Updating Choquet valuation and discounting information arrivals

A. Lapied, Robert Kast

Updating Choquet valuation and discounting information arrivals

June 14, 2005

André Lapied, Université d'Aix-Marseille, Greqam et Idep, lapied@ehess.univ-mars.fr
Robert Kast, CNRS, Lameta et Idep, kast@ensam.inra.fr

DT LAMETA - 2005 n°9

Abstract

Choquet expected utility has been convinced of being inconsistent within a dynamic framework by several authors. We explore different possible definitions for conditional Choquet integrals and their implications for updating capacities. We confront the definitions with dynamic consistency when information arrives along with time through a Choquet version of the Net Present Value. We get the intuition that only one definition is dynamically consistent and prove it in a decision model where time is discounted according to the agent's preferences. Our result is illustrated by a simplified real investment problem. Possible extensions to dynamically consistent valuation of uncertain cash flows is questioned in the conclusion.

Updating Choquet valuation and discounting information arrivals

Introduction

Choquet capacities and integrals extend probabilities and Lebesgue integrals to the non-additive case. They have been applied in decision theory since they were re-discovered by Schmeidler (1986) and (1989): the first "Choquet Expected Utility" model. They may be used to valuate future uncertain payoffs by economic agents who consider a set of probability distributions (instead of a unique one). However, in a dynamic decision or valuation problem, several difficulties arise. The main one is that some authors have shown that dynamic consistency (almost) implies expected utility (Border and Segal (1994)) and the Bayesian updating rule (Epstein and Le Breton (1993)). An other difficulty is that we have several updating formulas for capacities and no uncontroversial definitions of "conditional Choquet integrals". In this paper we review definitions and updating formulas and then we question their dynamic consistency in the setting of an uncertain cash flow valuation problem. We show that only one rule fits in a dynamically consistent model with Choquet integrals. Furthermore, we show that this result is not in contradiction with the previous non consistency results, because the rule violates an other axiom. An example illustrates why and how.

In a dynamic setting, the Bayesian updating rule is not always implied by the (implicit) definition of conditional expectations. This comes from the non linearity of Choquet integrals. Consider random variables X and Y on *(S, F, m)* where m is a probability distribution, the three equivalent formulations defining conditional expectations for any set C in F: $\int_C X dm = \int_C E(X/Y) dm$ (1), $\int_C [X - E(X/Y)] dm = 0$ (2) and $\int_C [E(X/Y) - X] dm = 0$ (3) are not equivalent anymore if integrals are Choquet's. Furthermore, if n is a Choquet capacity on F, $\int_C X dn = 0$, can be interpreted in two different ways: $\int_S [X.1_C] dn = 0$ or $\int_{-\infty}^{0} [1 - n([X > x] \cap C)] dx + \int_{0}^{+\infty} n([X > x] \cap C) dx = 0$ which may differ depending on the measurable set C.

If, or when, these formulas are equivalent, updating is necessarily Bayesian. However, if Choquet integrals represent preferences over a set of uncertain outcomes, each of the formulas reflects a particular way to value the future. Indeed, conditioning is a re-action in front of information arrivals and should be determined by some time consistency axioms on preferences.

Furthermore, the future is not made of uncertain states only: Time plays a role in the picture (at least as a parameter). Time is usually measured by discount factors, so that the valuation of a cash flow is obtained by its Net Present Value (NPV from now on). This formula is relevant in some cases. For instance, when financial markets define the term structure of interest rates representing trades in riskless bonds, rates or discount factors yield an economic measure of time. When time is appreciated by an individual ("preferences for present consumption"), however, the linear structure of the present Value is obtained under a separability axiom (Koopmans (1972)) which is questionable (see, for instance, Gilboa (1989), Shalev (1997), De Waegenaere and Wakker (2001) or Chateauneuf and Rebille (2003)). We shall not tackle the separability problem in this paper although we think it is at the heart of dynamic valuation[1]. We shall assume that time and uncertainty are subjectively measured by an agent's preferences if they satisfy some axioms and if they establish a hierarchy between these two components of the future: First, uncertainty is valued according to preferences over uncertain payoffs which satisfy specific axioms, and then, time is valued by discount factors on certain cash flows (cash flows of certainty equivalents).

In the first section of this paper, we investigate the different versions of conditional Choquet integrals and we derive the corresponding updating rules for capacities. This section extends previous works by Dempster (1967) and Shafer (1976), Gilboa and Schmeidler (1993), Cohen et al. (1993), and is mainly based on Denneberg (1994) and on Chateauneuf, Kast and Lapied (2001). Results are summarized in a tableau.

In the second section, we introduce information arrivals along with time and we interpret Choquet integrals as valuations of random payoffs within an extension of the Net Present Value formula. We get the intuition that only one definition of conditional Choquet integral can be consistent with time valuation: Not surprisingly we find the Chateauneuf, Kast and Lapied (1991) one (CKL) because it was constructed to satisfy a time consistency condition. The CKL updating rule enforces the role of comonotonicity between the information and the payoffs vectors.

[1] In the case where the economic measure of time is the financial markets' one (riskless bond prices), additivity of the present value is founded on no arbitrage in tight markets (no bid-ask spreads). However, lending and borrowing can never be obtained at the same rate in real markets because of transaction costs, therefore questioning the additivity assumption in the NPV.

This result is proved in section 3 where dynamic consistency, as defined by Kreps and Porteus (1978) for a decision making process, is adapted to our pure valuation problem. A simple real investment valuation problem is given to link valuation and an optimal decision strategy. The contradiction between dynamic consistency and non-Bayesian updating rules is questioned: We follow Karni and Schmeidler (1991), Machina (1998) and Sarin and Wakker (1998) to figure out which one of the set of consistency axioms these authors have put forward our rule violates. A counter-example shows that Consequentialism is the one: The role played by comonotonicity in CKL formula gives relevance to future payoffs, even though information excludes them.

1. Conditional Choquet integrals and updating rules for capacities

Let S be a finite space representing uncertain states $s \in S$ and $A, B \subset S$ be events. The decision maker's preferences over uncertain payoffs are assumed to be represented by a Choquet integral (Chateauneuf (1991), see section 3) and we note it I in this section with ν for the corresponding capacity. $I(1_A/1_B)$ is the individual value given to 1_A, i.e. payoff of 1€ if and only if event A occurs, conditional on the realisation of event B. Then:

$s \in B \Rightarrow I(1_A/1_B) = I(1_A/1_B=1) = \int_B 1_A d\nu_{/\sigma(B)} \equiv \nu(A/B)$, where $\sigma(B)$ is the σ-algebra generated by B and,

$s \in B^C \Rightarrow I(1_A/1_B) = I(1_A/1_B=0) = \int_{B^C} 1_A d\nu_{/\sigma(B)} \equiv \nu(A/B^C)$.

In Denneberg (1994) the conditional Choquet expectation was defined by:

$$\forall A, B, C \subset S, \quad \int_C [1_A - I(1_A/1_B)] d\nu = 0,$$

In Chateauneuf, Kast and Lapied (2001), the Choquet expectation was defined by:

$$\forall A, B, C \subset S, \quad \int_C 1_A d\nu = \int_C I(1_A/1_B) d\nu_{/\sigma(B)}.$$

Both formulas extend the implicit definition of mathematical (Lebesgue) expectation with respect to a probability μ:

$\int_C X d\mu = \int_C E(X/Y) d\mu$, $\int_C [X - E(X/Y)] d\mu = 0$ or $\int_C [E(X/Y) - X] d\mu = 0$, which are equivalent.

However, the equivalence doesn't hold for non linear integrals. Furthermore, neither Denneberg nor Chateauneuf et al. considered the general case with $C \in s(B)$ but restricted their attention to $C = S$, for the first one, to B and to B^c for the second authors.

In addition, $\int_C X dn = 0$ may be interpreted in two different ways for Choquet integrals: $\int_S [X.1_C] dn = 0$ or $\int_{-\infty}^{0} [1 - n([X > x] \cap C)] dx + \int_{0}^{+\infty} n([X > x] \cap C) dx = 0$ which may differ depending on the measurable set C. In the following propositions, we look through all cases.

Proposition 1.1[2]**:** *If conditional Choquet expectations* $I(1_A / 1_B)$ *are defined by: "A, B Ì S,*

(4) $\int_S 1_A dn = \int_S I(1_A/1_B) dn_{/s(B)}$,

then :

(i) *If* 1_A *and* 1_B *are comonotonic random variables,*

$$n(\mathrm{A/B}) = \frac{n(\mathrm{A} \cap \mathrm{B})}{n(\mathrm{B})}$$ *(Bayes updating rule).*

(ii) *If* 1_A *and* 1_B *are antimonotonic (i.e.* 1_A *and* -1_B *are comonotonic) random variables,*

$$n(\mathrm{A/B}) = \frac{n(\mathrm{A} \cup \mathrm{B^C}) - n(\mathrm{B^C})}{1 - n(\mathrm{B^C})}$$ *(Dempster-Schafer updating rule).*

Proof : All proofs of this section are in Appendix 1.

Note that the same results hold for $I(1_A / 1_B = 0) = n(A / B^C)$.

The general case where 1_A and 1_B are not comonotonic nor antimonotonic random variables cannot be solved by relation (4). Indeed, we would obtain one equation for two unknowns: $I(1_A / 1_B = 0)$ and $I(1_A / 1_B = 1)$ which are not necessarily equal to *0* nor to *1* as in the two previous cases and cannot be ranked, in general, to compute the Choquet integral.

Proposition 1. 2: *If conditional Choquet expectations* $I(1_A / 1_B)$ *are defined by: "A, B Ì S,*

(5) $\int_C 1_A dn = \int_C I(1_A/1_B) dn_{/s(B)}$,

then, for C Î {B, B^C*}:*

$$n(\mathrm{A/B}) = \frac{n(\mathrm{A} \cap \mathrm{B})}{n(\mathrm{B})}$$ *(Bayes updating rule).*

[2] The same result was obtained in Chateauneuf, Kast, Lapied (2001) under more restrictive assumptions.

$$n(\mathrm{A/B^C}) = \frac{n(\mathrm{A} \cap \mathrm{B^C})}{n(\mathrm{B^C})}$$ *(Bayes updating rule).*

Proposition 1. 3: *If the conditional Choquet Expectations* $I(1_A/1_B)$ *is defined by: "A, B Ì S,*

(6) $\int_S [1_A - I(1_A/1_B)]dn = 0$,

then:

If 1_A *and* 1_B *are comonotonic or antimonotonic random variables,*

$$n(\mathrm{A/B}) = \frac{n(\mathrm{A} \cap \mathrm{B})}{1 + n(\mathrm{A} \cap \mathrm{B}) - n(\mathrm{A} \cup \mathrm{B^C})}$$ *(Full Bayes updating rule).*

Notice that the same results hold for $I(1_A/1_B = 0) = n(A/B^C)$, and that the general case where 1_A and 1_B are not comonotonic nor antimonotonic random variables cannot be solved by relation (6).

Proposition 1. 4: *If conditional Choquet expectations* $I(1_A/1_B)$ *are defined by: "A, B Ì S,*

(7) $\int_C [1_A - I(1_A/1_B)]dn = 0$,

where $C \in \{B, B^C\}$, then:

(i) if $\int_C X dn = \int_S X.1_C dn$,

$$n(\mathrm{A/B}) = \frac{n(\mathrm{A} \cap \mathrm{B})}{1 + n(\mathrm{A} \cap \mathrm{B}) - n(\mathrm{A} \cup \mathrm{B^C})}$$ *, (Denneberg (1994))*

(ii) if $\int_C X dn = \int_{-\infty}^{0} [n(\{X \geq x\} \cap C) - n(C)]dx + \int_{0}^{+\infty} n(\{X \geq x\} \cap C)dx$,

$$n(\mathrm{A/B}) = \frac{n(\mathrm{A} \cap \mathrm{B})}{n(\mathrm{B})}.$$

Proposition 1. 5: *If conditional Choquet expectations* $I(1_A/1_B)$ *are defined by: "A, B Ì S,*

(8) $\int_S [I(1_A/1_B) - 1_A]dn = 0$,

then:

If 1_A *and* 1_B *are comonotonic or antimonotonic random variables,*

$$n(\mathrm{A/B}) = \frac{1 - n(\mathrm{A^C} \cup \mathrm{B^C})}{1 + n(\mathrm{A^C} \cap \mathrm{B}) - n(\mathrm{A^C} \cup \mathrm{B^C})}.$$

Proposition 1. 6: *If conditional Choquet expectations* $I(1_A/1_B)$ *are defined by: "A, B Ì S,*

(9) $\int_C [I(1_A/1_B) - 1_A] dn = 0,$

where $C \in \{B, B^C\}$, then:

(i) if $\int_C X\,dn = \int_S X.1_C dn$,

$$n\,(\mathrm{A/B}) = \frac{1 - n(\mathrm{A^C} \cup \mathrm{B^C})}{1 + n\,(\mathrm{A^C} \cap \mathrm{B}) - n\,(\mathrm{A^C} \cup \mathrm{B^C})}.$$

(ii) if $\int_C Xdn = \int_{-\infty}^{0} [n(\{X \geq x\} \cap C) - n(C)]dx + \int_0^{+\infty} n(\{X \geq x\} \cap C)dx,$

$$n\,(\mathrm{A/B}) = \frac{n(\mathrm{B}) - n(\mathrm{A^C} \cap \mathrm{B})}{n(\mathrm{B})}.$$

We can summarise the previous results in the tableau billow. We use the following notations:

(I) $\int_C 1_A dn = \int_C I(1_A/1_B) dn_{/s(B)}$

(II) $\int_C [1_A - I(1_A/1_B)] dn = 0$

(III) $\int_C [I(1_A/1_B) - 1_A] dn = 0$

(α) $\int_C X\,dn = \int_S X.1_C dn$

(β) $\int_C Xdn = \int_{-\infty}^{0} [n(\{X \geq x\} \cap C) - n(C)]dx + \int_0^{+\infty} n(\{X \geq x\} \cap C)dx$

Bayes' rule (Bayes): $n\,(\mathrm{A/B}) = \dfrac{n(\mathrm{A} \cap \mathrm{B})}{n(\mathrm{B})}$

Dempster-Shafer's rule (D-S): $n\,(\mathrm{A/B}) = \dfrac{n(\mathrm{A} \cup \mathrm{B^C}) - n\,(\mathrm{B^C})}{1 - n(\mathrm{B^C})}$

Full Bayesian Updating rule (FUBU):

$$n\,(\mathrm{A/B}) = \frac{n(\mathrm{A} \cap \mathrm{B})}{1 + n(\mathrm{A} \cap \mathrm{B}) - n(\mathrm{A} \cup \mathrm{B^C})}$$

FUBU on conjugate capacity[3] (FUBU/C):

[3] With $\bar{n}(\mathrm{A}) = 1 - n\,(\mathrm{A^C})$.

$$n(\mathrm{A/B}) = \frac{1 - n(\mathrm{A}^{\mathrm{C}} \cup \mathrm{B}^{\mathrm{C}})}{1 + n(\mathrm{A}^{\mathrm{C}} \cap \mathrm{B}) - n(\mathrm{A}^{\mathrm{C}} \cup \mathrm{B}^{\mathrm{C}})} = \frac{\bar{n}(\mathrm{A} \cap \mathrm{B})}{1 + \bar{n}(\mathrm{A} \cap \mathrm{B}) - \bar{n}(\mathrm{A} \cup \mathrm{B}^{\mathrm{C}})}$$

D-S on conjugate capacity (D-S/C):

$$n(\mathrm{A/B}) = \frac{n(\mathrm{B}) - n(\mathrm{A}^{\mathrm{C}} \cap \mathrm{B})}{n(\mathrm{B})} = \frac{\bar{n}(\mathrm{A} \cup \mathrm{B}^{\mathrm{C}}) - \bar{n}(\mathrm{B}^{\mathrm{C}})}{1 - \bar{n}(\mathrm{B}^{\mathrm{C}})}.$$

	(I)	(I)	(II)	(II)	(II)	(II)	(III)	(III)	(III)	(III)
			(α)	(α)	(β)	(β)	(α)	(α)	(β)	(β)
	C = B, B^{C}	C = S	C = B, B^{C}	C = S	C = B, B^{C}	C = S	C = B, B^{C}	C = S	C = B, B^{C}	C = S
1_A and 1_B como	Bayes	Bayes	FUBU	FUBU	Bayes	FUBU	FUBU /C	FUBU /C	D-S/C	FUBU /C
1_A and 1_B antimo	Bayes	D-S	FUBU	FUBU	Bayes	FUBU	FUBU /C	FUBU /C	D-S/C	FUBU /C
Gen case	Bayes		FUBU		Bayes		FUBU /C		D-S/C	

Obviously, such a tableau is informative but leaves us at loss as to which formula to apply in a valuation problem where information is taken into account. Taking information into account means that a decision is a process or a strategy where decisions are taken in accordance with information arrivals. In the next sections, time is explicitly represented in accordance with an information arrivals process.

2. Time discounting and conditioning valuation

We introduce the intuitions of our model and results through the classical setting of the Net Present Value (NPV). The extension to uncertainty of the classical NPV formula applied to a random payoff X_T at time T is:

$$NPV(X_T) = r(T)\, V(X_T).$$

This is easily interpreted: In the usual NPV setting, $V(X_T)$ is X_T's value in Euro at time T and $r(T)$ is a market interest rate. In our individual decision and valuation setting, V is an individual's valuation of an asset, i.e. $V(X_T)$ is X_T's certainty equivalent in Euro at time T and $r(T)$ is the individual's discount factor representing preferences for present consumption (wealth).

From now on, V will be a Choquet integral and this will be obtained from axioms on the agent's preferences in section 3.

Then $r(T)\ V(X_T)$ is the present value, i.e. in Euro today, of the certainty equivalent of X_T.

Note that, in the usual setting, the NPV formula requires no arbitrage on riskless bonds (the price of a marketed portfolio with non negative and at least one strictly positive payoffs is its strictly positive formation cost) in a frictionless market (i.e. borrowing and lending are priced the same). In an individual's valuation problem, time separability is obtained under a supplementary axiom on time preferences.

At a time $t < T$, let information Y_t arrive (Y_t is a random variable, e.g. an intermediary payoff at time t, or an index value) and let $NPV^{Yt}(X_T) = r^{Yt}(T)V^{Yt}(X_T)$ be the net value of X_T at time t (i.e. Euro at time t) with $r^{Yt}(T)$and $V^{Yt}(X_T)$ be the corresponding conditional discount factor and certainty equivalent, given the information. It is usually assumed that, seen from the present, $r^{Yt}(T) = r^t(T)$ is not random. In the literature on interest rates, this assumption is called "the rational expectation hypothesis", we shall justify this below.

With the previous interpretation in mind, a direct extension to NPV of conditional integral formulas given in the first section would be:

$$NPV(X_T) = NPV[NPV^{Yt}(X_T)] \qquad (1'),$$

$$NPV[X_T - NPV^{Yt}(X_T)] = 0 \qquad (2') \text{ and}$$

$$NPV[NPV^{Yt}(X_T) - X_T] = 0 \qquad (3').$$

Proposition 2.1: Definition (1') of conditional NPV implies:

$$r(T) = r(t)\ V[r^{Yt}(T)]\ .$$

Proof:

Formula (1') is: $r(T)\ V(X_T) = r(t)\ V[r^{Yt}(T)\ V^{Yt}(X_T)]$. If $X_T = x_T > 0$ is not random, then formula (1') becomes: $r(T)\ x_T = r(t)\ V[r^{Yt}(T)\ x_T]$ which is: $r(T) = r(t)\ V[r^{Yt}(T)]$.

Proposition 2.2: Under the assumption that the conditional expected rate is non random (rational expectation hypothesis):

$$\forall\, t < T,\ \forall\, Y_t, \quad r^{Yt}(T) = r^t(T) = \frac{r(T)}{r(t)},$$

definition (1') of conditional NPV holds if and only if conditional valuation satisfies:

$$\forall\, t < T, \quad V(X_T) = V[V^{Yt}(X_T)] \qquad \text{(CKL}^5\text{)}.$$

Proof: With a non random expected rate of return, formula (1') becomes:

$r(T)\ V(X_T) = r(t)\ r^t(T)\ V[V^{Yt}(X_T)]$ which implies :

$V(X_T) = V[V^{Yt}(X_T)]$ (CKL).

Remarks: We don't obtain the same consistency between formulas (2) or (3) in the introduction and the NPV ones (2') and (3'). Indeed, take (2'), for instance, we face a difficulty at first sight: the difference between a payoff in Euro at time T and a payoff at time t doesn't make sense as those two payoffs are not expressed in the same unit. Formula (2') could be reformulated consistently in Euro at time t:

(2'') $NPV[r^{Yt}(T)\ X_T - NPV^{Yt}(X_T)] = 0$, i.e.: $r(t)V[r^{Yt}(T)\ X_T - r^{Yt}(T)V^{Yt}(X_T)] = 0$,

or :

$$V[r^{Yt}(T)\ \{X_T - V^{Yt}(X_T)\}] = 0.$$

Equivalently, if $r^{Yt}(T)$ is non zero, equation (2), could be rewritten as:

(2''') $NPV[X_T - \frac{1}{r^{Yt}(T)}\ NPV^{Yt}(X_T)] = 0$, i.e.:

$$r(T)V[X_T - \frac{1}{r^{Yt}(T)}\ r^{Yt}(T)V^{Yt}(X_T)] = 0,$$

which collapses to formula (2).

The same applies to formula (3') and (3).

---

[5] This formula was proposed in Chateauneuf, Kast and Lapied (CKL) (2001) under an axiom called "time consistency" based on the intuition of the accountants' (past) time consistency : $r(T) = r(t)\ r^t(T)$.

However, none of these formulas say something about what conditions the conditional discount factors should satisfy. Indeed, if $X_T = x_T$ is a certain payoff as in proposition 2.2, all the formulas are tautological: 0 = 0!
These remarks give the intuition that formulas (2) and (3) cannot integrate information arriving along with time. Next section will show why more precisely.

3. Valuing uncertain future payoff flows

In this section, we extend the theory to cash flows valuation using CKL's definition (1) under the rational expectation hypothesis and show that this definition satisfies Dynamic Consistency.

Future, has, at least, two components: Times and States (uncertainty).
In this paper: $T=\{1, \ldots, T\}$ and $S=\{1, \ldots, S\}$ and let us add a present (certain) state where valuation is done (or when decisions are taken in accordance with this valuation).

3.1 Preferences and value functions

A decision maker has preferences over uncertain cash flows: $X: T \times S \to R$.
Preferences are represented by: $W: R^{T \times S} \to R$.
Note that: $X = \sum_S X(\,.\,, s) = \sum_T X(t,\,.\,)$ where $X(t,\,.\,)$ is the particular random cash flow:
$X(t,\,.\,) = (0, \ldots, 0, X_t, 0, \ldots, 0)$ and X_t is a random variable from S to R (and, similarly, $X(\,.\,, s)$ is a particular certain cash flow, i.e. a trajectory from T to R).

We assume the usual hierarchy between T and S in the sense that preferences can be decomposed according to three steps:

1- Preferences over certain cash flows are represented by : $P: R^T \to R$,
e.g. $P(c)=\sum_T r(t)c(t)$ according to axioms on preferences (Koopmans (1972)).

2- Preferences over uncertain payoffs are represented by: $V: R^S \to R$, e.g. $V(x) = \sum_S DX(s)n[X>X(s)]$ where $DX(s) = [X(s+1) - X(s)]$, $X(0) = 0$ and the s's are ordered according to the permutation on S which makes X non decreasing. n is a capacity on 2^S (Chateauneuf (1991)).

3- Preferences over uncertain cash flows which are represented by W, are such that:

$$W(X) = W[V(X_1), \ldots, V(X_T)].$$

(Under some assumptions on which we shall come back in the last section) .

Note that:

- If m is a lottery with payoffs contingent on the states in S, then capacity n could be interpreted as the result of a deformation function $n = f(m)$, (Yaari (1987)).
- More generally, we could consider $U(X(t,s))$ instead of $X(t,s)$ in equation (1) (Schmeidler (1989), Quiggin (1982) or Wakker (1989) if r or n are known).
- The model could be extended to non additive W (see, for instance, Gilboa (1989), Shalev (1997), De Waegenaere and Wakker (2001) or Chateauneuf and Rebille (2003)).
- r and/or n are not necessarily increasing (signed measures). See De Waegenaere and Wakker (2001) who give an interesting justification of negative discount factors. Furthermore, a signed n is a way to take transaction costs into account, see De Waegenare, Kast and Lapied (2003).

However, in the following, we shall keep to an additive W with discount factors (assuming Time Separability) and Choquet valuation for random variables (additive on comonotonic sets of random variables):

$$W(X) = \sum_{t=1}^{T} r(t)V(X_t).$$

3.2 Information

Information arrives along with time: $\forall t \in T\ Y_t : (S, 2^S) \rightarrow (I, 2^I)$, with $I \subset R$ and Y_t is such that: $Y_t^{-1}(2^I) \subset 2^S$.

Furthermore, we assume that $(Y_t^{-1}(2^I))_{t=1\ldots T}$ is a filtration. We assume also that:

$\forall X$, $\forall t \leq t$ $X_t^{-1}(2^I) \subset Y_t^{-1}(2^I)$. For instance, information could be given by the X_t's themselves[6].

Dynamical decision making takes future information arrivals into account, i.e. present preferences are consistent with future preferences conditional on information that is then available. If information arrives at time t, such preferences will yield "conditional valuations" of the future uncertain payoffs of X as from time t, i.e the uncertain cash flow $X^{t+} = (X_{t+1}, \dots , X_T)$:

$$W^{Yt}(X^{t+}) = W^{Yt}(X_{t+1}, \dots , X_T) : I \rightarrow R.$$

With: $W^{Yt}(X^{t+}) = \sum_{t=t+1}^{T} r^{Yt}(t)\, V^{Yt}(X_t)$ and $V^{Yt}(X_t) = \sum_{S} DX_t(s)\, n^{Yt}[X>X(s)]$ where n is a "conditional capacitiy" … to be defined.

Time consistency can only be understood in a decision making process, i.e. a strategy according to which decisions are adapted to information arrivals. In order to apply Bellman's principle of dynamic programming, Kreps and Porteus (1978) introduced time consistency in the decision process with known probability distributions. A decision d_t is taken at time t and modifies the probability distribution over future outcomes. Hence, the valuation is the expected utility (here, payoff expectation) with respect to the distribution determined by d_t. Let us note $V^{Yt}(d_t, X_t)$ the decision criterion, i.e. the expected value of the cash flow X_t, $t > t$, if decision d_t is taken with information Y_t. Then Kreps and Porteus time consistency condition for the decision criterion can be written as:

$\forall t \in T, \forall t < t', \forall Y_t, \forall Y_{t'}$,

$$\sum_{t=t+1}^{T} r^{Yt}(t)\, V^{Yt}(d_t, X_t) =$$

$$\sum_{t=t+1}^{t'-1} r^{Yt}(t)\, V^{Yt}(d_t, X_t) + r^{Yt}(t')\, V^{Yt}[d_t, X_{t'} + Max_{dt'} \sum_{t=t'+1}^{T} r^{Yt'}(t)\, V^{Yt'}(d_{t'}, X_t)].$$

Let assume decision d_0 is taken at the initial time in order to simplify the notations, we have Kreps and Porteus Time Consistency (TC):

[6] These assumptions were those of CKL and the interpretation of their condioning rules is founded on comonotonicity of information and uncertain payoffs.

$\forall t \in T, \forall Y_t$,

$$\sum_{t=1}^{T} r(t)\, V(d_0, X_t) = \sum_{t=1}^{t-1} r(t)\, V(d_0, X_t) + r(t)\, V[d_0, X_t + Max_{dt} \sum_{t=t+1}^{T} r^{Yt}(t)\, V^{Yt}(d_t, X_t)] \quad (KP).$$

In our pure valuation context, decisions do not appear explicitly. Valuation bears upon the cash flow and is meant to represent preferences. In such a context, (KP) can be interpreted the following way:

Let X be a cash flow and X' an other one which differs from X only from time t on under information $\{Y_t \in B\}$:

$\forall t \leq t \ \forall s \in S,\ X_t(s) = X'_t(s)$ and $\forall t > t \ \forall s \in \{Y_t \in B^c\},\ X_t(s) = X'_t(s)$. If X is optimal, then $W(X') \leq W(X)$ and (KP) implies that, given information Y_t , we have : $W^{Yt}(X_t^{'+}) \leq W^{Yt}(X_t^+)$. Reciprocally, if $W^{Yt}(X_t^{'+}) \leq W^{Yt}(X_t^+)$ then: $W(X') \leq W(X)$, otherwise X wouldn't be optimal.

This interpretation is the one taken as a definition by Karni and Schmeidler (1991) as well as by Sarin and Wakker (1998) for Dynamic Consistency (DC), expressed in terms of valuation instead of preferences:

Definition (Dynamic Consistency):

$\forall t \in T, \forall Y_t$, $\forall B \in I\ \forall X,\ X': T \times S \rightarrow R$ such that $\forall t \leq t\ \ \forall s \in S,\ X_t(s) = X'_t(s)$ and $\forall t > t\ \forall s \in \{Y_t \in B^c\},\ X_t(s) = X'_t(s)$:

$$W(X') \leq W(X) \Leftrightarrow W^{\{Y_t \in B\}}(X_t^{'+}) \leq W^{\{Y_t \in B\}}(X_t^{+}) \qquad \text{(DC)}.$$

Notice that this equivalence makes no reference to the axioms necessary for the preferences' representation, when in (KP) valuation was the discounted expected payoffs (or expected utility). In particular, (DC) is consistent with our assumptions about W.

More directly, (KP) Time Consistency formula expressed in terms of cash flow valuation becomes:

$$\forall t \in T, \forall Y_t : \sum_{t=1}^{T} r(t)\, V(d_0, X_t) = \sum_{t=1}^{t-1} r(t)\, V(d_0, X_t) + r(t)\, V[d_0, X_t + \sum_{t=t+1}^{T} r^{Yt}(t)\, V^{Yt}(d_t, X_t)].$$

With our decomposition for W in terms of discount payoffs and Choquet integrals, we obtain a definition of Time Consistency similar to what was called dynamic consistency in Chateauneuf, Kast and Lapied (2001):

Definition (Time Consistency):

$$\forall t \in T, \forall Y_t \quad W(X) = W[X_1, \dots, X_t + W^{Y_t}(0, \dots, 0, X_{t+1}, \dots, X_T), 0, \dots, 0] \qquad \text{(TC)}.$$

Notice that both (TC) and (DC) are derived from (KP) in a pure valuation context and under time separability. Furthermore, they obviously imply (KP) if preferences satisfy the assumptions made in 3.1.

However, in (KP), conditional expectations are well defined by probability theory when, in a general preference representation model, conditional valuation has yet to be defined. This is this paper's goal, indeed, and is worked on in next sub-section.

3.3 Conditioning uncertain payoff valuation

In this sub-section, we concentrate on the particular case where $X = (0, \dots, 0, X_T) = (0, X_T)$, and we assume that preferences satisfy Model Consistency: separability in time and Choquet integral over payoffs at each information date. This yields:

$$\forall t \in T, \forall Y_t \quad W^{Y_t}(X) = r^{Y_t}(T) V^{Y_t}(X_T) = r^{Y_t}(T) \sum_S DX_T(s)\, n^{Y_t}(s).$$

It is no wonder then, if the intuitions we put forward in section 2 can be easily proved in this setting.

Indeed, Time Consistency: $W(X) = W[(0, W^{Y_t}(0, X_T), 0)]$, becomes:

$r(T)\, V(X_T) = r(t)\, V[r^{Y_t}(T)\, V^{Y_t}(X_T)]$.

Then, the following proposition and corollaries are straightforward.

Proposition 3.1: If $\forall t \in T$, $X_t = x_t$ is non random, then under (TC):

$$\forall Y_t, \quad r(T) = r(t)\, V[r^{Y_t}(T)].$$

Note that proposition 3.3 yields an (implicit) definition of $r^{Y_t}(T)$

Corollary 3.1.1: If, $\forall t < T$, $\forall Y_t$, $r^{Y_t}(T) = r^t(T)$ is non random, then, under (TC):

$\forall t < T$, $r(T) = r(t)\, r^t(T)$ (accountants' time consistency).

Corollary 3.1.2: If $\forall t < T$, $\forall Y_t$, $r^{Y_t}(T) = r^t(T)$ is non random, then (TC) obtains if and only if: $V(X_T) = V[V^{Y_t}(X_T)]$ (CKL).

More generally, assume that:

$$\forall t < T, \forall Y_t, \quad V[r^{Y_t}(T)\, V^{Y_t}(X_T)] = V[r^{Y_t}(T)]\, V[V^{Y_t}(X_T)],$$

i.e., in the case where n is a probability, $r^{Yt}(T)$ and $V^{Yt}(X_T)$ are "*n-independent*" (the implications of such an assertion has still to be found in terms of capacities[7]!). In accordance with our assumption on the hierarchy between time and uncertainty, this formula can be interpreted as a discrepancy between the treatment of information on time and on uncertainty. For instance, an information could contain two "independent variables", one affecting preferences over present consumption (or one relative to the bond market) and one affecting preferences over uncertain payoffs (or one about the stock market trends).

Corollary 3.1.3: If $\forall t<T$, $\forall Y_t$, $V[r^{Yt}(T)\, V^{Yt}(X_T)] = V[r^{Yt}(T)]V[V^{Yt}(X_T)]$, then (TC) obtains if and only if (CKL) is satisfied.

In accordance with the intuition obtained within the classical NPV model in section 2, the two last corollaries disqualify alternative definitions of conditional Choquet integrals when information arrivals and discounting are taken into account. The two sided updating rule we introduced in section one is then derived from the necessary conditions of our dynamic decision model.

However, many authors have claimed that non expected models are dynamically inconsistent: Epstein and Le Breton (1993), Border and Segal (1994), notably. However other authors have moderated this claim by analysing the many axioms or particular conditions on models (often implicit) that may interfere: Karni and Schmeidler (1991), Jaffray (1994), Machina (1998), Sarin and Wakker (1998) among the more prominent ones. Our model satisfies Time Consistency and Model Consistency (sequential consistency in Sarin and Wakker (1998), i.e. here: both W and W^{Yt} are decomposed into a Choquet integral and a linear discount factor). We know from Sarin and Wakker, for instance, that if we add Consequentialism, to the previous two conditions, only the multipriors model fits. This would require us to introduced pessimism (in the sense of Schmeidler (1989)) instead of comonotone additivity. In fact, the problem is simpler: our model doesn't satisfy Consequentialism.

Definition (Consequentialism): *Let B be a set in* 2^S *and X, X', Y, Y' be random payoffs such that* $\forall s \in B^c$, $X(s) = X'(s)$ *and* $Y(s) = Y'(s)$ *and* $\forall s \in B$ $X(s) = Y(s)$ *and* $X'(s) = Y'(s)$ *then:*

$X \succsim_B X' \Leftrightarrow Y \succsim_B Y'$.

And we have:

[7] Mathematically, this may introduce a condition on capacities similar to $m^B(A)=m(A)$ for a probability.

Proposition 3.2:

(CKL) violates Consequentialism.

The proof, in appendix 2, is a counter example in which: $X \prec_B X'$ and $Y \succ_B Y'$.

The intuition behind this result justifies our updating rule and relies on the dependence between the information and the payoff variables. When both variables are comonotonic (hence are positively correlated for any additive probability distribution, Chateauneuf et al. (1994)) information and payoffs are dependent and then the information (the set *B*, in section 1) is enforced, hence Bayes rule applies. Indeed Bayes rule measures the set on which payoffs obtain (*A*) and the information set (*B*) coincide. When they are antimonotonic, hence negatively correlated, then the information is contrary to payoffs, it's the set B^C which is enforced and Dempster-Schaffer rule applies: it measures the set on which payoffs obtain when the contrary of information is true.

More precisely, our updating rule depends on the ranking of payoffs and the ranking of information values, including payoffs which are not concerned by information. This is the reason why Consequentialism may not be satisfied.

3.4. Applying time consistent valuation to a real investment problem

Mining is often referred to as an example of real investment valuation, (e.g., Pindyck and Dixit (1994)). Assume in a very simplified setting that production costs are known and constant: *c*. The future ore price is uncertain and we assume we are looking two periods ahead (two years, say), with ore prices going up (*u*) or down (*d*) each year. Thus, at cost *c*, the extracted ore value in one year time is P_u or P_d and in two years time: P_{uu}, P_{ud}, P_{du} *or* P_{dd}. If the price in period 1 is too low, i.e. if the information is *d* (for going down) it may not be worth producing anymore if, for instance, $P_{dd} - c < 0$. This is a "real option" (the option not to produce if the ore price is too low) and its value adds to the investment value in a dynamic setting. The option value is the difference between a dynamic valuation and static a one.

Following our notations, let $X = (X_1, X_2)$ be the investment, its uncertain payoffs at time T =2 are assumed to be such that: $P_{uu} - c > P_{ud} - c > P_{du} - c > 0 > P_{dd} - c$. The investor's preferences are represented by a Choquet expectation with respect to a capacity *n* at time 0 (present) and by a Choquet expectation with respect to capacity n^{Y1} where information at time 1 is "up": $Y^1 = 1_B$ with $B = \{uu, ud\}$ and $B^c = \{du, dd\}$. We have:

$$V^{[1B=1]}(X_2) = (P_{dd} - c)n^B(X_2 > P_{dd}) + (P_{ud} - P_{du})\, n^B(X_2 > P_{du}) + (P_{uu} - P_{ud})\, n^B(X_2 > P_{ud})$$

$$= (P_{du} - c)n^B(A_1) + (P_{ud} - P_{du})\, n^B(A_2) + (P_{uu} - P_{ud})\, n^B(A_3)$$

with $A_3 \subset A_2 = B \subset A_1$.

Hence, from section 1, we have:

$n^B(A_1) = n^B(A_2) = \frac{n(B)}{n(B)} = 1$ and

$$V^{[1B=1]}(X_2) = P_{ud} - c + (P_{uu} - P_{ud})\, \frac{n(A_3)}{n(B)}.$$

Similarly, given that the mine will not be exploited in the P_{dd} case, we have:

$V^{[1B=0]}(X_2) = (P_{du} - c)\frac{n(A_1) - n(B)}{1 - n(B)} + 0 + 0$ because the two last capacities are $\frac{n(B) - n(B)}{1 - n(B)} = 0$.

It is easy to see that $V^{[1B=1]}(X_2) = V^{[1B=0]}(X_2)$ and then:

$$V[V^{1B}(X_2)] = V^{[1B=0]}(X_2)(1 - n(B)) + V^{[1B=1]}(X_2)\, n(B) = V(X_2).$$

This yields the present expected value of the investment, i.e. its cash flow, including the option not to exploit the mine. This means that $V^{Y1}(X_2)$ is, in fact, the value of the optimal investment given information Y_1. It is also equal to the net present value of the optimal cash flow (under the assumption that the discount factors do not depend on information), in accordance with Kreps and Porteus definition of dynamic consistency:

$$W(X) = W[X_1 + W^{Y1}(X_2)] = r(1)\{V(X_1) + r^1(2)\, V[V^{Y1}(X_2)]\} = r(1)\, V(X_1) + r(2)\, V(X_2),$$

and then $V[V^{1B}(X_2)] = V(X_2)$, as in the previous equation.

If we take another conditioning rule, the last equality is not satisfied anymore, in general. For instance, with FUBU rule, as in Denneberg (1994), we have:

$$V^{[1B=1]}(X_2) = P_{ud} - c + (P_{uu} - P_{ud})\, \frac{n(A_3)}{1 + n(A_3) - n(\{uu, du, dd\})},$$

$$V^{[1B=0]}(X_2) = (P_{du} - c)\frac{n(\{du\})}{1 + n(\{du\}) - n(A_1)},$$

$$V[V^{1B}(X_2)] = (P_{du} - c)\frac{n(\{du\}) + n(B)[1 - n(A_1)]}{1 + n(\{du\}) - n(A_1)} + (P_{ud} - P_{du})\, n(A_2)$$

$$+ (P_{uu} - P_{ud})\, \frac{n(A_3)n(B)}{1 + n(A_3) - n(\{uu, du, dd\})}.$$

In general, the last term differs from the first side of the dynamic consistency condition:

$V(X_2) = (P_{du} - c)\, n(A_1) + (P_{ud} - P_{du})\, n(A_2) + (P_{uu} - P_{ud})\, n(A_3)$.

Hence, Dynamic Consistency is not satisfied is general.

4. Conclusion and possible extensions

Let us recall the Dynamic Consistency condition for cash flows:

$\forall t \in T,\ \forall Y_t,\quad W(X) = W[X_1, \dots, X_t + W^{Y_t}(0, X_{t+1}, \dots, X_T), 0]$, or under the assumption of time additivity on W:

$$\forall t \in T,\ \forall Y_t,\quad \sum_{t=1}^{T} r(t)\, V(X_t) = \sum_{t=1}^{t-1} r(t)\, V(X_t) + r(t) V[X_t + \sum_{t=t+1}^{T} r^{Y_t}(t)\, V^{Y_t}(X_t)].$$

The formula cannot be simplified because V is generally not additive. This would lead to the conclusion that, although we have a dynamically consistent valuation, we couldn't use it in a practical way in a valuation problem.

If we want to stay close to practice, however, note from the tableau in section 1 that we only know how to calculate the conditional capacities in two cases: the case where information is comonotonic with future payoffs and the case where it is antimonotonic. Given that Choquet integrals are additive on comonotonic sets of random variables, the dynamically consistent valuation of a cash flow can be computed for cash flows which only consist of comonotonic or antimonotonic payoffs at each dates, with information arrivals which are comonotononic or antimonotonic with the payoffs as in the example in section 3.4. Obviously, such a condition seems very restrictive. However, in many practical problems, this restriction is satisfied or acceptable as an approximation (e.g as the valuation of a super hedge, see Dhaene et al. (2002)a and (2002)b). Indeed, take the case of an investment similar to that of example 3.4, and assume uncertainty is described by binomial lattices (or Bernoullian, if arrows don't meet in the lattice). This representation of uncertainty is merely a discrete version of the usual assumption made in the real investment models where uncertainty is defined by a Brownian motion and binomial lattices are used to approximate non analytical solutions, in practice. In such models, payoffs increase or decrease at each period as do information arrivals, so that valuation arrivals can be computed at each date by backward induction.

Nevertheless, the limitation of the formula handiness, added to the non realistic separability assumption on time preferences, are incentives to enrich our representation model for preferences over future uncertain cash flows. Let us mention two extensions of the hierarchical model which have been investigated:

- We mentioned in the introduction that time separability can be replaced by a comonotonic additivity axiom which takes into account hedging opportunities for cash flows which do not vary in the same sense (Gilboa (1989), Shalev (1997), De Waegenaere and Wakker (2001) and Chateauneuf and Rebille (2003)).
- In a an other type of model, Heal (1998) and Chichilnisky (1996) proposed a criterion mixing a discount factor and an absolute value for the far future. This criterion was introduced to minimise the vanishing effect of discount factors on future payoffs, an effect incompatible with sustainable development (see Chichilnisky and Heal eds. (1998) for a an overview of the related literature).

Even if the hierarchy between time and uncertainty is maintained for valuation, none of the models introduce both non additivity with respect to time AND with respect to uncertainty, even in the case of uncertain payoffs.

For uncertain cash flows, the conclusion of our model indicates that no tractable solutions can be hoped for if the hierarchy is maintained, even in the case where discount factors are not random. Our next goal in this trend of research is to tackle the problem of valuing the future as a whole. In the simple setting of our model, this amounts to have preferences over two component payoffs. A more ambitious goal is to forget about the decomposition of the future into these two components. There may be more components (e.g. time, random time, uncertain payoffs and random payoffs) or no components at all in a subjective vision of states of the future world (in the sense of Savage's Big World).

References:

- Border K., Segal U. (1994): « Dynamic consistency implies expected utility (almost) », *Journal of Economic Theory*, 63, 170-188.
- Chateauneuf A. (1991): « On the use of capacities in modelling uncertainty aversion and risk aversion », *Journal of Mathematical Economics*, 20, 349-369.
- Chateauneuf A., Kast R., Lapied A. (1994): «Market preferences revealed by prices : Non linear pricing in slack markets », in Machina M. and Munier B. eds. *Models and experiments in risk and rationality*, Kluwer Academic Publisher.

- Chateauneuf A., Kast R., Lapied A. (2001): « Conditioning capacacities and Choquet integrals: the role of comonotony », *Theory and Decision*, 51, 367-386.
- Chateauneuf A., Rebille Y. (2003): « Some characterization of non additive multiperiod models », *W.P.* CERMSEM, Universtiy of Paris 1.
- Chichilninsky G. (1996): « What is sustainable development », *Social Choice and Welfare*, 13, 2, 231-257.
- Chichilninsky G., Heal G. eds. (1998): *Sustainabilty:Dynamics and Uncertainty*, Kluwer Academic Publishers.
- Cohen M., Gilboa I., Jaffray J-Y., Schmeidler D. (1997): « An experimental study of updating ambiguous beliefs », W.P. CEME, University of Paris 1.
- Dempster A. (1967): « Upper and lower probabilities induced by a multivalued mapping », *Annals of Mathematical Statistics*, 38, 325-339.
- Denneberg D. (1994): « Conditioning (updating) non additive measures », *Annals of Operations research*, 52, 21-42.
- De Waegenaere A., Kast R., Lapied A. (2003): « Choquet pricing equilibrium », *Insurance: Mathematics and Economics*, 32/3, 359-370.
- De Waegenare A., Wakker P. (2001): « Non monotonic Choquet integrals », *Journal of Mathematical Economics*, 36, 45-60.
- Dhaene J., Denuit M., Goovaerts M.J., Kaas R., Vyncke D. (2002)a: « The concept of comonotonicity in actuarial science and finance: theory », *Insurance: Mathematics and Economics*, 31, 3-33.
- Dhaene J., Denuit M., Goovaerts M.J., Kaas R., Vyncke D. (2002)b: « The concept of comonotonicity in actuarial science and finance: applications », *Insurance: Mathematics and Economics*, 31, 133-161.
- Epstein L., Le Breton M. (1993) : « Dynamically consistent beliefs must be Bayesian », *Journal of Economic Theory*, 61, 1-22.
- Gilboa I. (1989): « Expectations and variations in multi-period decisions », *Econometrica*, 57, 1153-59.
- Gilboa I., Schmeidler D. (1993): « Updating ambiguous beliefs », *Journal of Econonomic Theory*, 59, 33-49.
- Heal G. (1998): *Valuing the future*, Columbia University Press.
- Jaffray J-Y. (1994): « Dynamic decision making with belief functions », in Yager R. et al., eds., *Advances in the Dempster-Schaffer Theory of Evidence*, 331-352.

- Karni E., Schmeidler D. (1991): « Atemporal Dynamic Consistency and Expected Utility Theory », *Journal of Economic Theory*, 54, 401-408.
- Koopmans T. (1972): « Representation of preference orderings over time », in Mac Guire C. and Radner R. eds., *Decision and organization*, North Holland, 79-100.
- Kreps D., Porteus E. (1978) : « Temporal resolution of uncertainty and dynamic choice theory », *Econometrica*, 46, 1, 185-200.
- Machina M. (1998): « Dynamic consistency and non expected utility models of choice under uncertainty », *Journal of Economic Literature*, 22, 1622-1668.
- Pindyck A., Dixit R. (1994): *Investments under uncertainty*, Princeton University Press.
- Quiggin J. (1982): « A theory of anticipated utility », *Journal of Economic Behavior and Organisations*, 3, 4, 323-342.
- Sarin R., Wakker P.(1998): « Dynamic choice and NonExpected Utility », *Journal of Risk and Uncertainty*, 17, 87-119.
- Schmeidler D. (1986): « Integral representation without additivity », *Proceedings of the American Mathematical Society*, 97, 2, 255-261.
- Schmeidler D. (1989): « Subjective probability and expected utility without additivity », *Econometrica* , 57, 3, pp.571-587.
- Shafer G. (1967): *A mathematical theory of evidence*, Princeton University Press.
- Shalev J. (1997): « Loss aversion in a multiperiod model », *Mathematical Social Sciences*, 33, 203-226.
- Wakker P. (1989): « Continuous subjective expected utility with non additive probabilities », *Journal of Mathematical Economics*, 18, 1-27.
- Yaari M. (1987): « The dual theory of choice under risk », *Econometrica*,55, 95-115.

Appendix 1

Proof of proposition 1.1: By definition $n(A) = \int_S 1_A dn$.

(i) Comonotonicity of 1_A and 1_B implies A Ì B or B Ì A. For conditioning, consider the non trivial case where A Ì B.

$s \in B^C \Rightarrow s \in A^C \Rightarrow 1_A = 0$,

$s \in B^C \Rightarrow I(1_A / 1_B) = I(1_A / 1_B = 0)$,

then: $s \in B^C \Rightarrow I(1_A / 1_B) = I(1_A / 1_B = 0) = 0$.

$s \in B \Rightarrow I(1_A/1_B) = I(1_A/1_B = 1) \geq 0 = I(1_A/1_B = 0)$.

It follows that $\int_S I(1_A/1_B)dn_{/S(B)} = I(1_A/1_B = 1)n(B)$ and relation (1) implies:

$$\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 1) = \frac{n(\mathrm{A})}{n(\mathrm{B})} = \frac{n(\mathrm{A} \cap \mathrm{B})}{n(\mathrm{B})}.$$

(ii) Antimonotonicity of 1_A and 1_B implies A^C Ì B or B Ì A^C. For conditioning, consider the non trivial case where A^C Ì B.

$s \in B^C \Rightarrow s \in A \Rightarrow 1_A = 1$,

$s \in B^C \Rightarrow I(1_A/1_B) = I(1_A/1_B = 0)$,

then: $s \in B^C \Rightarrow I(1_A/1_B) = I(1_A/1_B = 0) = \int_{B^C} dn_{B^C} = 1$.

$s \in B \Rightarrow I(1_A/1_B) = I(1_A/1_B = 1) \leq 1 = I(1_A/1_B = 0)$.

It follows that $\int_S I(1_A/1_B)dn_{/S(B)} = I(1_A/1_B = 1) + [1 - I(1_A/1_B = 1)n(B^C)$ and relation (1) implies:

$$\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 1) = \frac{n(\mathrm{A}) - n(\mathrm{B}^\mathrm{C})}{1 - n(\mathrm{B}^\mathrm{C})} = \frac{n(\mathrm{A} \cup \mathrm{B}^\mathrm{C}) - n(\mathrm{B}^\mathrm{C})}{1 - n(\mathrm{B}^\mathrm{C})},$$ **QED**.

Proof of proposition 1.2: We have two candidates for $\int_C X dn$:

(3) $\int_C X dn = \int_S X 1_C dn$, and

(4) $\int_C X dn = \int_{-\infty}^{0} [n(\{X \geq x\} \cap C) - n(C)]dx + \int_0^{+\infty} n(\{X \geq x\} \cap C)dx$.

First, consider definition (3).

$\int_C 1_A dn = \int_S 1_A 1_C\, dn = \int_S 1_{A \cap C} dn = n(A \cap C)$,

$\int_C I(1_A/1_B)dn_{/S(B)} = \int_S I(1_A/1_B)1_C dn$.

If $C = B$:

$s \in B^C \Rightarrow I(1_A/1_B)1_B = 0$,

$s \in B \Rightarrow I(1_A/1_B)1_B = I(1_A/1_B = 1) \geq 0$.

then: $\int_B I(1_A/1_B)dn_{/S(B)} = I(1_A/1_B = 1)n(B)$, and (2) implies:

$$\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 1) = \frac{n(\mathrm{A} \cap \mathrm{B})}{n(\mathrm{B})}.$$

If $C = B^C$:

$s \in B \Rightarrow I(1_A / 1_B) 1_{B^C} = 0$,

$s \in B^C \Rightarrow I(1_A / 1_B) 1_{B^C} = I(1_A / 1_B = 0) \geq 0$.

then: $\int_{B^C} I(1_A / 1_B) dn_{/S(B)} = I(1_A / 1_B = 0) n(B^C)$, and (2) implies:

$$\mathrm{I}(1_\mathrm{A} / 1_\mathrm{B} = 0) = \frac{n(\mathrm{A} \cap \mathrm{B}^\mathrm{C})}{n(\mathrm{B}^\mathrm{C})}.$$

Now, consider definition (4).

$$\int_C 1_A dn = \int_0^1 n(A \cap C) dx = n(A \cap C),$$

If $I(1_A / 1_B = 0) \leq I(1_A / 1_B = 1)$,

$$\int_C I(1_A / 1_B) dn = \int_0^{I(1_A / 1_B = 0)} n(C) dx + \int_{I(1_A / 1_B = 0)}^{I(1_A / 1_B = 1)} n(B \cap C) dx$$

$$= I(1_A / 1_B = 0)[n(C) - n(B \cap C)] + I(1_A / 1_B = 1) n(B \cap C).$$

With $C = B$ relation (2) yields: $\mathrm{I}(1_\mathrm{A} / 1_\mathrm{B} = 1) = \dfrac{n(\mathrm{A} \cap \mathrm{B})}{n(\mathrm{B})}$,

and with $C = B^C$ we have: $\mathrm{I}(1_\mathrm{A} / 1_\mathrm{B} = 0) = \dfrac{n(\mathrm{A} \cap \mathrm{B}^\mathrm{C})}{n(\mathrm{B}^\mathrm{C})}$.

If $I(1_A / 1_B = 1) \leq I(1_A / 1_B = 0)$,

$$\int_C I(1_A / 1_B) dn = I(1_A / 1_B = 1)[n(C) - n(B^C \cap C)] + I(1_A / 1_B = 0) n(B^C \cap C).$$

Relation (2) yields the same results, **QED**.

Proof of proposition 1.3: First, suppose that 1_A and 1_B are comonotonic random variables, then A Ì B or B Ì A. We consider the non trivial case where A Ì B.

$s \in B^C \Rightarrow s \in A^C \Rightarrow 1_A = 0$,

$s \in B^C \Rightarrow I(1_A / 1_B) = I(1_A / 1_B = 0)$,

then: $s \in B^C \Rightarrow 1_A - I(1_A / 1_B) = 0$.

$s \in B \Rightarrow I(1_A / 1_B) = I(1_A / 1_B = 1)$,

$s \in A \Rightarrow 1_A - I(1_A / 1_B) = 1 - I(1_A / 1_B = 1) \geq 0$,

$s \in A^C \cap B \Rightarrow 1_A - I(1_A / 1_B) = -I(1_A / 1_B = 1) \leq 0$.

It follows that

$$\int_S [1_A - I(1_A/1_B)]dn = -I(1_A/1_B = 1) + I(1_A/1_B = 1)n(A \cup B^C) + [1 - I(1_A/1_B = 1)]n(A),$$

and relation (5) implies:

$$\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 1) = \frac{n(\mathrm{A})}{1 + n(\mathrm{A}) - n(\mathrm{A} \cup \mathrm{B}^\mathrm{C})} = \frac{n(\mathrm{A} \cap \mathrm{B})}{1 + n(\mathrm{A} \cap \mathrm{B}) - n(\mathrm{A} \cup \mathrm{B}^\mathrm{C})}.$$

Now, suppose that 1_A and 1_B are antimonotonic random variables, then $A^C \subset B$ or $B \subset A^C$. We consider the non trivial case where $A^C \subset B$.

$s \in B^C \Rightarrow s \in A \Rightarrow 1_A = 1,$

$s \in B^C \Rightarrow I(1_A/1_B) = I(1_A/1_B = 0),$

then: $s \in B^C \Rightarrow 1_A - I(1_A/1_B) = 1 - \int_{B^C} dn_{B^C} = 0.$

$s \in B \Rightarrow I(1_A/1_B) = I(1_A/1_B = 1),$

$s \in A^C \Rightarrow 1_A - I(1_A/1_B) = -I(1_A/1_B = 1) \le 0,$

$s \in A \cap B \Rightarrow 1_A - I(1_A/1_B) = 1 - I(1_A/1_B = 1) \ge 0.$

It follows that

$$\int_S [1_A - I(1_A/1_B)dn = -I(1_A/1_B = 1) + I(1_A/1_B = 1)n(A) + [1 - I(1_A/1_B = 1)]n(A \cap B),$$

and relation (5) implies:

$$\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 1) = \frac{n(\mathrm{A} \cap \mathrm{B})}{1 + n(\mathrm{A} \cap \mathrm{B}) - n(\mathrm{A} \cup \mathrm{B}^\mathrm{C})}$$, **QED**.

Proof of proposition 1.4: (i) The case where $C = B$ is Denneberg (1994) Proposition 2.2. Let us consider the case where $C = B^C$.

$s \in B \Rightarrow [1_A - I(1_A/1_B)]1_{B^C} = 0,$

$s \in A \cap B^C \Rightarrow [1_A - I(1_A/1_B)]1_{B^C} = 1 - I(1_A/1_B = 0) \ge 0,$

$s \in A^C \cap B^C \Rightarrow [1_A - I(1_A/1_B)]1_{B^C} = -I(1_A/1_B = 0) \le 0,$

then:

$$\int_{B^C} [1_A - I(1_A/1_B)]dn = -I(1_A/1_B = 0) + I(1_A/1_B = 0)n(A \cup B) + [1 - I(1_A/1_B = 0)]n(A \cap B^C),$$

and (6) implies:

$$\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 0) = \frac{n(\mathrm{A} \cap \mathrm{B}^\mathrm{C})}{1 + n(\mathrm{A} \cap \mathrm{B}^\mathrm{C}) - n(\mathrm{A} \cup \mathrm{B})}.$$

(ii) We have:

$s \in A \cap B \Rightarrow 1_A - I(1_A/1_B) = 1 - I(1_A/1_B = 1) \geq 0$,

$s \in A \cap B^C \Rightarrow 1_A - I(1_A/1_B) = 1 - I(1_A/1_B = 0) \geq 0$,

$s \in A^C \cap B \Rightarrow 1_A - I(1_A/1_B) = -I(1_A/1_B = 1) \leq 0$,

$s \in A^C \cap B^C \Rightarrow 1_A - I(1_A/1_B) = -I(1_A/1_B = 0) \leq 0$.

If $I(1_A/1_B = 0) \leq I(1_A/1_B = 1)$,

$$\int_C [1_A - I(1_A/1_B)]dn = \int_{-I(1_A/1_B=1)}^{-I(1_A/1_B=0)} \{n[(A \cup B^C) \cap C] - n(C)\}dx + \int_{-I(1_A/1_B=0)}^{0} [n(A \cap C) - n(C)]dx$$

$$+ \int_0^{1-I(1_A/1_B=1)} n(A \cap C)dx + \int_{1-I(1_A/1_B=1)}^{1-I(1_A/1_B=0)} n(A \cap B^C \cap C)dx.$$

With $C = B$ relation (6) gives: $\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 1) = \dfrac{n(\mathrm{A} \cap \mathrm{B})}{n(\mathrm{B})}$,

and with $C = B^C$, we have: $\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 0) = \dfrac{n(\mathrm{A} \cap \mathrm{B}^\mathrm{C})}{n(\mathrm{B}^\mathrm{C})}$.

If $I(1_A/1_B = 1) \leq I(1_A/1_B = 0)$,

$$\int_C [1_A - I(1_A/1_B)]dn = \int_{-I(1_A/1_B=0)}^{-I(1_A/1_B=1)} \{n[(A \cup B) \cap C] - n(C)\}dx + \int_{-I(1_A/1_B=1)}^{0} [n(A \cap C) - n(C)]dx$$

$$+ \int_0^{1-I(1_A/1_B=0)} n(A \cap C)dx + \int_{1-I(1_A/1_B=0)}^{1-I(1_A/1_B=1)} n(A \cap B \cap C)dx.$$

Relation (6) yields the same results, **QED**.

Proof of proposition 1.5: With the same method as in the proof of proposition 3, we have:

- when 1_A and 1_B are comonotonic random variables (A Ì B),

$s \in B^C \Rightarrow I(1_A/1_B) - 1_A = 0$,

$s \in A \cap B \Rightarrow I(1_A/1_B) - 1_A = I(1_A/1_B = 1) - 1 \leq 0$,

$s \in A^C \cap B \Rightarrow I(1_A/1_B) - 1_A = I(1_A/1_B = 1) \geq 0$,

then:

$$\int_S [I(1_A/1_B) - 1_A]dn = I(1_A/1_B = 1) - 1 + [1 - I(1_A/1_B = 1)]n(A^C) + I(1_A/1_B = 1)n(A^C \cap B),$$

and relation (7) implies:

$$\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 1) = \frac{1 - n(\mathrm{A}^\mathrm{C})}{1 + n(\mathrm{A}^\mathrm{C} \cap \mathrm{B}) - n(\mathrm{A}^\mathrm{C})} = \frac{1 - n(\mathrm{A}^\mathrm{C} \cup \mathrm{B}^\mathrm{C})}{1 + n(\mathrm{A}^\mathrm{C} \cap \mathrm{B}) - n(\mathrm{A}^\mathrm{C} \cup \mathrm{B}^\mathrm{C})},$$

- when 1_A and 1_B are anticomonotonic random variables (A^C Ì B),

$s \in B^C \Rightarrow I(1_A/1_B) - 1_A = 0$,

$s \in A^C \cap B \Rightarrow I(1_A/1_B) - 1_A = I(1_A/1_B = 1) \geq 0$,

$s \in A \cap B \Rightarrow I(1_A/1_B) - 1_A = I(1_A/1_B = 1) - 1 \leq 0$,

then:

$$\int_S [I(1_A/1_B) - 1_A]dn = I(1_A/1_B = 1) - 1 + [1 - I(1_A/1_B = 1)]n(A^C \cup B^C) + I(1_A/1_B = 1)n(A^C),$$

and relation (7) implies:

$$\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 1) = \frac{1 - n(\mathrm{A}^\mathrm{C})}{1 + n(\mathrm{A}^\mathrm{C} \cap \mathrm{B}) - n(\mathrm{A}^\mathrm{C})}, \textbf{ QED}.$$

Proof of proposition 1.6: (i) - when $C = B$:

$s \in B^C \Rightarrow [I(1_A/1_B) - 1_A]1_B = 0$,

$s \in A \cap B \Rightarrow [I(1_A/1_B) - 1_A]1_B = I(1_A/1_B = 1) - 1 \leq 0$,

$s \in A^C \cap B \Rightarrow [I(1_A/1_B) - 1_A]1_B = I(1_A/1_B = 1) \geq 0$,

then:

$$\int_B [I(1_A/1_B) - 1_A]dn = I(1_A/1_B = 1) - 1 + [1 - I(1_A/1_B = 1)]n(A^C \cup B^C) + I(1_A/1_B = 1)n(A^C \cap B)$$

and (8) implies:

$$\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 1) = \frac{1 - n(\mathrm{A}^\mathrm{C} \cup \mathrm{B}^\mathrm{C})}{1 + n(\mathrm{A}^\mathrm{C} \cap \mathrm{B}) - n(\mathrm{A}^\mathrm{C} \cup \mathrm{B}^\mathrm{C})}.$$

- when $C = B^C$:

$s \in B \Rightarrow [I(1_A/1_B) - 1_A]1_{B^C} = 0$,

$s \in A \cap B^C \Rightarrow [I(1_A/1_B) - 1_A]1_{B^C} = I(1_A/1_B = 0) - 1 \leq 0$,

$s \in A^C \cap B^C \Rightarrow [I(1_A/1_B) - 1_A]1_{B^C} = I(1_A/1_B = 0) \geq 0$,

then:

$$\int_{B^C} [I(1_A/1_B) - 1_A]dn = I(1_A/1_B = 0) - 1 + [1 - I(1_A/1_B = 0)]n(A^C \cup B) + I(1_A/1_B = 0)n(A^C \cap B^C)$$

and (8) implies:

$$\mathrm{I}(1_\mathrm{A}/1_\mathrm{B} = 0) = \frac{1 - n(\mathrm{A}^\mathrm{C} \cup \mathrm{B})}{1 + n(\mathrm{A}^\mathrm{C} \cap \mathrm{B}^\mathrm{C}) - n(\mathrm{A}^\mathrm{C} \cup \mathrm{B})}.$$

(ii) We have:

$s \in A \cap B \Rightarrow I(1_A/1_B) - 1_A = I(1_A/1_B = 1) - 1 \leq 0$,

$s \in A \cap B^C \Rightarrow I(1_A/1_B) - 1_A = I(1_A/1_B = 0) - 1 \leq 0$,

$s \in A^C \cap B \Rightarrow I(1_A/1_B) - 1_A = I(1_A/1_B = 1) \geq 0$,

$s \in A^C \cap B^C \Rightarrow I(1_A / 1_B) - 1_A = I(1_A / 1_B = 0) \geq 0$.

- If $I(1_A / 1_B = 0) \leq I(1_A / 1_B = 1)$,

$$\int_C [I(1_A/1_B) - 1_A] dn = \int_{I(1_A/1_B=0)-1}^{I(1_A/1_B=1)-1} \{n[(A^C \cup B) \cap C] - n(C)\} dx + \int_{I(1_A/1_B=1)-1}^{0} [n(A^C \cap C) - n(C)] dx$$

$$+ \int_0^{I(1_A/1_B=0)} n(A^C \cap C) dx + \int_{I(1_A/1_B=0)}^{I(1_A/1_B=1)} n(A^C \cap B \cap C) dx.$$

With $C = B$ relation (8) gives: $I(1_A/1_B = 1) = \dfrac{n(B) - n(A^C \cap B)}{n(B)}$,

and with $C = B^C$, we have: $I(1_A/1_B = 0) = \dfrac{n(B^C) - n(A^C \cap B^C)}{n(B^C)}$.

- If $I(1_A / 1_B = 1) \leq I(1_A / 1_B = 0)$,

$$\int_C [I(1_A/1_B) - 1_A] dn = \int_{I(1_A/1_B=1)-1}^{I(1_A/1_B=0)-1} \{n[(A^C \cup B^C) \cap C] - n(C)\} dx + \int_{I(1_A/1_B=0)-1}^{0} [n(A^C \cap C) - n(C)] dx$$

$$+ \int_0^{I(1_A/1_B=1)} n(A^C \cap C) dx + \int_{I(1_A/1_B=1)}^{I(1_A/1_B=0)} n(A^C \cap B^C \cap C) dx.$$

Relation (8) yields the same results, **QED**.

Appendix 2

A counter-example: CKL formula doesn't satisfy Consequentialism (Karni and Schmeidler (1991)

Let us consider $S = \{s_1, s_2, s_3, s_4\}$ and four risks *X, X', Y, Y'* with payoffs:

$x(\{s_1\}) = 1, x(\{s_2\}) = 2, x(\{s_3\}) = 7.5, x(\{s_4\}) = 8$

$x'(\{s_1\}) = 1, x'(\{s_2\}) = 2, x'(\{s_3\}) = 3, x'(\{s_4\}) = 13.5$

$y(\{s_1\}) = 15, y(\{s_2\}) = 14, y(\{s_3\}) = 7.5, y(\{s_4\}) = 8$

$y'(\{s_1\}) = 15, y'(\{s_2\}) = 14, y'(\{s_3\}) = 3, y'(\{s_4\}) = 13.5$

After event $B = \{s_3, s_4\}$ has been realised, the value of these risks are respectively:

$V^B(X) = 1 + n(\{s_2, s_3, s_4\}/B) + 5.5\, n(\{s_3, s_4\}/B) + 0.5\, n(\{s_4\}/B)$

$V^B(X') = 1 + n(\{s_2, s_3, s_4\}/B) + n(\{s_3, s_4\}/B) + 10.5\, n(\{s_4\}/B)$

$V^B(Y) = 7.5 + 0.5\, n(\{s_1, s_2, s_4\}/B) + 6\, n(\{s_1, s_2\}/B) + n(\{s_1\}/B)$

$V^B(Y') = 3 + 10.5\, n(\{s_1, s_2, s_4\}/B) + 0.5\, n(\{s_1, s_2\}/B) + n(\{s_1\}/B)$

From Proposition 1.1:

- Because $B \subset \{s_2, s_3, s_4\}$, $B = \{s_3, s_4\}$, $\{s_4\} \subset B$, the characteristic functions of these three sets are comonotonic with the characteristic function of *B*. Then, their conditional capacities are given by Bayes updating rule: $n(\{s_2, s_3, s_4\}/B) = n(\{s_3, s_4\}/B) = 1$, $n(\{s_4\}/B) = \dfrac{n(\{s_4\})}{n(B)}$.

- Because $\{s_1\} \subset B^C$, $B^C = \{s_1, s_2\}$, $B^C \subset \{s_1, s_2, s_4\}$, the characteristic functions of these three sets are antimonotonic with the characteristic function of *B*. Then, their conditional capacities are given by Dempster-Schafer updating rule: $n(\{s_1\}/B) = n(\{s_1, s_2\}/B) = 0$,

$$n(\{s_1, s_2, s_4\}/B) = \frac{n(\{s_1, s_2, s_4\} - n(B^C))}{1 - n(B^C)}.$$

Therefore, we have:

$$V^B(X) = 7.5 + 0.5\,\frac{n(\{s_4\})}{n(B)},\quad V^B(X') = 3 + 10.5\,\frac{n(\{s_4\})}{n(B)}$$

$$V^B(Y) = 7.5 + 0.5\,\frac{n(\{s_1, s_2, s_4\} - n(B^C))}{1 - n(B^C)},\quad V^B(Y') = 3 + 10.5\,\frac{n(\{s_1, s_2, s_4\} - n(B^C))}{1 - n(B^C)}$$

Let n be a capacity such that:

$n(\{s_4\}) = 0.3$, $n(\{s_3, s_4\}) = 0.6$, $n(\{s_1, s_2\}) = 0.5$, $n(\{s_1, s_2, s_4\}) = 0.7$.

We obtain: $V^B(X) = 7.75 < V^B(X') = 8.25$ and $V^B(Y) = 7.7 > V^B(Y') = 7.2$,

which is equivalent to: $X \prec_B X'$ and $Y \succ_B Y'$: A contradiction to Consequentialism.