Hicks (1935, 1962) to Marschak (1938), from Lange (1944) to Simpson (1950), expressed intriguing suggestions on variance, dispersion and higher-order moments as relevant parameters directly influencing the agent's decisions under uncertainty.

In this study, we develop the foundations of an ordinal utility of moments as a rational and autonomous criterion of choice under uncertainty, showing that it dissolves all the best known behavioral paradoxes which are still embarrassing the expected utility theory. This ordinal approach is strongly reminiscent of standard microeconomic theory and it could be used to reset and generalize both assets demand theory and asset pricing models.

FOUNDATIONS

It is well known that the theory of choice under uncertainty assumes that preferences are defined over the set of probability distribution functions (Savage, 1954; DeGroot, 1970).

Confining ourselves, for ease of exposition, to the case of univariate distributions, we assume that the essential information concerning any distribution F is contained in the m -dimensional vector of moments $M \equiv (\mu, \mu^{(2)}, \mu^{(3)}, \dots, \mu^{(m)})$ where, μ is the mean and $\mu^{(s)}$ is the s -order central moment in original units, so that $(\mu^{(s)})^s$ is the usual central moment of order $s \geq 2$.

Note that, instead of central moments, noncentral moments could, equivalently, be used. Moreover, scale, location and dispersion parameters can be considered in the case of distributions (e.g., stable) for which moments do not exist.

The existence of an ordinal utility of moments is obtained under the assumption of a preference order satisfying the axioms of (1) asymmetry, (2) transitivity and (3) continuity (Appendix).

Assumption of Preference Order

Let $Q \subseteq \mathbb{R}^m$ be a rectangular subset of \mathbb{R}^m (the cartesian product of m real intervals), whose elements are the m -dimensional vectors of moments, $M \in Q$. Let $>$ be a preference order i.e., a binary relation defined by a subset \mathfrak{R} of the Cartesian product $Q \times Q$, whose elements are the ordered pairs of vectors (M_a, M_b) .

We write $M_a > M_b$ instead of $(M_a, M_b) \in \mathfrak{R}$ and we say that M_a is preferred to M_b , corresponding to F_a is preferred to F_b for distribution functions.

Clearly, or $M_a > M_b$ or $M_a \not> M_b$ and both cannot hold: in fact, or $(M_a, M_b) \in \mathfrak{R}$ or $(M_a, M_b) \notin \mathfrak{R}$.

We write $M_a \sim M_b$ (equivalence) if and only if $M_a \not> M_b$ and $M_b \not> M_a$.

Theorem 1: Of Complete Preferences

The preference order is complete.

Theorem 2: Of Negatively Transitive Preferences

The preference order is negatively transitive.

Theorem 3: Of Equivalence Classes

The equivalence \sim is reflexive, symmetric and transitive.

Theorem 4: Of Ordinal Utility on Moments

Under Axioms I, II, III there is a real function $H: Q \rightarrow R$ which represents the preferences $>$, i.e., such that, for every $M_a, M_b \in Q$:

$$M_a > M_b \text{ if and only if } H(M_a) > H(M_b) \tag{1}$$

The function H is unique up to any order-preserving transformation \mathcal{O} :

$$H(M_a) > H(M_b) \text{ if and only if } \mathcal{O}(H(M_a)) > \mathcal{O}(H(M_b)) \tag{2}$$

The function H is called an ordinal utility because it just represents the given preference order $>$ in terms of otherwise arbitrary real numbers (utils).

Theorem 5: Of Continuous Utility

Under Axioms I, II, III and the usual relative topology for Q (intersection of Q with the set of all open rectangles in R^m , including arbitrary unions and finite intersections; Debreu (1959), Rader (1963)), the utility function H is continuous if and only if, for every $M_a \in Q$, the sets $\{M_b \in Q: M_a > M_b\}$ and $\{M_b \in Q: M_b > M_a\}$ belong to the topology.

In the following, we assume that H is continuous with bounded first order partial derivatives. The usual concepts available in the theory of choice under uncertainty can be extended to our approach.

Let F, G be two probability distributions with relevant moment vectors:

$M_F \equiv (\mu_F, \mu_F^{(2)}, \mu_F^{(3)}, \dots, \mu_F^{(m)})$ and $M_G \equiv (\mu_G, \mu_G^{(2)}, \mu_G^{(3)}, \dots, \mu_G^{(m)})$, respectively. Let $H(M)$ be a differentiable utility function.

Definition of Non Satiation

The utility H is non satiated if, for every $\delta > 0$:

$$H(\mu_F, \mu_F^{(2)}, \mu_F^{(3)}, \dots, \mu_F^{(m)}) < H(\mu_F + \delta, \mu_F^{(2)}, \mu_F^{(3)}, \dots, \mu_F^{(m)}) \tag{3}$$

In differential terms: $\frac{\partial H}{\partial \mu} > 0$.

Definition of Risk Aversion

A utility H is risk averse if, for every F :

$$H(\mu_F, 0, 0, \dots, 0) > H(\mu_F, \mu_F^{(2)}, \mu_F^{(3)}, \dots, \mu_F^{(m)}) \quad (4)$$

Note that for $m = 2$, risk aversion means $\frac{\partial H}{\partial \mu^{(2)}} < 0$; for $m \geq 3$, a negative marginal utility of volatility $\frac{\partial H}{\partial \mu^{(2)}} < 0$ does not imply risk aversion.

Definition of Certainty Equivalent

The certainty equivalent of F is defined as the amount C_F such that:

$$H(\mu_F, \mu_F^{(2)}, \mu_F^{(3)}, \dots, \mu_F^{(m)}) = H(C_F, 0, 0, \dots, 0) \quad (5)$$

Definition of Risk

Given two distributions F, G with equal mean, $\mu_F = \mu_G$, we say that F is less risky than G if $H(M_F) > H(M_G)$ for every risk averse utility H .

Definition of Stochastic Dominance

Given two distribution functions, F, G defined over the same support, we have m th-order stochastic dominance of F over G , $F >_m G$, $m \geq 1$, if $F \neq G$ and:

$$F^{(m)}(x) \leq G^{(m)}(x) \quad \forall x$$

where:

$$F^{(m)}(x) = \begin{cases} F(x) & m = 1 \\ \int_{-\infty}^x F^{(m-1)}(y) dy & m \geq 2 \end{cases}$$

Clearly, if $F >_m G$ then $F >_{m+1} G$.

We recall a well known result linking stochastic dominance and VNM expected utility functions.

Theorem 6 on Stochastic Dominance and Expected Utility

We have:

- $F >_1 G \iff E_F(U(x)) \geq E_G(U(x))$ for every U with $U' \geq 0$
- $F >_2 G \iff E_F(U(x)) \geq E_G(U(x))$ for every U with $U' \geq 0, U'' \leq 0$
- $F >_3 G \iff E_F(U(x)) \geq E_G(U(x))$ for every U with $U' \geq 0, U'' \leq 0, U''' \geq 0$

Therefore, first order stochastic dominance means increasing VNM utilities; second order stochastic dominance means increasing and concave VNM utilities.

This is no longer true for the moment ordinal utility.

Theorem 7 on Stochastic Dominance and Moment Utility

If $F >_m G$ then $M_F \equiv (\mu_F, \mu_F^{(2)}, \mu_F^{(3)}, \dots, \mu_F^{(m)}) \neq (\mu_G, \mu_G^{(2)}, \mu_G^{(3)}, \dots, \mu_G^{(m)}) \equiv M_G$ and $(-1)^{k-1} \mu_F^{(k)} > (-1)^{k-1} \mu_G^{(k)}$ for the smallest k for which $\mu_F^{(k)} \neq \mu_G^{(k)}$.

As special cases we have:

- If $F \succ_1 G$ then $\mu_F > \mu_G$
- If $F \succ_2 G$ then $(\mu_F > \mu_G)$ or $(\mu_F = \mu_G \text{ and } \mu_F^{(2)} < \mu_G^{(2)})$
- If $F \succ_3 G$ then $(\mu_F > \mu_G)$ or $(\mu_F = \mu_G \text{ and } \mu_F^{(2)} < \mu_G^{(2)})$ or $(\mu_F = \mu_G \text{ and } \mu_F^{(2)} = \mu_G^{(2)} \text{ and } \mu_F^{(3)} > \mu_G^{(3)})$

In the first case, in particular, it does not necessarily follow (even if it is, however, plausible) that $H(M_F) > H(M_G)$ whenever higher order moments are relevant. In particular, Allais (1953, 1979) assumes that $F \succ_1 G$ implies $H(F) > H(G)$ (axiom of absolute preference). In the expected utility approach it is equivalent to assume $U' > 0$ (nonsatiation).

INDIFFERENCE PRICING AND THE (DIS)SOLUTION OF PARADOXES

The Von Neumann-Morgenstern (1944) theory of choice under uncertainty has long been the standard approach to model the maximizing behavior of agents in financial markets. The basic result is the existence of a utility function $U(\cdot)$ (the VNM utility) describing the optimal decisions of an investor as those maximizing the expected utility of his or her future wealth $E(U(\tilde{W}))$. This beautiful result was obtained at the cost of a special and controversial axiom concerning the structure of preferences in the case of uncertainty, the so-called independence axiom, inducing a linear dependence between utility and probabilities.

In fact, the independent axiom assumes that randomization is irrelevant: it says that if $A \succ B$ then $pA + (1-p)C \succ pB + (1-p)C$ (Machina, 1987; Fishburn, 1982, 1988 and the seminal work in Von Neumann and Morgenstern, 1944, especially chapter 1 and Appendix, where, considering their penetrating results, they ask themselves, p.28, Have we not shown too much?).

The approach followed above replaces the VNM expected utility with the more fundamental and less demanding ordinal utility $H(E(\tilde{W}), \text{Std}(\tilde{W}), \dots)$, allowing us to avoid the main flaws of the former and providing us with a generalized framework for optimal behavior in both complete and incomplete markets. In particular, note that H is a utility of expectations and not an expectation of utilities.

Note also that this approach is different from (and much simpler than) other generalizations (non-expected utility approaches) suggested in the literature after Allais (1953) contribution to the empirical critique of the VNM expected utility theory.

In fact, in most cases, including Allais (1979), Kahneman and Tversky (1979), Chew (1983), Fishburn (1983), Machina (1982) and many others, the utility function is a function of moments of a Bernoullian utility U , or a nonlinear function of both Bernoullian utilities and subjective probabilities:

$$\sum_i U(x_i) \pi(p_i), \frac{\sum_i U(x_i) p_i}{\sum_i T(x_i) p_i}, \sum_i U(x_i) p_i + \left[\sum_i T(x_i) p_i \right]^2, \dots \text{etc.}$$

In our case, we obtain a more general result based on more intuitive elements:

$$H \left(\sum_i x_i p_i, \left(\sum_i x_i^2 p_i \right)^{1/2}, \dots, \left(\sum_i x_i^m p_i \right)^{1/m} \right) \quad (6)$$

Moreover, it is well known that the expected utility can be approximated by a particular function of m central moments:

$$\sum_i U(x_i)P_i \approx U(\mu) + \frac{1}{2}U^{(2)}(\mu)(\mu^{(2)})^2 + \dots + \frac{1}{m!}U^{(m)}(\mu)(\mu^{(m)})^m \quad (7)$$

where, $U^{(j)}$ is the j -th derivative of U , but this is only a special (polynomial) case of H and is not always able to account for the observed phenomena.

The utility of moments is perfectly compatible with the empirical observations in all well known behavioral paradoxes, described in terms of games and lotteries.

In order to show this compatibility we use the so called utility-indifference pricing (Henderson and Hobson, 2009), which can be applied in all cases of personal valuation of non-traded assets and incomplete markets.

For the sake of simplicity, let us consider the two-moment ordinal utility.

Let W be the current wealth of the decision maker and \tilde{G} be the random variable representing the game, with mean M_G and vol Σ_G . Future wealth, in case of a decision to gamble, is given by:

$$\tilde{W} = \frac{W - P_G}{P_0} + \tilde{G} \quad (8)$$

where, $W - P_G$ is wealth left after payment for the game, invested at the riskless rate $r = 1/P_0 - 1$ (often set to zero) and the personal indifference price P_G is defined as the price at which the agent is indifferent to paying the price and entering the game or paying nothing and avoiding the game:

$$H\left(\frac{W - P_G}{P_0} + M_G, \Sigma_G\right) = H\left(\frac{W}{P_0}, 0\right) \quad (9)$$

In the l.h.s., using a Taylor series approximation for small risks (Pratt, 1964), we have:

$$H\left(\frac{W - P_G}{P_0} + M_G, \Sigma_G\right) = H\left(\frac{W}{P_0}, 0\right) + \frac{\partial H}{\partial M}(M_G - \frac{P_G}{P_0}) + \frac{\partial H}{\partial \Sigma}\Sigma_G \quad (10)$$

so that, imposing Eq. 9 and simplifying, we obtain:

$$P_G = M_G P_0 + \left(\frac{\partial H / \partial \Sigma}{\partial H / \partial M} P_0\right) \Sigma_G \quad (11)$$

where, the term in brackets is the (negative) personal price of the vol, P_σ and it is independent of monotonic transformation of H . Therefore, the indifference price, P_G , of the game is obtained as moment quantities, M_G , Σ_G , times subjective moment prices, P_0 , P_σ .

Moreover, Eq. 11 can be easily generalized to higher moments (skewness Γ and kurtosis Ψ):

$$P_G = M_G P_0 + \Sigma_G P_\sigma + \Gamma_G P_\gamma + \Psi_G P_\kappa \quad (12)$$

and this version will be used in the following to show that the behavior of a moment-utility maximizer is perfectly compatible with all the proposed paradoxes, from St. Petersburg (1713) to Friedman and Savage (1948), Allais (1953, 1979) and Kahneman and Tversky (1979).

Table 1: The St. Petersburg game

Flip	Probability of a head	Prize
1	$\frac{1}{2}$	1
2	$\frac{1}{4}$	2
....
n	$\frac{1}{2^n}$	2^{n-1}
....

The St. Petersburg Paradox

The name of the paradox is due to the solution, proposed by Daniel Bernoulli in 1738 to a question posed by his cousin Nicolas Bernoulli, in a letter dated September 9th, 1713. Note that at that time, the method to value an uncertain prospect was established by Christian Huygens in 1657 as the mathematical expectation of the gain (Hacking, 1975). Peter tosses a coin and continues to do so until it should land heads when it comes to the ground. He agrees to give Paul one ducat if he gets heads on the very first throw, two ducats if he gets it on the second, four if on the third, eight if on the fourth and so on, so that with each additional throw, the number of ducats he must pay is doubled.

Paul is the player: if he obtains heads at the first flip he wins 1, at the second flip he wins 2, ..., at the n-th flip he wins 2^{n-1} and so on. The question is to determine a fair price, P(G), to enter the game.

Clearly, the price must be at least 1 (the minimum gain), but the expected gain is infinite (Table 1).

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

However, as Nicolas Bernoulli observed in stating the paradox, nobody would pay an arbitrarily large amount to play the game: it has, he said, to be admitted that any fairly reasonable man would sell his chance, with great pleasure, for twenty ducats.

The famous Bernoulli solution, in log terms, provided a path-breaking device, introducing the concept of utility (moral expectation) and reducing expectation to a finite value:

$$U(G) = \sum_{n=1}^{\infty} U(2^{n-1}) \frac{1}{2^n} < +\infty$$

Alternatively, the price, P(G), of the game can be obtained, in our approach, by considering the sequential game as a one-shot game (a lottery) with an infinity of tickets, identified by natural numbers (1, 2, ..., n, ...) with decreasing probability of extraction ($\frac{1}{2}, \frac{1}{4}, \dots, \frac{1}{2^n}, \dots$) and increasing rewards (1, 2, ..., $2^{n-1}, \dots$). This means that the original game is a portfolio of sub-games $G_n, n \geq 1$ or Arrow-Debreu securities (one for each row in the table above), the n-th of which implies a prize of 2^{n-1} if we get heads at the n-th flip and zero otherwise.

Clearly, each ticket could be sold separately at its price P(G_n) and the price of the lottery is the sum of the prices of all tickets.

We show that a two-moment utility approach is sufficient to solve the paradox.

For each ticket n, the expected value is always $\frac{1}{2}$ and the standard deviation is $0.5(2^n - 1)^{0.5}$.

Table 2: The first two moments of the St. Petersburg game

Ticket	Probability	Prize	Expected Prize	Standard Deviation
1	1/2	1	1/2	0.50
2	1/4	2	1/2	0.87
...
n	1/2 ⁿ	2 ⁿ⁻¹	1/2	0.5(2 ⁿ⁻¹) ^{0.5}
...

$$E(G_n) = 2^{n-1} \frac{1}{2^n} + 0(1 - \frac{1}{2^n}) = \frac{1}{2}$$

$$\text{Var}(G_n) = (2^{n-1} - \frac{1}{2})^2 \frac{1}{2^n} + (0 - \frac{1}{2})^2 (1 - \frac{1}{2^n}) = \frac{2^n - 1}{4}$$

Considering the first two moments (Table 2) the price of the n-th ticket is, from Eq. 12:

$$P(G_n) = \max \left(0, \frac{1}{2}P_0 + \frac{\sqrt{2^n - 1}}{2}P_\sigma \right)$$

where, the limited liability provision has been applied and the price of the game is simply the sum of the (non negative) prices of all tickets:

$$P(G) = \sum_{n=1}^{\infty} P(G_n) = \sum_{n=1}^{\infty} \max \left(0, \frac{P_0}{2} + \frac{P_\sigma}{2} \sqrt{2^n - 1} \right)$$

For example, if $P_0 = 1$ and $P_\sigma = -0.134$ then $P(G_n) = 0$ for $n > 5$ and $P(G) = 1.507$. Coin tosses beyond the fifth have no economic value. Note that in this game-situation, P_σ is not a proper market price but just the gambler's personal price of volatility. Analogously for higher-order moments; a refined price $P(G)$ could also be obtained using higher moments: the skewness of the n-th ticket is $[0.25 (2^n - 1) (2^{n-1} - 1)]^{1/3}$, the kurtosis is $[(2^n - 1) ((2^n - 1)^3 + 1) / 2^{n+4}]^{1/4}$.

The Friedman and Savage (1948) Paradox

In their classical article, Friedman and Savage (1948) observed the difficulty of combining the belief in diminishing marginal utility and the observation that the same individual buy insurance as well as lottery tickets. Clearly, the first choice is evidence of risk aversion but the second one can be rationalized only by a risk loving behavior, being well known that lotteries are largely unfair games, accepted only by individuals having a strong preference for risk.

The clever solution they proposed, in the framework of the Von Neumann-Morgenstern theory, just published a few years before, was based on a complex hypothesis concerning the shape of the utility function, made by three segments, concave-convex-concave, where current income is in the initial convex segment (ivi, paragraph IV).

In terms of moment utility the solution is much simpler and it clarify that skewness is the key element behind the observed behavior.

A typical lottery represents a large chance of losing a small amount (the price of the lottery ticket) plus a small change of winning a large amount (a prize).

Viceversa, the game against which you buy insurance, paying the premium π , contains a small chance of a much larger loss and a large chance of no loss. Let L and J be, respectively, the two random variables so that L (lottery) is preferred to 0 as well as $-\pi$ is preferred to J:

Table 3: The first three moments of the Friedman-Savage lottery (L) and insurance (J)

Moments	L	J
Mean (μ)	-0.50	-0.50
Standard deviation (σ)	50.00	50.00
Skewness (ζ)	232.10	-232.10

$$L = \begin{cases} -1 & 99.99\% \\ 4999 & 0.01\% \end{cases} \quad J = \begin{cases} -5000 & 0.01\% \\ 0 & 99.99\% \end{cases}$$

Let us calculate the first three moments of L and J (Table 3).

Mean and standard deviation are the same but skewness is reversed in sign, with L having a large, positive skewness and J a large, negative one. Assuming the following prices of the three moments: $P_\mu = 1$, $P_\sigma = -0.34$, $P_\zeta = 0.1$ we obtain the prices $P(L) = 5.71 > 0$ and $P(J) = -40.70 < 0$ so that the ticket is bought with pleasure (and considered cheap) and up to 40.7 money units could be willingly paid to avoid the insurable risk.

The Allais Paradox

The Allais (1953) paradox was the first factual evidence against expected utility. In fact, asking people to choose between games A and B, where A gives 1 million with certainty and B gives 1 million with 89% probability and 0 or 5 millions with, respectively, 1 and 10% probabilities:

$$A = \begin{cases} 1 & 100\% \end{cases} \quad B = \begin{cases} 0 & 1\% \\ 1 & 89\% \\ 5 & 10\% \end{cases}$$

people prefer in large majority A to B: $A \succ B$

Then, asking them to choose between A' and B' defined by:

$$A' = \begin{cases} 0 & 89\% \\ 1 & 11\% \end{cases} \quad B' = \begin{cases} 0 & 90\% \\ 5 & 10\% \end{cases}$$

the same people very often prefer B' to A': $B' \succ A'$.

The paradox stems from the fact that, from $A \succ B$, the expected utility approach deduces $A' \succ B'$, which is at variance with the experimental evidence (Allais reports 53% of cases of violation of the logical implication).

In fact, $A \succ B$ means:

$$U(1) > 0.01U(0) + 0.89U(1) + 0.10U(5)$$

but collecting $U(1)$ and adding to both members of the inequality $0.89U(0)$ you obtain, algebraically:

$$0.89U(0) + 0.11U(1) > 0.9U(0) + 0.10U(5)$$

i.e., $A' \succ B'$, against the empirical evidence.

Table 4: The first four moments of the Allais games

	A	B	A'	B'
Mean (μ)	1.000	1.390	0.110	0.500
Standard deviation (σ)	0.000	1.207	0.313	1.500
Skewness (ζ)	0.000	1.666	0.424	2.080
Kurtosis (κ)	0.000	2.032	0.513	2.531

According to Allais, either people in experimental situations do not use the rational thinking used in real world decision-making, or people do not follow the expected utility paradigm.

In fact, using our approach, the rationality of the actual choices may be easily recognized.

Considering each lottery as an asset, the first four moments are in Table 4.

Assuming the following prices of the four moments: $P_0 = 1$, $P_\sigma = -0.34$, $P_\zeta = 0.01$, $P_\kappa = -0.001$ we obtain the prices of the lotteries: $P(A) = 1 > P(B) = 0.994$ and $P(A') = 0.007 < P(B') = 0.008$, in accordance with the Allais experiments.

This means that, using the Marschak triangle as in Machina (1987), the indifference curves in our approach are nonlinear in the probabilities and may display a ‘fanning out’ effect from the sure event A, as implied by actual behavior.

The Kahneman and Tversky (1979) Paradox

In a famous experiment, a systematic violation of the independence axiom was documented: 80% of 95 respondents preferred A to B where:

$$A = \left\{ \begin{array}{ll} 3000 & 100\% \end{array} \right. \quad B = \left\{ \begin{array}{ll} 0 & 20\% \\ 4000 & 80\% \end{array} \right.$$

65% preferred B' to A' where:

$$A' = \left\{ \begin{array}{ll} 0 & 75\% \\ 3000 & 25\% \end{array} \right. \quad B' = \left\{ \begin{array}{ll} 0 & 80\% \\ 4000 & 20\% \end{array} \right.$$

and more than 50% of respondents violated the independent axiom, given that, if Q pays 0 for sure, then:

$$A'' = \left\{ \begin{array}{ll} Q & 75\% \\ A & 25\% \end{array} \right. \quad B'' = \left\{ \begin{array}{ll} Q & 75\% \\ B = \left\{ \begin{array}{ll} 0 & 20\% \\ 4000 & 80\% \end{array} \right. & 25\% \end{array} \right.$$

and B'' is considered equal to B' in terms of outcomes and probabilities.

Note that treating lotteries as assets implies that linear combinations such as $0.75Q + 0.25A$ are meaningful and $P(0.75Q + 0.25A) = 0.25P(A) + P(A')$.

The point is that, in terms of valuation, B' and B'' are not the same asset and B'' is equivalent to:

$$B'' = \left\{ \begin{array}{ll} 0 & 75\% \\ P(B) & 25\% \end{array} \right. \neq B' = \left\{ \begin{array}{ll} 0 & 80\% \\ 4000 & 20\% \end{array} \right.$$

Table 5: The first four moments of Kahneman-Tversky games

	A'	B	B'
Mean (μ)	750.00	3200.00	800.00
Standard deviation (σ)	1299.04	1600.00	1600.00
Skewness (ζ)	1362.84	-1831.54	1831.54
Kurtosis (κ)	1605.52	2148.28	2148.28

Table 6: The first four moments of the Tversky-Kahneman games

	A = A'	B	B''
Mean (μ)	7.500	9.000	36.000
Standard deviation (σ)	12.990	18.000	18.000
Skewness (ζ)	13.628	20.605	-20.605
Kurtosis (κ)	16.055	24.168	24.168

Using the first four moments in Table 5 and assuming the following prices of the four moments: $P_0 = 1, P_\sigma = -0.2, P_\zeta = 0.1, P_\kappa = -0.001$ we obtain the prices of the lotteries: $P(A) = 3000 > P(B) = 2694.70$ and $P(A') = 624.87 < P(B') = 661.01$, in accordance with the experimental results. Note also that $P(B'') = 561.28 < P(A') < P(B')$.

The Tversky and Kahneman (1981) Paradox

Most subjects, confronted with the following alternatives: A versus B and A' versus B', prefer B to A but also A' to B' where:

$$A = \begin{cases} 0 & 75\% \\ 30 & 25\% \end{cases} \quad B = \begin{cases} 0 & 80\% \\ 45 & 20\% \end{cases}$$

$$A' = \begin{cases} 0 & 75\% \\ 30 & 25\% \end{cases} \quad B' = \begin{cases} 0 & 75\% \\ B'' = \begin{cases} 0 & 20\% \\ 45 & 80\% \end{cases} & 25\% \end{cases}$$

The paradox (reversal or isolation effect) stems from the fact that not only $A = A'$ but also $B = B'$ in terms of ultimate outcomes and probabilities.

However, from the point of view of our theory of valuation and choice, the two-stage frame in B' is not irrelevant: in B, not 0 means 45; in B', not 0 means a new game B'', which can be sold for a certain price.

Using the first four moments in Table 6 and assuming the following prices of the four moments: $P_0 = 1, P_\sigma = -0.2, P_\zeta = 0.05, P_\kappa = -0.1$, we obtain the prices of the lotteries: $P(A) = 3.98 < P(B) = 4.01$ and $P(A') = 3.98 > P(B') = 3.84$, being $30 > P(B'') = 28.95$. This result is in accordance with the Tversky and Kahneman (1981) experiment, showing that, in effect, no paradox is implied in the observed behavior.

In particular, note, once again, that in terms of valuation:

$$B' = \begin{cases} 0 & 75\% \\ B'' = \begin{cases} 0 & 20\% \\ 45 & 80\% \end{cases} & 25\% \end{cases} \quad \text{is} \quad B' = \begin{cases} 0 & 75\% \\ P(B'') & 25\% \end{cases} \quad \text{and not} \quad B = \begin{cases} 0 & 80\% \\ 45 & 20\% \end{cases}$$

and the risk-neutral probabilities for B'', for which the personal price of B'' is the expected value:

$$P(B'') = (1 - \hat{p})0 + \hat{p}45$$

are given by:

$$B'' = \begin{cases} 0 & 1 - \frac{28.95}{45} \approx 36\% \\ 45 & \frac{28.95}{45} \approx 64\% \end{cases}$$

This observation also holds for the Markowitz (1959) formulation of Allais's experiment:

$$B' = \begin{cases} 1 & 89\% \\ 5 & 91\% \\ 0 & 9\% \end{cases} \quad \text{is } B' = \begin{cases} 1 & 89\% \\ P(B'') & 11\% \end{cases} \quad \text{and not } B = \begin{cases} 1 & 89\% \\ 5 & 10\% \\ 0 & 1\% \end{cases}$$

CONCLUSIONS

After Tobin (1958), considerable effort has been devoted to connecting the expected utility approach to a utility function directly expressed in terms of moments. In contrast with this approach, we have provided the theoretical foundation of an ordinal utility function of moments which is free of any independence axiom and compatible with all the behavioral paradoxes documented in recent and less recent works on decisions under uncertainty. This moment-utility can be used as the starting point for a new formulation of asset demand models and asset pricing, having more general properties and greater flexibility than existing expected utility results. Future research could be addressed in this promising direction.

APPENDIX: BASIC AXIOMS AND PROOFS OF THE THEOREMS

Let $(\Omega, \mathfrak{S}, \wp)$ be a standard probability space, Ω being the set of elementary events (states of the world), \mathfrak{S} the set (σ -algebra) of subsets of Ω (events), \wp a (subjective) probability measure of the events. Given the set A of all possible actions or decisions, all couples (ω, a) with $\omega \in \Omega$ and $a \in A$, are mapped onto a real vector of monetary consequences $c \in \mathbb{R}^n$, the Euclidean space of n -dimensional real vectors, so that $X(\omega, a) = c$ or $X_a(\omega) = c$ is a random variable and $F_a \in \mathcal{F}$ is its probability distribution function. Clearly, the preferences over acts in A are, equivalently, preferences over the set of random variables X_a as well as preferences over the set \mathcal{F} of distribution functions. Let us confine ourselves, for ease of exposition, to the case of univariate distributions ($n = 1$) and assume that the essential information concerning any distribution F is contained in the m -dimensional vector of moments $M \equiv (\mu, \mu^{(2)}, \mu^{(3)}, \dots, \mu^{(m)})$ where μ is the mean and $\mu^{(s)}$ is the s -order central moment in original units:

Definition of s -order modified central moment: $\mu^{(s)} \equiv \begin{cases} 1 & s = 0 \\ \mu & s = 1 \\ \left(\int (x - \mu)^s dF \right)^{\frac{1}{s}} & s \geq 2 \end{cases} \quad (\text{A.1})$

Note that $(\mu^{(s)})^s$ is the usual central moment of order $s \geq 2$.

Let $Q \subseteq \mathbb{R}^m$ be a rectangular subset of \mathbb{R}^m (the Cartesian product of m real intervals), whose elements are the m -dimensional vectors of moments, $M \in Q$.

Assumption of Preference Order

Let $>$ be a preference order i.e., a binary relation defined by a subset \mathfrak{R} of the Cartesian product $Q \times Q$, whose elements are the ordered pairs of vectors (M_a, M_b) .

We write $M_a > M_b$ instead of $(M_a, M_b) \in \mathfrak{R}$ and we say that M_a is preferred to M_b , corresponding to F_a is preferred to F_b .

Clearly, or $M_a > M_b$ or $M_a \not> M_b$ and both cannot hold: in fact, or $(M_a, M_b) \in \mathfrak{R}$ or $(M_a, M_b) \notin \mathfrak{R}$.

I. Axiom of Asymmetric Preferences

We assume that $>$ is asymmetric i.e., that:

$$\text{if } M_a > M_b \text{ then } M_b \not> M_a \tag{A.2}$$

The relation is therefore irreflexive and, moreover, if $M_b \not> M_a$ then two alternative cases are possible: either $M_a > M_b$ or $M_a \not> M_b$.

In the latter case we say that M_a and M_b are equivalent and we write $M_a \sim M_b$.

Definition of Equivalence

$M_a \sim M_b$ if and only if $M_a \not> M_b$ and $M_b \not> M_a$

Theorem 1 of Complete Preferences

Given $M_a, M_b \in Q$ then one and only one case holds: $M_a > M_b$ or $M_b > M_a$ or $M_a \sim M_b$.

Proof: It is easy to show that any two cases are a contradiction. ■

II. Axiom of Transitive Preferences

We assume that $>$ is transitive:

$$\text{if } M_a > M_b \text{ and } M_b > M_c \text{ then } M_a > M_c \tag{A.3}$$

Definition of Weak Order

The preference order $>$ is a weak order if it is asymmetric and transitive.

Definition of Negatively Transitive Preferences

If $M_a \not> M_b$ and $M_b \not> M_c$ then $M_a \not> M_c$.

Lemma 1

$>$ is negatively transitive if and only if, for every $M_a, M_b, M_c \in Q$, $M_a > M_b$ implies $M_a > M_c$ or $M_c > M_b$.

Proof: Under negative transitivity, if $M_a > M_b$ but $M_a \not> M_c$ and $M_c \not> M_b$ then by negative transitivity $M_a \not> M_b$ against the assumption.

Viceversa, if $M_a > M_b$ implies $M_a > M_c$ or $M_c > M_b$ and negative transitivity is false we have from $M_a \not> M_c$ and $M_c \not> M_b$ that $M_a > M_b$ so that $M_a > M_c$ or $M_c > M_b$, in both cases a contradiction. ■

Theorem 2 of Negatively Transitive Preferences

Asymmetric and transitive preferences are equivalent to asymmetric and negatively transitive.

Proof: Under transitivity if $M_a \succ M_b$ and $M_b \succ M_c$ then $M_a \succ M_c$; therefore, by asymmetry, if $M_b \not\succeq M_a$ and $M_c \not\succeq M_b$ then $M_c \not\succeq M_a$ which is negative transitivity.

Viceversa, under negative transitivity, if $M_a \succ M_b$ and $M_b \succ M_c$ then, from Lemma 1, ($M_a \succ M_c$ or $M_c \succ M_b$) and ($M_b \succ M_a$ or $M_a \succ M_c$). But, by asymmetry, $M_b \not\succeq M_a$ and $M_c \not\succeq M_b$ so that $M_c \succ M_b$ and $M_b \succ M_a$ are false. Therefore $M_a \succ M_c$ which means transitivity. ■

Theorem 3 of Equivalence Classes

The equivalence \sim is reflexive, symmetric and transitive and we have:

$$\text{if } M_a \succ M_b \text{ and } M_a \sim M_c \text{ then } M_c \succ M_b; \text{ if } M_a \succ M_b \text{ and } M_b \sim M_c \text{ then } M_a \succ M_c \quad (\text{A.4})$$

Moreover, \succ on Q/\sim (the set of equivalence classes of Q under \sim) is a strict order i.e., it is a weak order and for every equivalence class $M_A, M_B \in Q/\sim$ one and only one case holds: $M_A \succ M_B$ or $M_B \succ M_A$ (weak connectedness).

Proof: The equivalence is clearly reflexive and symmetric. Suppose it is not transitive: $M_a \sim M_b$ and $M_b \sim M_c$ but $M_a \sim M_c$ is false. Then, by definition, either $M_a \succ M_c$ or $M_c \succ M_a$. From Lemma 1, in the first case, $M_a \succ M_b$ or $M_b \succ M_c$; in the second case $M_c \succ M_b$ or $M_b \succ M_a$, in contradiction with the hypothesis.

If $M_a \succ M_b$ and $M_a \sim M_c$ then by Theorem 1 and Lemma 1 we have $M_c \succ M_b$. If $M_a \succ M_b$ and $M_b \sim M_c$ then by Theorem 1 and Lemma 1 we have $M_a \succ M_c$.

Finally, \succ on Q/\sim is a weak order, being asymmetric and negative transitive: under symmetry, if $M_A \succ M_B$ and $M_B \succ M_A$ then exist M_a, M_a' in M_A and M_b, M_b' in M_B such that $M_a \sim M_a', M_b \sim M_b'$ and $M_a \succ M_b$ and $M_b' \succ M_a'$. From (A.4) $M_a \succ M_b'$ and $M_b' \succ M_a'$ which is a contradiction; for negative transitivity, if $M_A \succ M_B$ then $M_a \succ M_b$ for any M_a in M_A and M_b in M_B and if M_c is any vector in M_C then we have, from Lemma 1, $M_c \succ M_b$ or $M_a \succ M_c$. Therefore, $M_c \succ M_B$ or $M_A \succ M_C$.

For weak connectedness, given that M_A and M_B are disjoint, from Theorem 1, either $M_a \succ M_b$ or $M_b \succ M_a$ for every M_a in M_A and M_b in M_B . Therefore, either $M_A \succ M_B$ or $M_B \succ M_A$. ■

III. Axiom of Continuity

There is a countable subset $D \subset Q/\sim$ that is \succ -dense in Q/\sim i.e., for every $M_A, M_C \in Q/\sim \setminus D$, $M_A \succ M_C$ there is $M_B \in D$ such that:

$$M_A \succ M_B \text{ and } M_B \succ M_C \quad (\text{A.5})$$

Note that the subset of rational numbers is \succ -dense and \prec -dense in the set of real numbers.

Theorem 4 of Ordinal Utility on Moments

Proof: The proof follows the steps as in Fishburn (1970). ■

Theorem 5 of Continuous Utility

Proof: See Debreu (1964). ■

Theorem 6 on Stochastic Dominance and Expected Utility

Proof: (i) using integration by parts:

$$E_F(U(x)) - E_G(U(x)) = \int U(x)dF(x) - \int U(x)dG(x) = \int U(x)d(F(x) - G(x)) = - \int (F(x) - G(x))U'(x)dx$$

Therefore, if $F(x) \leq G(x)$ then $E_F(U(x)) \geq E_G(U(x))$. Viceversa, if $E_F(U(x)) \geq E_G(U(x))$ let I be an interval in which $F(x) > G(x)$ and let χ_I be the indicator function of I. Define:

$$U(x) = \int \chi_I(x)dx$$

so that $U'(x) = \chi_I(x) \geq 0$ and inserted into the above equation gives a contradiction. For (ii) and (iii) Fishburn and Vickson (1978) and Whitmore (1970). ■

Theorem 7 on Stochastic Dominance and Moment Utility

Proof: Apply Fishburn (1980, theorem 1) and the relation between central, μ and non central, v , moments:

$$(\mu_F^{(s)})^s = \sum_{j=0}^s \binom{s}{j} (v_F^{(s-j)})^{s-j} (-\mu_F)^j \quad \forall s \geq 2. \quad \blacksquare$$

REFERENCES

Allais, M., 1953. Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine. *Econometrica*, 21: 503-546.

Allais, M., 1979. The So-called Allais Paradox and Rational Decisions Under Uncertainty, Expected Utility. In: *Hypotheses and the Allais Paradox*, Allais, M. and O. Hagen (Eds.), Reidel Pub. Co., Dordrecht, The Netherlands, pp: 437-681.

Arrow, K.J., 1965. Aspects of the Theory of Risk-bearing. *Yrjö Jahnsson Lectures*, Helsinki, Finland.

Baron, D.P., 1977. On the utility theoretic foundations of mean-variance analysis. *J. Finance*, 32: 1683-1697.

Bigelow, J.P., 1993. Consistency of mean-variance analysis and expected utility analysis: A complete characterization. *Econ. Lett.*, 43: 187-192.

Borch, K., 1969. A note on uncertainty and indifference curves. *Rev. Econ. Stud.*, 36: 1-4.

Borch, K., 1973. Uncertainty and indifference curves: A correction. *Rev. Econ. Stud.*, 40: 141-141.

Chamberlain, G., 1983. A characterization of the distributions that imply mean-variance utility functions. *J. Econ. Theory*, 29: 185-201.

Chew, S.H., 1983. A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica*, 51: 1065-1092.

DeGroot, M.H., 1970. *Optimal Statistical Decisions*. McGraw-Hill, New York.

Debreu, G., 1959. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*. Cowles Foundation Monograph, New Haven, YUP.

Debreu, G., 1964. Continuity properties of Paretian utility. *Int. Econ. Rev.*, 5: 285-293.

Fama, E.F., 1971. Risk, return and equilibrium. *J. Polit. Econ.*, 79: 30-55.

- Feldstein, M.S., 1969. Mean-variance analysis in the theory of liquidity preference and portfolio selection. *Rev. Econ. Stud.*, 36: 5-12.
- Fishburn, P.C., 1970. *Utility Theory for Decision Making*. Wiley, New York.
- Fishburn, P.C. and R.G. Vickson, 1978. Theoretical Foundations of Stochastic Dominance. In: *Stochastic Dominance*, Whitmore, G.A. and M.C. Findlay (Eds.). Lexington Books, Lexington, MA, pp: 37-113.
- Fishburn, P.C., 1980. Stochastic dominance and moments of distributions. *Math. Operat. Res.*, 5: 94-100.
- Fishburn, P.C., 1982. *The Foundations of Expected Utility*. Reidel, Dordrecht.
- Fishburn, P.C., 1983. Transitive measurable utility. *J. Econ. Theory*, 31: 293-317.
- Fishburn, P.C., 1988. *Nonlinear Preference and Utility Theory*. Wheatsheaf, Brighton, Brighton, Hove UK.
- Friedman, M. and L.J. Savage, 1948. The utility analysis of choices involving risk. *J. Polit. Econ.*, 56: 279-304.
- Hacking, I., 1975. *The Emergence of Probability*. Cambridge University Press, Cambridge.
- Hakansson, N.H., 1972. Mean-variance analysis in a finite world. *J. Financ. Quant. Anal.*, 7: 1873-1880.
- Henderson V. and D. Hobson, 2009. Utility Indifference Pricing: An Overview. In: *Indifference Pricing: Theory and Applications*, Carmona, R. (Ed.). Princeton University Press, Princeton, pp: 44-74.
- Hicks, J.R., 1935. A suggestion for simplifying the theory of money. *Economica*, 2: 1-19.
- Hicks, J.R., 1962. Liquidity. *Econ. J.*, 72: 787-802.
- Kahneman, D. and A. Tversky, 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, 47: 263-291.
- Keynes, J.M., 1921. *A Treatise on Probability*. Macmillan, London.
- Knight, F.H., 1921. *Risk, Uncertainty and Profit*. Houghton Mifflin, Boston.
- Kroll, Y., H. Levy and H.M. Markowitz, 1984. Mean-variance versus direct utility maximization. *J. Finance*, 39: 47-61.
- Lange, O., 1944. *Price Flexibility and Employment*. Cowles Commission Monograph, Principia Press, Bloomington, IN USA.
- Levy, H. and H.M. Markowitz, 1979. Approximating expected utility by a function of mean and variance. *Am. Econ. Rev.*, 69: 308-317.
- Levy, H., 1989. Two-moment decision models and expected utility maximization: Comment. *Am. Econ. Rev.*, 79: 597-600.
- Machina, M.J., 1982. Expected utility' analysis without the independence axiom. *Econometrica*, 50: 277-323.
- Machina, M.J., 1987. Choice under uncertainty: Problems solved and unsolved. *Econ. Perspect.*, 1: 121-154.
- Markowitz, H.M., 1952. Portfolio selection. *J. Finance*, 7: 77-91.
- Markowitz, H.M., 1959. *Portfolio Selection: Efficient Diversification of Investments*. 2nd Edn., Wiley, New York.
- Marschak, J., 1938. Money and the theory of assets. *Econometrica*, 6: 311-325.
- Markowitz, H., 1987. *Mean-variance Analysis in Portfolio Choice and Capital Markets*. Basil Blackwell Inc., Oxford and New York.
- Meyer, J., 1987. Two-moment decision models and expected utility maximization. *Am. Econ. Rev.*, 77: 421-430.
- Owen, J. and R. Rabinovitch, 1983. On the class of elliptical distributions and their applications to the theory of portfolio choice. *J. Finance*, 38: 745-752.

- Pratt, J., 1964. Risk aversion in the small and in the large. *Econometrica*, 32: 122-136.
- Rader, T., 1963. The existence of a utility function to represent preferences. *Rev. Econ. Stud.*, 30: 229-232.
- Reid, D.W. and B.V. Tew, 1986. Mean-variance versus direct utility maximization: A comment. *J. Finance*, 41: 1177-1179.
- Rubinstein, M., 1973. The fundamental theorem of parameter-preference security valuation. *J. Financ. Quant. Anal.*, 1: 61-69.
- Samuelson, P.A., 1967. General proof that diversification pays. *J. Financ. Quant. Anal.*, 2: 1-13.
- Samuelson, P.A., 1970. The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. *Rev. Econ. Stud.*, 37: 537-542.
- Savage, L.J., 1954. *The Foundation of Statistics*. 2nd Edn., Wiley, New York.
- Simpson, P.B., 1950. Risk allowances for price expectations. *Econometrica*, 18: 253-259.
- Sinn, H.W., 1983. *Economic Decisions under Uncertainty*. North-Holland, Amsterdam, The Netherlands.
- Tobin, J., 1958. Liquidity preference as behavior toward risk. *Rev. Econ. Stud.*, 25: 65-86.
- Tsiang, S.C., 1972. The rationale of the mean-standard deviation analysis, skewness preference and the demand for money. *Am. Econ. Rev.*, 62: 354-371.
- Tsiang, S.C., 1974. The rationale of the mean-standard deviation analysis: Reply and errata for original article. *Am. Econ. Rev.*, 64: 442-450.
- Tversky, A. and D. Kahneman, 1981. The framing of decisions and the psychology of choice. *Science*, 211: 453-458.
- Von Neumann, J. and O. Morgenstern, 1944. *Theory of Games and Economic Behavior*. 1st Edn., Princeton University Press, Princeton, UK., ISBN: 0691003629.
- Whitmore, G.A., 1970. Third-degree stochastic dominance. *Am. Econ. Rev.*, 60: 457-459.