COMMONOTONICITY AND TIME CONSISTENCY FOR LEBESGUE CONTINUOUS MONETARY UTILITY FUNCTIONS

FREDDY DELBAEN

Abstract. It is proved that commonotonicity and time consistence for monetary utility functions do not go together. I also give additional results on atomless and conditionally atomless probability spaces.

1. Notation

The purpose of this paper[1] is to investigate the relation between commonotonicity and time consistency of monetary utility functions[2]. It will turn out that it is sufficient to have a two period model. In this setting we will work with a probability space equipped with three sigma algebras $(\Omega, \mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2, \mathbb{P})$. The sigma algebra $\mathcal{F}_0$ is supposed to be trivial $\mathcal{F}_0 = \{\emptyset, \Omega\}$ whereas the sigma algebra $\mathcal{F}_2$ is supposed to express innovations with respect to $\mathcal{F}_1$. Since we do not put topological properties on the set Ω we will make precise definitions later that do not use conditional probability kernels. But essentially we could say that we suppose that conditionally on $\mathcal{F}_1$ the probability $\mathbb{P}$ is atomless on $\mathcal{F}_2$. We will show that such an hypothesis implies that there is an atomless sigma algebra $\mathcal{B} \subset \mathcal{F}_2$ that is independent of $\mathcal{F}_1$. The space $L^\infty(\mathcal{F}_i)$ will denote the space of bounded $\mathcal{F}_i$ measurable random variables, modulo equality almost surely, a.s. . We will also suppose that there is a time consistent utility function $u_{0,2}\colon L^\infty(\mathcal{F}_2) \to \mathbb{R}$. As shown in [7] this means that we also have utility functions $u_{1,2}\colon L^\infty(\mathcal{F}_2) \to L^\infty(\mathcal{F}_1)$ and $u_{0,1}\colon L^\infty(\mathcal{F}_1) \to L^\infty(\mathcal{F}_0) = \mathbb{R}$ such that $u_{0,2} = u_{0,1} \circ u_{1,2}$. In particular $u_{0,1}$ is simply the restriction of $u_{0,2}$ to $L^\infty(\mathcal{F}_1)$. Our utility functions are monetary and concave which is expressed in the following list of properties, valid for all $0 \le i < j \le 2$:

(1) $u_{i,j}\colon L^\infty(\mathcal{F}_j) \to L^\infty(\mathcal{F}_i)$, if $\xi \ge 0$ then also $u_{i,j}(\xi) \ge 0$ and $u_{i,j}(0) = 0$.

(2) For $\xi, \eta \in L^\infty(\mathcal{F}_j)$, $0 \le \lambda \le 1$ and $\mathcal{F}_i$ measurable, we have
$$u_{i,j}(\lambda\xi + (1-\lambda)\eta) \ge \lambda u_{i,j}(\xi) + (1-\lambda)u_{i,j}(\eta).$$

(3) Since commonotonicity implies positive homogeneity we will use a stronger property and suppose coherence. For $\xi \in L^\infty(\mathcal{F}_j)$, $0 \le \lambda$ and $\mathcal{F}_i$ measurable, we have
$$u_{i,j}(\lambda\xi) = \lambda u_{i,j}(\xi).$$

(4) For $\xi \in L^\infty(\mathcal{F}_j)$ and $a \in L^\infty(\mathcal{F}_i)$ we have
$$u_{i,j}(\xi + a) = u_{i,j}(\xi) + a.$$

Date: First version December 2018, this version December 16, 2020.

[1]AMS-classification primary 91B10, secondary 91B16, 60A99

[2]This research was done while the author was visiting Tokyo Metropolitan University in October and November 2018. I thank the staff of TMU for the many fruitful discussions and in particular I thank Prof. Adachi for the many critical remarks.

(5) We will need Lebesgue continuity which means: if $\xi_n \in L^\infty(\mathcal{F}_j)$ is a uniformly bounded sequence such that $\xi_n \to \eta$ in probability then $u_{i,j}(\xi_n)$ tends to $u_{i,j}(\eta)$ in probability.
(6) The Lebesgue property is stronger than the Fatou property which says that for a sequence $\xi_n \in L^\infty$ such that a.s. $\xi_n \downarrow \eta \in L^\infty$ we have $u_{ij}(\xi_n) \to u_{ij}(\eta)$, a.s. .

The utility functions we need are coherent and hence we can use the dual representation. There is a set of probability measures, $\mathcal{S}$, absolutely continuous with respect to $\mathbb{P}$ such that

$$u_{0,2}(\xi) = \inf_{\mathbb{Q}\in\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\xi].$$

The set $\mathcal{S}$ is seen as a subset of L^1 and is supposed to be convex and closed. The Lebesgue continuity is equivalent to the weak compactness of $\mathcal{S}$. We will suppose that our utility functions are *relevant*, i.e. for each A with $\mathbb{P}[A] > 0$ we have $u(-\mathbf{1}_A) < 0$. see [7]. By the Halmos-Savage theorem this means that $\mathcal{S}$ contains an equivalent probability measure. We need this property in order to avoid some problems with negligible sets appearing in the definition and comparison of conditional expectations.

Without further notice we will always assume that our utility functions are relevant and Lebesgue continuous. These assumptions are not always needed, sometimes Fatou continuity is sufficient. Since we want to put more emphasis on the methods of proof, we will not aim for the most general results.

The time consistency also allows to construct the utility functions $u_{i,j}$ once $u_{0,2}$ is known. The construction is easier when $u_{0,2}$ is relevant. The Fatou or Lebesgue property is less important for this development. As shown in [7], there is a way to check whether the utility function $u_{0,2}$ can be extended to a time consistent family of utility functions. To do this we introduce the acceptability cones $\mathcal{A}_{0,2} = \{\xi \mid u_{0,2}(\xi) \geq 0\}$, $\mathcal{A}_{0,1} = \{\xi \in L^\infty(\mathcal{F}_1) \mid u_{0,2}(\xi) \geq 0\}$, $\mathcal{A}_{1,2} = \{\xi \in L^\infty(\mathcal{F}_2) \mid \text{ for all } A \in \mathcal{F}_1 : u_{0,2}(\xi \mathbf{1}_A) \geq 0\}$. The necessary and sufficient condition for the existence of a time consistent extension is $\mathcal{A}_{0,2} = \mathcal{A}_{0,1} + \mathcal{A}_{1,2}$. If this is fulfilled we put $u_{1,2}(\xi) = \text{ess.inf}\{\eta \in L^\infty(\mathcal{F}_1) \mid \xi - \eta \in \mathcal{A}_{1,2}\}$. $u_{0,1}$ is simply the restriction of $u_{0,2}$ to $L^\infty(\mathcal{F}_1)$. This gives sense to expressions such as: $u_{0,2}$ is time consistent.

We say that two random variables ξ, η are *commonotonic* if there are two nondecreasing functions $f, g\colon \mathbb{R} \to \mathbb{R}$ and a random variable ζ such that $\xi = f(\zeta), \eta = g(\zeta)$. *Commonotonicity can be seen as the opposite of diversification.* If ζ increases then both ξ and η increase (or better do not decrease). By the way in case ξ and η are commonotonic then one can choose $\zeta = \xi + \eta$, see [7]. It can be shown that in this case one can choose representatives — still denoted (ξ, η) — such that $(\xi(\omega) - \xi(\omega'))(\eta(\omega) - \eta(\omega')) \geq 0$ for all ω, ω'. Since we do not need this result, we do not include a proof. We say that a set $E \subset \mathbb{R}^2$ is commonotonic if $(x, y), (x', y') \in E$ implies $(x - x')(y - y') \geq 0$ [3]. Random variables ξ, η are commonotonic if and only if the support of the image measure of (ξ, η) is a commonotonic set. We will use a special commonotonic subset of $\mathbb{R}^2$.

Already in case the utility functions are expected value and conditional expectations, the main theorem leads to the following result. (The notion conditionally atomless will be explained and analysed in the next section.)

Theorem 1. *If $\mathcal{F}_2$ is conditionally atomless with respect to $\mathcal{F}_1$ then for any couple (f, g) of $\mathcal{F}_1$-measurable finitely valued random variables, there is a commonotonic couple (ξ, η)*

[3] In convex function theory, such sets are also called monotone or monotonic sets.

of $\mathcal{F}_2$-measurable random variables such that (in an extended sense, made precise later) $f = \mathbb{E}[\xi \mid \mathcal{F}_1], g = \mathbb{E}[\eta \mid \mathcal{F}_1]$. Furthermore for every norm on $\mathbb{R}^2$ there is a constant C, such that $\|(\xi, \eta)\| \leq C\|(f, g)\|$ almost surely.

Both concepts, time consistency and commonotonicity, are important in the theory of risk evaluation. The concept of time consistency (and inconsistency) was introduced and investigated by Koopmans, [12]. The role of commonotonicity found its way in insurance and is present in several papers. The use of Choquet integration as premium principle was emphasized by Denneberg, [9]. Denneberg was inspired by the pioneering work of Yaari, [21]. Schmeidler proved the relation between commonotonic principles, convex games and Choquet integration, [14]. Modern uses can be found in for instance [17] and [18]. For more references and different proofs of these results I refer to [7]. Although commonotonicity seems to be a desirable property, there might be some difficulties when insurance contracts are priced in this way, see [5] for some unexpected consequences.

The concept of risk measures (up to sign changes monetary utility functions) was introduced in [1] and [2].

Using the general version of the theorem above we will show that except in very restrictive cases, a utility function $u_{0,2}$ cannot be time consistent and commonotonic at the same time. It seems that time consistency is a strong property that excludes some other *desirable* properties. For instance in [11] it is shown that in a filtration with innovations (comparable to the requirement of being conditionally atomless) utility functions that are time consistent and law determined, are necessarily of entropic type. We refer to [11] for the details and for the precise form of the innovations. The present paper studies time consistent utility functions that might depend on past history and are not necessarily law determined. The methods we use are different from the approaches used for law determined or law invariant utility functions. Among the many papers on these utility functions, we could refer the reader to e.g. [3],[4],[5],[11],[19],[20],[22].

2. Atomless Extension of Sigma Algebras

In this section we work with a probability space $(\Omega, \mathcal{F}_2, \mathbb{P})$ equipped with the filtration $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2$.

Definition 1. We say that $\mathcal{F}_2$ is atomless conditionally to $\mathcal{F}_1$ if the following holds. For every $A \in \mathcal{F}_2$ there exists a set $B \subset A$, $B \in \mathcal{F}_2$, such that $0 < \mathbb{E}[\mathbf{1}_B \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]$ on the set $\{\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1] > 0\}$.

In case the conditional expectation could be calculated with a – under extra topological conditions – regular probability kernel, say $K(\omega, A)$, then the above definition is a measure theoretic way of saying that the probability measure $K(\omega, .)$ is atomless for almost every $\omega \in \Omega$. The precise relation between these two notions is not the topic of this paper. See [8] for the details.

Theorem 2. *$\mathcal{F}_2$ is atomless conditionally to $\mathcal{F}_1$ if for every $A \in \mathcal{F}_2$, $\mathbb{P}[A] > 0$, there is $B \subset A$ such that*

$$\mathbb{P}\left[0 < \mathbb{E}[\mathbf{1}_B \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]\right] > 0.$$

Proof The proof is a standard exhaustion argument. For completeness we give the details. Let $\mathcal{D}$ be the collection of $\mathcal{F}_1-$measurable sets:

$$\mathcal{D} = \{\{0 < \mathbb{E}[\mathbf{1}_B \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]\} \mid B \subset A\}$$

We show that there is a biggest set in $\mathcal{D}$ and this set must then be equal to $\{\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1] > 0\}$. To show that there is a biggest set in $\mathcal{D}$ it is sufficient to show that $\mathcal{D}$ is stable for countable unions. Let D_n be a sequence in $\mathcal{D}$ and suppose that for each n we have a set $B_n \subset A$ such that $D_n = \{0 < \mathbb{E}[\mathbf{1}_{B_n} \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]\}$. Now take

$$B = \cup_n \left(B_n \cap \left(D_n \setminus \left(\cup_{k\le n-1} D_k\right)\right)\right).$$

It is easy to check that $\{0 < \mathbb{E}[\mathbf{1}_B \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]\} = \cup_n D_n$ and therefore $\cup_n D_n \in \mathcal{D}$. Let now $D = \{0 < \mathbb{E}[\mathbf{1}_B \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]\}$ be a maximum in $\mathcal{D}$. Suppose that $\mathbb{P}[\{\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1] > 0\} \setminus D] > 0$. This implies that $\mathbb{P}[A \setminus D] > 0$. According to the hypothesis of the theorem, there will be a set $B' \subset (A \setminus D)$ with $D' = \{0 < \mathbb{E}[\mathbf{1}_{B'} \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_{A\setminus D} \mid \mathcal{F}_1]\}$ and non-negligible. Since $D \cup D' \in \mathcal{D}$ and $D \cap D' = \emptyset$, the element D is not a maximum, a contradiction.

The main result of this section is the following

Theorem 3. *$\mathcal{F}_2$ is atomless conditionally to $\mathcal{F}_1$ if and only if there exists an atomless sigma algebra $\mathcal{B} \subset \mathcal{F}_2$ that is independent of $\mathcal{F}_1$.*

The "if" part is easy but requires some continuity argument. Because $\mathcal{B}$ is atomless, there is a $\mathcal{B}$-measurable, $[0,1]$ uniformly distributed random variable U. The sets $B_t = \{U \le t\}, 0 \le t \le 1$ form an increasing family of sets with $\mathbb{P}[B_t] = t$. Let $A \in \mathcal{F}_2$ and let $F = \{0 < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]\}$. We may suppose that $\mathbb{P}[F] > 0$ since otherwise there is nothing to prove. We will show that there is $t \in]0,1[$ with $\mathbb{P}\left[0 < \mathbb{E}[\mathbf{1}_{A\cap B_t} \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]\right] > 0$. According to the previous theorem, $\mathcal{F}_2$ is conditionally atomless with respect to $\mathcal{F}_1$. Obviously for $0 \le s \le t \le 1$ we have, by independence of $\mathcal{B}$ and $\mathcal{F}_1$:

$$\|\mathbb{E}[\mathbf{1}_{A\cap B_t} \mid \mathcal{F}_1] - \mathbb{E}[\mathbf{1}_{A\cap B_s} \mid \mathcal{F}_1]\|_\infty \le \|\mathbb{E}[\mathbf{1}_{B_t\setminus B_s} \mid \mathcal{F}_1]\|_\infty = t - s.$$

It follows that there is a set of measure 1, say Ω', such that for all $s \le t$, rational, and all $\omega \in \Omega'$, $\mathbb{E}\left[\mathbf{1}_{A\cap B_t} \mid \mathcal{F}_1\right](\omega)$ can be taken to satisfy

$$|\mathbb{E}[\mathbf{1}_{A\cap B_t} \mid \mathcal{F}_1](\omega) - \mathbb{E}[\mathbf{1}_{A\cap B_s} \mid \mathcal{F}_1](\omega)| \le t - s.$$

For each $\omega \in \Omega'$ we can extend the function

$$\{q \in [0,1] \mid q \text{ rational }\} \to \mathbb{E}[\mathbf{1}_{A\cap B_q} \mid \mathcal{F}_1](\omega)$$

to a continuous function on $[0,1]$. The resulting continuous extension then represents $(\mathbb{E}[\mathbf{1}_{A\cap B_t} \mid \mathcal{F}_1])_t$. For $t = 0$ we have zero and for $t = 1$ we find $\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]$. Because for $\omega \in \Omega'$, the trajectories are continuous, a simple application of Fubini's theorem shows that the real valued function

$$t \to \mathbb{P}\left[0 < \mathbb{E}[\mathbf{1}_{A\cap B_t} \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]\right]$$

becomes strictly positive for some t. With some extra work – as will be done later – one can even show that there is $G \subset A$ such that $\mathbb{E}[\mathbf{1}_G \mid \mathcal{F}_1] = (1/2)\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]$. For completeness let us now give the details of the application of Fubini's theorem. Suppose on the contrary that for all $t \in [0,1]$ we have

$$\mathbb{P}\left[0 < \mathbb{E}[\mathbf{1}_{A\cap B_t} \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1]\right] = 0.$$

Then on the product space $[0,1] \times \Omega'$ we find that the (clearly measurable) set

$$\{(t,\omega) \mid 0 < \mathbb{E}[\mathbf{1}_{A\cap B_t} \mid \mathcal{F}_1](\omega) < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1](\omega)\}$$

has $m \times \mathbb{P}$ measure zero (m denotes Lebesgue measure). By Fubini's theorem we have that for almost all $\omega \in \Omega'$, the set

$$\{t \mid 0 < \mathbb{E}[\mathbf{1}_{A\cap B_t} \mid \mathcal{F}_1](\omega) < \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1](\omega)\}$$

must have Lebesgue measure zero. However, for $\omega \in \Omega'$, this contradicts the continuity of the mapping

$$t \to \mathbb{E}[\mathbf{1}_{A\cap B_t} \mid \mathcal{F}_1](\omega).$$

The proof of the “only if" part is broken down in several steps stated in the lemmata that follow. We will without further notice, always suppose that $\mathcal{F}_2$ is atomless conditionally to $\mathcal{F}_1$.

Lemma 1. *Suppose $A \in \mathcal{F}_1$ and $C \subset A$ is such that $\mathbb{E}[\mathbf{1}_C \mid \mathcal{F}_1] > 0$ on A. Then we can construct a decreasing sequence of sets $(B_n)_{n\geq 0}$, $B_n \subset C$, such that $0 < \mathbb{E}[\mathbf{1}_{B_n} \mid \mathcal{F}_1] \leq 2^{-n}$ on A.*

Proof The statement is obviously true for $n = 0$ since we can take $B_0 = C$. We now proceed by induction and suppose the statement holds for n. So the set $B_n \subset A$ satisfies $0 < \mathbb{E}[\mathbf{1}_{B_n} \mid \mathcal{F}_1] \leq 2^{-n}$ on A. Clearly $A \subset \{\mathbb{E}[\mathbf{1}_{B_n} \mid \mathcal{F}_1] > 0\}$. By assumption there is a set $D \subset B_n$ such that on $A \subset \{\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1] > 0\}$ we have

$$0 < \mathbb{E}[\mathbf{1}_D \mid \mathcal{F}_1] < \mathbb{E}[\mathbf{1}_{B_n} \mid \mathcal{F}_1].$$

We now take

$$B_{n+1} = \left(D \cap \left\{\mathbb{E}[\mathbf{1}_D \mid \mathcal{F}_1] \leq \frac{1}{2}\mathbb{E}[\mathbf{1}_{B_n} \mid \mathcal{F}_1]\right\}\right) \cup \left((B_n \setminus D) \cap \left\{\mathbb{E}[\mathbf{1}_D \mid \mathcal{F}_1] > \frac{1}{2}\mathbb{E}[\mathbf{1}_{B_n} \mid \mathcal{F}_1]\right\}\right).$$

The set B_{n+1} satisfies the requirements.

Lemma 2. *Let $C \in \mathcal{F}_2$ and let $h\colon \Omega \to [0,1]$ be $\mathcal{F}_1$ measurable. Then there is a set $B \subset C$ such that $\mathbb{E}[\mathbf{1}_B \mid \mathcal{F}_1] = h\,\mathbb{E}[\mathbf{1}_C \mid \mathcal{F}_1]$.*

Proof Let $B_0 = \emptyset$. Inductively we define for $n \geq 1$, classes $\mathcal{B}_n$ and sets $B_n \in \mathcal{B}_n$. For $n \geq 1$ let

$$\mathcal{B}_n = \{B_{n-1} \subset B \subset C \mid B \in \mathcal{F}_2,\, \mathbb{E}[\mathbf{1}_B \mid \mathcal{F}_1] \leq h\,\mathbb{E}[\mathbf{1}_C \mid \mathcal{F}_1]\}.$$

Let $\beta_n = \sup\{\mathbb{P}[B] \mid B \in \mathcal{B}_n\}$ and take $B_n \in \mathcal{B}_n$ such that $\mathbb{P}[B_n] \geq (1 - 2^{-n})\beta_n$. Clearly B_n is non-decreasing and let $B_\infty = \cup_n B_n$. Obviously

$$\mathbb{P}[B_\infty] \geq \limsup \beta_n \geq \liminf \beta_n \geq \lim \mathbb{P}[B_n] = \mathbb{P}[B_\infty].$$

We claim the $\mathbb{E}[\mathbf{1}_{B_\infty} \mid \mathcal{F}_1] = h\,\mathbb{E}[\mathbf{1}_C \mid \mathcal{F}_1]$. We already have that $\mathbb{E}[\mathbf{1}_{B_\infty} \mid \mathcal{F}_1] \leq h\,\mathbb{E}[\mathbf{1}_C \mid \mathcal{F}_1]$. If $\mathbb{P}\left[\mathbb{E}[\mathbf{1}_{B_\infty} \mid \mathcal{F}_1] < h\,\mathbb{E}[\mathbf{1}_C \mid \mathcal{F}_1]\right] > 0$ then $\mathbb{P}[B_\infty] < \mathbb{P}[C]$ and there must be $m \geq 1$ such that $\mathbb{P}\left[\mathbb{E}[\mathbf{1}_{B_\infty} \mid \mathcal{F}_1] < h\,\mathbb{E}[\mathbf{1}_C \mid \mathcal{F}_1] - 2^{-m}\right] > 0$. The previous lemma allows to find $D \subset C \setminus B_\infty$, $\mathbb{P}[D] = \eta > 0$, such that $0 < \mathbb{E}[\mathbf{1}_D \mid \mathcal{F}_1] \leq 2^{-m}$ on the set $\{\mathbb{E}[\mathbf{1}_B \mid \mathcal{F}_1] < h\,\mathbb{E}[\mathbf{1}_C \mid \mathcal{F}_1] - 2^{-m}\}$ and zero elsewhere. The set $D \cup B_\infty$ is in all classes $\mathcal{B}_n$ and for n big enough:

$$\beta_n \geq \mathbb{P}[D \cup B_\infty] \geq \mathbb{P}[B_n] + \eta \geq (1 - 2^{-n})\beta_n + \eta \geq \beta_n + \eta - 2^{-n} > \beta_n,$$

yielding a contradiction. So we must have $\mathbb{E}[\mathbf{1}_{B_\infty} \mid \mathcal{F}_1] = h\,\mathbb{E}[\mathbf{1}_C \mid \mathcal{F}_1]$.

Remark 1. The lemma above is a variant of Sierpiński's theorem, [15]. This theorem states that in an atomless probability space $(\Omega, \mathcal{E}, \mathbb{P})$, for every set $A \in \mathcal{E}$ and every $0 < t < 1$, there is a set $B \subset A$ with $\mathbb{P}[B] = t\mathbb{P}[A]$. The usual proof — presented in many probability courses — uses the Axiom of Choice (AC). A referee pointed out that for many people AC – or Zorn's lemma – is an extra assumption. To prove Sierpiński's theorem we only need the Axiom of Countable Dependent Choice, which is a countable form of the axiom of choice. In analysis this is the axiom that is usually needed and used. The proof above follows the approach given by Lorenc and Witula, [13].

Lemma 3. *There is an increasing family of sets* $(B_t)_{t\in[0,1]}$ *such that* $\mathbb{E}[\mathbf{1}_{B_t} \mid \mathcal{F}_1] = t$. *The sigma algebra* $\mathcal{B}$, *generated by the family* $(B_t)_t$ *is independent of* $\mathcal{F}_1$. *The system* $(B_t)_t$ *can also be described as* $B_t = \{U \le t\}$ *where* U *is a random variable that is independent of* $\mathcal{F}_1$ *and uniformly distributed on* $[0, 1]$.

Proof The proof is a repeated use of the previous lemma where we take $h = 1/2$. We start with $B_0 = \emptyset, B_1 = \Omega$. Suppose that for the diadic numbers $k2^{-n}, k = 0, \dots 2^n$ the sets are already defined. Then we consider the set $B_{(k+1)2^{-n}} \setminus B_{k2^{-n}}$ and apply the previous lemma with $h = 1/2$. We get a set $D \subset B_{(k+1)2^{-n}} \setminus B_{k2^{-n}}$ with $\mathbb{E}[\mathbf{1}_D \mid \mathcal{F}_1] = 2^{-(n+1)}$. We then define $B_{(2k+1)2^{-(n+1)}} = B_{k2^{-n}} \cup D$. For non-diadic numbers t we find a sequence of diadic numbers d_n such that $d_n \uparrow t$. Then we define $B_t = \cup_n B_{d_n}$. This completes the construction. Since the system $(B_t)_t$ is trivially stable for intersection, the relation $\mathbb{E}[\mathbf{1}_{B_t} \mid \mathcal{F}_1] = t$ shows that the sigma algebra $\mathcal{B}$ generated by $(B_t)_t$, is independent of $\mathcal{F}_1$. The construction of U is standard. At level n we put $U_n = \sum_{k=1,\dots 2^n} k2^{-n}\mathbf{1}_{B_{k2^{-n}} \setminus B_{(k-1)2^{-n}}}$. U_n then decreases to a random variable U that satisfies the needed properties. The proof of the theorem is now completed.

Remark 2. Suppose that for the probability $\mathbb{P}$, there is an atomless sigma algebra $\mathcal{B} \subset \mathcal{F}_2$ that is independent of $\mathcal{F}_1$. Suppose now that $\mathbb{Q} \sim \mathbb{P}$ is an equivalent probability measure. Clearly the definition of being conditionally atomless is invariant for equivalent measure changes. Hence there is an atomless sigma algebra $\mathcal{B}' \subset \mathcal{F}_2$ that is independent of $\mathcal{F}_1$ for the probability $\mathbb{Q}$. Proving this directly does not seem easy.

The following proposition is Lemma 2 where we take $C = \Omega$. For didactical reasons we give another proof that directly uses the existence of an independent sigma algebra. We use the same assumptions and notations as in the theorem above.

Proposition 1. *For every* $\mathcal{F}_1$ *measurable function* $h\colon \Omega \to [0, 1]$, *there is a set* $B_h \in \mathcal{F}_2$ *such that* $\mathbb{E}[\mathbf{1}_{B_h} \mid \mathcal{F}_1] = h$.

Proof The idea is to use the set B_t on the set $\{h = t\}$, i.e. $B = \cup_t(\{h = t\} \cap B_t)$. However, because the set of real numbers is uncountable, this definition is not good enough to obtain a set in $\mathcal{F}_2$. So we need a trick. Let ϕ be the mapping

$$\phi\colon (\Omega, \mathcal{F}_2) \to (\Omega, \mathcal{F}_1) \times (\Omega, \mathcal{B}), \phi(\omega) = (\omega, \omega).$$

This mapping is obviously measurable and the image measure is – because of independence – the product measure. We also define $h_1(\omega, \omega') = h(\omega)$ and $U_2(\omega, \omega') = U(\omega')$. For $A \in \mathcal{F}_1$ we denote $A_1 = A \times \Omega$. We define $B_h = \{U \le h\} = \phi^{-1}\{U_2 \le h_1\}$. We now verify that

$\mathbb{E}[\mathbf{1}_{B_h} \mid \mathcal{F}_1] = h$. To do this we calculate for a set $A \in \mathcal{F}_1$ the probability $\mathbb{P}[B_h \cap A]$.

$$\begin{aligned}
\mathbb{P}[B_h \cap A] &= \mathbb{P} \times \mathbb{P}[(U_2 \le h_1) \cap A_1] \\
&= \int \mathbb{P}[d\omega'] \int \mathbb{P}[d\omega] \mathbf{1}_{\{U_2 \le h_1\}}(\omega, \omega') \mathbf{1}_{A_1}(\omega, \omega') \\
&= \int \mathbb{P}[d\omega'] \, \mathbb{P}[\{h \ge U(\omega'\} \cap A] \\
&= \int_0^1 dt \, \mathbb{P}[\{h \ge t\} \cap A] \\
&= \mathbb{E}[h \mathbf{1}_A],
\end{aligned}$$

showing $\mathbb{E}[\mathbf{1}_{B_h} \mid \mathcal{F}_1] = h$.

Remark 3. The previous theorem is not actually needed. We will need the stronger version where the conditional expectation is replaced by the utility function $u_{1,2}$. To prove this stronger version, we will use a slightly different approach. However in the case where we are only interested in conditional expectations the above proof might be of some didactical interest.

Remark 4. After the first version was made available, I got the remark that the paper [16] of Shen, J., Shen, Y., Wang, B., and Wang, R. contains similar concepts and results.[4] In their notation they work with a measurable space $(\Omega, \mathcal{A})$ on which they have a finite number of probability measures $\mathbb{Q}_1, \ldots, \mathbb{Q}_n$.[5] They introduce

Definition 2. The set $(\mathbb{Q}_1, \ldots, \mathbb{Q}_n)$ is conditionally atomless if there exists a dominating measure $\mathbb{Q}$ (i..e $\mathbb{Q}_k \ll \mathbb{Q}$ for each $k \le n$) as well as a continuously distributed random variable X (for the measure $\mathbb{Q}$) such that the vector of Radon-Nikodym derivatives $\left(\frac{d\mathbb{Q}_k}{d\mathbb{Q}}\right)_k$ is independent of X.

They then prove the following

Proposition 2. *Are equivalent*

(1) $(\mathbb{Q}_1, \ldots, \mathbb{Q}_n)$ *is conditionally atomless*
(2) *in the definition we can take* $\mathbb{Q} = \frac{1}{n}(\mathbb{Q}_1 + \ldots + \mathbb{Q}_n)$
(3) X *can be taken as uniformly distributed over* $[0, 1]$.

There are several differences with my approach. There is the technical difference that they suppose the existence of a continuously distributed random variable X. In doing so they avoid the technical points between the more conceptual definition using conditional expectations and the construction of a suitable sigma-algebra with a uniformly distributed random variable. A further difference is that they use a dominating measure that later can be taken as the mean of $(\mathbb{Q}_1, \ldots, \mathbb{Q}_n)$. Of course their result together with the results here show that the definition of $(\mathbb{Q}_1, \ldots, \mathbb{Q}_n)$ being conditionally atomless, is equivalent to the statement that for the measure $\mathbb{Q}_0 = \frac{1}{n}(\mathbb{Q}_1 + \ldots + \mathbb{Q}_n)$, the sigma algebra $\mathcal{A}$ is conditionally

[4]I thank Ruodu Wang for pointing out these relations and for the subsequent discussions we had on the topic.

[5]Their paper also considers an infinite number of measures but to clarify the relation between their paper and my approach, I only consider a finite number of measures.

atomless with respect to the sigma-algebra generated by the Radon-Nikodym derivatives $\left(\frac{d\mathbb{Q}_k}{d\mathbb{Q}_0}\right)_k$. In [16] it is also shown that one can take any strictly positive convex combination of the measures $(\mathbb{Q}_1,\dots,\mathbb{Q}_n)$. Below we will show that this sigma-algebra in some sense has a minimal property, a result that clarifies the relation between the two approaches. Before doing so, let us recall two easy results from introductory probability theory.

Result 1. For a given probability space $(\Omega,\mathcal{A},\mathbb{Q})$ let us denote $\mathcal{N}=\{N\in\mathcal{A}\mid \mathbb{Q}[N]=0\}$. Suppose that a sub sigma-algebra $\mathcal{F}\subset\mathcal{A}$ is given and that $\mathcal{G}$, $\mathcal{F}\subset\mathcal{G}$, is another sub sigma-algebra which is included in the sigma-algebra generated by $\mathcal{F}$ and $\mathcal{N}$. Then for each $\xi\in L^1(\Omega,\mathcal{A},\mathbb{Q})$

$$\mathbb{E}_{\mathbb{Q}}[\xi\mid\mathcal{F}]=\mathbb{E}_{\mathbb{Q}}[\xi\mid\mathcal{G}]\quad \text{a.s.}$$

Result 2. With the notation in the previous exercise let $F\colon\Omega\to\mathbb{R}^n$ and $F'\colon\Omega\to\mathbb{R}^n$ be two vectors that are equal a.s. . Let $\mathcal{F}$ be generated by F and $\mathcal{G}$ be generated by F'. Then $\mathcal{F}$ and $\mathcal{G}$ are equal up to sets in $\mathcal{N}$. More precisely $\mathcal{G}$ is included in the sigma-algebra generated by $\mathcal{F}$ and $\mathcal{N}$ (and of course conversely), i.e. $\sigma(\mathcal{F},\mathcal{N})=\sigma(\mathcal{G},\mathcal{N})$.

Proposition 3. *Let $\mathbb{Q}_1,\dots,\mathbb{Q}_n$ be probability measures on a measurable space $(\Omega,\mathcal{A})$. Let $\mathbb{Q}_0$ denote a convex combination of these measures $\mathbb{Q}_0=\sum_k\lambda_k\mathbb{Q}_k$ where each $\lambda_k>0$. Let f_k denote an $\mathcal{A}$ measurable version $\frac{d\mathbb{Q}_k}{\mathbb{Q}_0}$. Let $\mathbb{Q}$ be another dominating measure with g_k an $\mathcal{A}$ measurable version of $\frac{d\mathbb{Q}_k}{d\mathbb{Q}}$. Let $\mathcal{N}=\{N\in\mathcal{A}\mid\mathbb{Q}_0[N]=0\}$. Let $\mathcal{F}$ be generated by $f_k, k=1\dots n$ and let $\mathcal{G}$ be generated by $g_k, k=1\dots n$. Then $\mathcal{F}\subset\sigma(\mathcal{G},\mathcal{N})$*

Proof Clearly $\mathbb{Q}_0\ll\mathbb{Q}$ so let $h=\frac{d\mathbb{Q}_0}{d\mathbb{Q}}$. It is now immediate that $g_k=f_kh$ $\mathbb{Q}$ a.s. . To see this, observe that the values of f_k on $\{h=0\}$ do not matter. The functions g_k and h are $\mathcal{G}$ measurable since h can be taken as $h=\sum_k\lambda_kg_k$. Then we define $f'_k=\frac{g_k}{h}$ on $\{h>0\}$ and $f'_k=0$ on $\{h=0\}$. This choice shows that the f'_k are $\mathcal{G}$ measurable. It is immediate that $f_k=f'_k$ $\mathbb{Q}_0$ a.s. . The result now follows.

From the theorem it follows that the sigma-algebra augmented with the class $\mathcal{N}$ is the same for all strictly positive convex combinations. The theorem shows that in the definition of conditionally atomless with respect to $\mathcal{F}$, we can also add the null sets $\mathcal{N}$ to $\mathcal{F}$. To check that $\mathcal{A}$ is conditionally atomless with respect to a sigma-algebra $\mathcal{F}$ it is clear that the smaller $\mathcal{F}$, the easier it is to satisfy the condition. In my opinion the above clarifies the relation between this paper and [16].

3. A continuity Result

Let us recall the standing assumptions. $\mathcal{F}_2$ is conditionally independent of $\mathcal{F}_1$. U is independent of $\mathcal{F}_1$ and is uniformly distributed on $[0,1]$. The utility function $u_{1,2}\colon L^\infty(\mathcal{F}_2)\to L^\infty(\mathcal{F}_1)$ is coherent and is Lebesgue continuous. For each $h\colon\Omega\to[0,1]$ that is $\mathcal{F}_1$ measurable we put $\phi(h)=u_{1,2}(\mathbf{1}_{\{U\le h\}})$. Clearly ϕ takes values in the space $L^\infty(\mathcal{F}_1)$. We have the following continuity result.

Proposition 4. *If $h_n\downarrow h$ or $h_n\uparrow h$, we have $\phi(h_n)\to\phi(h)$.*

Proof If $h_n\downarrow h$ then $\mathbf{1}_{\{U\le h_n\}}\downarrow\mathbf{1}_{\{U\le h\}}$ and the Fatou property gives the desired result. For the upward convergence we must be more careful. Because U has a continuous distribution function and is independent of $\mathcal{F}_1$, we conclude that $\mathbb{P}[U=h]=0$ and hence $\mathbf{1}_{\{U\le h_n\}}\uparrow\mathbf{1}_{\{U\le h\}}$ a.s. . The Lebesgue property then allows to conclude.

Theorem 4. *If $h\colon \Omega \to [0,1]$ is $\mathcal{F}_1$ measurable, there is an $\mathcal{F}_1$ measurable function $g\colon \Omega \to [0,1]$ such that the set $B_g = \{U \le g\}$ satisfies $u_{1,2}(\mathbf{1}_{B_g}) = h$.*

Proof The statement can be rewritten as $\phi(g) = h$. Let us introduce the class

$$\mathcal{G} = \{g \mid g \text{ is } \mathcal{F}_1 \text{ measurable and } u_{1,2}(\mathbf{1}_{B_g}) = \phi(g) \ge h\}.$$

$\mathcal{G}$ is nonempty since $1 \in \mathcal{G}$. Furthermore $\mathcal{G}$ is stable for taking the minimum. Indeed, let $g_1, g_2 \in \mathcal{G}$ and put $g = g_1\mathbf{1}_A + g_2\mathbf{1}_{A^c}$ where $A = \{g_1 < g_2\}$. Since $u_{1,2}(\mathbf{1}_{B_g}) = u_{1,2}(\mathbf{1}_{B_{g_1}})\mathbf{1}_A + \mathbf{1}_{A^c}u_{1,2}(\mathbf{1}_{B_{g_2}}) \ge h$ we have that $g \in \mathcal{G}$. Let now $g_n \downarrow g$ where $g_n \in \mathcal{G}$ and $\mathbb{E}[g_n] \downarrow \inf\{\mathbb{E}[g'] \mid g' \in \mathcal{G}\}$. The continuity for decreasing sequences then shows that $g \in \mathcal{G}$. The previous lines are enough to show that $\mathcal{G}$ has a minimum. Let g be the smallest function in $\mathcal{G}$. The continuity for increasing sequences (the Lebesgue property) will show that actually $u_{1,2}(\mathbf{1}_{B_g}) = h$. Suppose on the contrary that the set $\{u_{1,2}(\mathbf{1}_{B_g}) > h\}$ has non zero measure. This assumption trivially implies that $\mathbb{P}[g > 0] > 0$. Take now a sequence $g_n \uparrow g$ such that on $\{g > 0\}$, $g_n < g$. By the previous theorem $u_{1,2}(\mathbf{1}_{B_{g_n}}) \uparrow u_{1,2}(\mathbf{1}_{B_g})$. Hence, there must exist n such that $A_n = \{u_{1,2}(\mathbf{1}_{B_{g_n}}) > h\}$ has nonzero measure. On A_n, we have $g_n > 0$ hence also $g > 0$ and therefore also $g_n < g$. Put now $g' = g_n\mathbf{1}_{A_n} + g\mathbf{1}_{A_n^c}$. We have $\mathbb{E}[g'] < \mathbb{E}[g]$ but also $g' \in \mathcal{G}$ a contradiction to the minimality of g.

Remark 5. Although “intuitively clear", the continuity of the process $u_{1,2}(\mathbf{1}_{B_t})$ is not an easy result. First of all, we are working with random variables identified under the equivalence a.s. . That means that we must first select or construct measurable functions instead of classes of measurable functions. Then we must show that with respect to t these outcomes are continuous. The general theory of stochastic processes gives us the necessary tools to achieve this goal. We do not really need these finer results so if you do not belong to the amateurs of the general theory of stochastic processes, the remark can be skipped, see [10] for the necessary details. First we will construct a process $\alpha(t,\omega)$. For each rational point $q \in [0,1]$ we select an $\mathcal{F}_1$ measurable function $\alpha'(q)$ that represents $u_{1,2}(\mathbf{1}_{B_q})$. Because of monotonicity we can – if needed – change these selections on a set of zero measure, to make sure that a.s. the mapping $\mathbb{Q} \cap [0,1] \to \mathbb{R}; q \to \alpha'(q)$ is increasing. For each $t \in [0,1]$ we now define $\alpha(t) = \inf_{q \text{ rational },q\ge t} \alpha'(q)$. The functions $\alpha(t)$ are of course $\mathcal{F}_1$ measurable and represent $u_{1,2}(\mathbf{1}_{B_t})$ by the Fatou property. We may also suppose that $\alpha(0) = 0, \alpha(1) = 1$ a.s. . It is clear that α is a.s. nondecreasing in t and is right continuous. This means there is a set (independent of t) such that on this set $t \to \alpha(t,\omega)$ is right continuous and nondecreasing.

The function α also satisfies $\alpha(h) = u_{1,2}(\mathbf{1}_{\{U\le h\}}) = \phi(h)$ for each $\mathcal{F}_1$ measurable function $h\colon \Omega \to [0,1]$. [6] The statement is easy to verify for elementary functions h and the general statement trivially follows by approximating h from *above* by elementary functions. Let us give the details. For en elementary function $h = \sum_{k=1}^K t_k\mathbf{1}_{A_k}$ (the sets A_k are disjoint and in

[6]To avoid misunderstandings the random variable $\alpha(h)$ is defined as $\alpha(h)(\omega) = \alpha(h(\omega),\omega)$. Such a practice is common in stochastic process theory.

$\mathcal{F}_1$), we have

$$\begin{aligned}
\alpha(h) &= \sum_k \alpha(t_k)\mathbf{1}_{A_k} \\
&= \sum_k u_{1,2}(\mathbf{1}_{B_{t_k}})\mathbf{1}_{A_k} \\
&= \sum_k u_{1,2}(\mathbf{1}_{B_{t_k}}\mathbf{1}_{A_k})\mathbf{1}_{A_k} \\
&= \sum_k u_{1,2}\left(\mathbf{1}_{B_{t_k}\cap A_k}\right)\mathbf{1}_{A_k} \\
&= \sum_k u_{1,2}\left(\left(\sum_l \mathbf{1}_{B_{t_l}\cap A_l}\right)\mathbf{1}_{A_k}\right)\mathbf{1}_{A_k} \\
&= \sum_k u_{1,2}(\mathbf{1}_{\{U\le h\}}\mathbf{1}_{A_k})\mathbf{1}_{A_k} \\
&= u_{1,2}(\mathbf{1}_{\{U\le h\}}) = \phi(h).
\end{aligned}$$

As indicated above, the Fatou property then completes the proof using right continuity. Indeed, let $h\colon \Omega \to [0,1]$ be $\mathcal{F}_1$ measurable and let $h_n \downarrow h$ be a sequence of elementary functions, that are $\mathcal{F}_1$ measurable. Since $\mathbf{1}_{\{U\le h_n\}} \downarrow \mathbf{1}_{\{U\le h\}}$, the Fatou property and the right continuity of $\alpha(t)$ give us $\phi(h) = u_{1,2}(\mathbf{1}_{\{U\le h\}})$.

The proof of the left continuity can be done using ideas from the general theory of stochastic processes. For $\varepsilon > 0$ we define

$$h = \inf\{t \mid \lim_{s\to t; s<t} \alpha(s) \le \alpha(t) - \varepsilon\} \wedge 1.$$

Observe that by construction $h > 0$. Suppose now that at the point h, the probability that α has a jump of size at least ε is nonzero. Take $h_n \uparrow h$; $h_n < h$. The continuity result gives us that $\alpha(h_n) \uparrow \alpha(h)$ which is a contradiction to α having a jump. So for almost every $\omega \in \Omega$, $\alpha(., \omega)$ has no jumps of size at least ε. Since the latter was arbitrary, the a.s. continuity of the process α is proved.

4. Some Special commonotonic Set

In this section we will define a special norm on $\mathbb{R}^2$. Part of its unit sphere will then be used as a commonotonic set. The reader could make some drawings to help visualise the constructions. The construction will be done in several steps. The first step consists in taking the curve obtained as the concatenation of the convex intervals that join the points

$$(-4,-4) \to (-4,-2) \to (0,0) \to (4,2) \to (4.4).$$

The convex hull of this set is a parallelogram P_0, with parallel vertical sides given by the line segments

$$((-4,-4) \to (-4,-2) \text{ and } (4,2) \to (4,4).$$

The set P_0 will be used as the unit ball of a norm on $\mathbb{R}^2$. More precisely we use the Minkowski functional:

$$\|(x,y)\| = \inf\{\alpha \mid \alpha > 0, (x,y) \in \alpha P_0\}.$$

Note that every point of P_0 is the convex combination of points taken on the vertical sides. An easy and continuous way to obtain such convex combination goes as follows. Through a point in P_0 take a line parallel to the “skew" sides of P_0 and see where it intersects the vertical sides. Elementary calculations give us that for $(x, y) \in P_0$ we may write $(x, y) = (1 - \lambda_0)(u_1^0, u_2^0) + \lambda_0(v_1^0, v_2^0)$ or

$$(x, y) = \frac{4 - x}{8}\left(-4, y - 3 - \frac{3x}{4}\right) + \frac{4 + x}{8}\left(4, y + 3 - \frac{3x}{4}\right).$$

For each $n \in \mathbb{Z}$ we now define $P_n = 2^n P_0$ and similarly as for $n = 0$ we define λ_n, (u_1^n, u_2^n), (v_1^n, v_2^n). These functions are obviously continuous. The set E consists of all the vertical segments with the origin added. It forms a commonotonic set. This follows from the equality

$$E = \{(0, 0)\} \cup \cup_{n\in\mathbb{Z}} \left(2^n \left([(-4, -4), (-4, -2)] \cup [(2, 4), (4, 4)]\right)\right).$$

We now define functions Λ, U, V on $\mathbb{R}^2$. For $(x, y) \in P_n \setminus P_{n-1}$ we define $\Lambda(x, y) = \lambda_n(x, y), U(x, y) = u^n(x, y), V(x, y) = v^n(x, y)$. We also put $\Lambda(0, 0) = 1, U(0, 0) = (0, 0) = V(0, 0)$. These functions are no longer continuous but are certainly Borel measurable. They satisfy:

(1) $\Lambda \colon \mathbb{R}^2 \to [0, 1]$
(2) $U \colon \mathbb{R}^2 \to E$, $V \colon \mathbb{R}^2 \to E$
(3) We have $\|U(x, y)\| \le 2\|(x, y)\|$ and $\|V(x, y)\| \le 2\|(x, y)\|$. Indeed for $(x, y) \in P_n \setminus P_{n-1}$ we have $2^n = \|U(x, y)\| \ge \|(x, y)\| \ge 2^{n-1}$ and the same for V.
(4) For all $(x, y) \in \mathbb{R}^2$, $(x, y) = (1 - \Lambda(x, y))U(x, y) + \Lambda(x, y)V(x, y)$.
(5) The coordinates of $V - U$, $V_1(x, y) - U_1(x, y)$ and $V_2(x, y) - U_2(x, y)$, are nonnegative.

5. The main Result

We start by giving an extension of the usual definition of conditional expectation.

Definition 3. We say that an $\mathcal{F}_2$ measurable random variable, ξ, has an *extended* conditional expectation with respect to $\mathcal{F}_1$ if there is a countable $\mathcal{F}_1$ measurable partition, $(A_n)_n$, such that each $\mathbf{1}_{A_n}\xi$ is integrable. The conditional expectation is then defined as $\sum_n \mathbb{E}[\mathbf{1}_{A_n}\xi \mid \mathcal{F}_1]$.

The reader can check that the existence and the definition of an extended conditional expectation is independent of the choice of the $\mathcal{F}_1$ measurable partition. We will sometimes drop the word *extended*.

Again we suppose that $\mathcal{F}_2$ is conditionally atomless with respect to $\mathcal{F}_1$. The utility function $u_{1,2}$ is Lebesgue continuous.

Before giving the main result of the paper we first prove a special case.

Theorem 5. *For every couple (f, g) of $\mathcal{F}_1$ measurable finitely valued random variables, there is a commonotonic couple (ξ, η) of $\mathcal{F}_2$ measurable random variables such that $f = \mathbb{E}[\xi \mid \mathcal{F}_1], g = \mathbb{E}[\eta \mid \mathcal{F}_1]$. Furthermore $\|(\xi, \eta)\| \le 2\|(f, g)\|$ almost surely.*

Proof The proof is almost given in the previous sections. Let $(f, g) \colon \Omega \to \mathbb{R}^2$ be $\mathcal{F}_1$ measurable. Using the functions Λ, U, V of the previous section we can then write

$$(f, g) = \Lambda(f, g)V(f, g) + (1 - \Lambda(f, g))U(f, g).$$

Because $\Lambda(f,g):\Omega\to[0,1]$ is $\mathcal{F}_1$ measurable and because $\mathcal{F}_2$ is conditionally atomless with respect to $\mathcal{F}_1$, there is an $\mathcal{F}_2$ measurable set B such that $\mathbb{E}[\mathbf{1}_B\mid\mathcal{F}_1]=\Lambda(f,g)$. The random variables (ξ,η) are now defined as

$$\xi=\mathbf{1}_B V_1(f,g)+\mathbf{1}_{B^c}U_1(f,g) \text{ and } \eta=\mathbf{1}_B V_2(f,g)+\mathbf{1}_{B^c}U_2(f,g),$$

in other words

$$(\xi,\eta)=\mathbf{1}_B V(f,g)+\mathbf{1}_{B^c}U(f,g).$$

Both random variables have in an extended sense, conditional expectations and because $U(f,g),V(f,g)$ are $\mathcal{F}_1$ measurable we get $(f,g)=\mathbb{E}[(\xi,\eta)\mid\mathcal{F}_1]$. Because (ξ,η) takes its values in the commonotonic set E (introduced above) we get that ξ and η are commonotonic. The estimate of the norms follow from the estimates for U and V.

Corollary 1. *The random variable (ξ,η) has the same integrability properties as the couple (f,g). In particular if (f,g) is bounded, the couple (ξ,η) is bounded.*

Remark 6. In case one wants to use another norm than the Minkowski functional of P_0, one must adapt the constant. Because all norms on $\mathbb{R}^2$ are equivalent this is an exercise in linear algebra. I did not try to find the best estimates for e.g. the Euclidean norm, where a rough calculation gave $10\sqrt{2}$. This problem would require to find a better commonotonic set than the one used above.

The next theorem is an improvement of the preceding result in the sense that we replace the conditional expectation by a more general utility function. The proof follows the same lines.

Theorem 6. *For every couple (f,g) of $\mathcal{F}_1$ measurable bounded valued random variables, there is a commonotonic couple (ξ,η) of $\mathcal{F}_2$ measurable random variables such that $f=u_{1,2}(\xi),g=u_{1,2}(\eta)$. Furthermore $\|(\xi,\eta)\|\le 2\|(f,g)\|$ almost surely.*

Proof We use the same notation (Λ,U,V) as in the previous proof. But this time we take a set B such that $u_{1,2}(\mathbf{1}_B)=\Lambda$. Again we define

$$(\xi,\eta)=\mathbf{1}_B V(f,g)+\mathbf{1}_{B^c}U(f,g)=U(f,g)+\mathbf{1}_B(V(f,g)-U(f,g)).$$

We then have:

$$\begin{aligned}u_{1,2}(\xi)&=u_{1,2}(U_1(f,g)+\mathbf{1}_B(V_1(f,g)-U_1(f,g))\\&=U_1(f,g)+u_{1,2}(\mathbf{1}_B)(V_1(f,g)-U_1(f,g))\\&=U_1(f,g)+\Lambda(f,g)(V_1(f,g)-U_1(f,g))=f,\end{aligned}$$

and similarly for g and the second coordinates. Remark that we could apply the positive homogeneity of $u_{1,2}$ because $(V_1(f,g)-U_1(f,g))\ge 0$.

Remark 7. If (f,g) is only finitely valued, we can write

$$(f,g)=\mathbf{1}_{\{(f,g)=(0,0)\}}(f,g)+\sum_{n\in\mathbb{Z}}\mathbf{1}_{\{(f,g)\in P_n\setminus P_{n-1}\}}(f,g)$$

and this is a sum of bounded random variables. For each n we can define ξ_n,η_n as in the theorem. These random variables are zero outside $\{(f,g)\in P_n\setminus P_{n-1}\}$ and hence the sum $(\xi,\eta)=\sum_{n\in\mathbb{Z}}(\xi_n,\eta_n)$ is defined. We could then extend $u_{1,2}$ as we did for conditional expectations. Finally we get $u_{1,2}(\xi)=f,u_{1,2}(\eta)=g$. This extension is important when the utility functions are defined on e.g. Orlicz spaces or Riesz spaces. Important for such extensions is the pointwise (almost surely) estimate $\|(\xi,\eta)\| \le 2\|(f,g)\|$.

6. Commonotonicity and Time Consistency

In this section we use the same hypothesis on the filtration $(\mathcal{F}_0,\mathcal{F}_1,\mathcal{F}_2)$. In particular we suppose that $\mathcal{F}_2$ is atomless conditionally to $\mathcal{F}_1$. We start with a monetary coherent utility function $u_{0,2}\colon L^\infty(\mathcal{F}_2)\to\mathbb{R}$. We suppose – as in the rest of the paper – that $u_{0,2}$ is relevant.

Theorem 7. *Suppose*

1. *$\mathcal{F}_2$ is atomless conditionally to $\mathcal{F}_1$*
2. *$u_{0,2}$ is coherent and relevant*
3. *$u_{0,2}$ is time consistent*
4. *$u_{0,2}$ is commonotonic, i.e. if $\xi,\eta\in L^\infty(\mathcal{F}_2)$ are commonotonic, then $u_{0,2}(\xi+\eta) = u_{0,2}(\xi)+u_{0,2}(\eta)$*
5. *$u_{0,2}$ is Lebesgue continuous.*

Then there is a probability $\mathbb{Q}\sim\mathbb{P}$ such that for all $f\in L^\infty(\mathcal{F}_1)$ we have $u_{0,1}(f)=\mathbb{E}_{\mathbb{Q}}[f]$.

Proof According to the previous section, for each $f,g\in L^\infty(\mathcal{F}_1)$ there are *commonotonic* $\xi,\eta\in L^\infty(\mathcal{F}_2)$ with $u_{1,2}(\xi)=f$, $u_{1,2}(\eta)=g$ and $u_{1,2}(\xi+\eta)=f+g$. We then have $u_{0,1}(f)=u_{0,1}\left(u_{1,2}(\xi)\right)=u_{0,2}(\xi)$ and similarly for g. The combination with commonotonicity then gives

$$\begin{aligned} u_{0,1}(f+g) &= u_{0,1}(u_{1,2}(\xi+\eta)) \\ &= u_{0,2}(\xi+\eta) \\ &= u_{0,2}(\xi)+u_{0,2}(\eta) \\ &= u_{0,1}(u_{1,2}(\xi))+u_{0,1}(u_{1,2}(\eta)) \\ &= u_{0,1}(f)+u_{0,1}(g) \end{aligned}$$

This shows that $u_{0,1}$ is additive (therefore linear) and hence is given by a finitely additive probability measure. But Lebesgue continuity implies that this measure, say $\mathbb{Q}$, should be sigma additive and absolutely continuous with respect to $\mathbb{P}$. Because $u_{0,2}$ and hence $u_{0,1}$ are relevant we must have $\mathbb{Q}\sim\mathbb{P}$.

Remark 8. For ξ,η commonotonic (and not just for the ones used in the proof of the theorem) we can now prove that $u_{1,2}(\xi+\eta)=u_{1,2}(\xi)+u_{1,2}(\eta)$. We already know that $u_{1,2}(\xi+\eta)\ge u_{1,2}(\xi)+u_{1,2}(\eta)$. If $\mathbb{Q}[u_{1,2}(\xi+\eta)>u_{1,2}(\xi)+u_{1,2}(\eta)]>0$, then we have

$$\begin{aligned} u_{0,2}(\xi+\eta) &= u_{0,1}(u_{1,2}(\xi+\eta)) = \mathbb{E}_{\mathbb{Q}}\left[u_{1,2}(\xi+\eta)\right] \\ &> \mathbb{E}_{\mathbb{Q}}\left[u_{1,2}(\xi)\right]+\mathbb{E}_{\mathbb{Q}}\left[u_{1,2}(\eta)\right] = u_{0,1}(u_{1,2}(\xi))+u_{0,1}(u_{1,2}(\eta)) \\ &= u_{0,2}(\xi)+u_{0,2}(\eta) \end{aligned}$$

which is a contradiction to $u_{0,2}(\xi+\eta)=u_{0,2}(\xi)+u_{0,2}(\eta)$. The strict inequality on the second line follows from the fact that $u_{0,1}$ is the expectation with respect to the equivalent probability measure $\mathbb{Q}$.

Remark 9. If the assumption of relevancy is dropped, we must start with a time consistent system of utility functions $u_{0,2},u_{0,1},u_{1,2}$. In that case we only have that $\mathbb{Q}\ll\mathbb{P}$ and the result of the previous remark only holds $\mathbb{Q}$ a.s.

Remark 10. There is no reason that $u_{0,2}$ is additive on $L^\infty(\mathcal{F}_2)$ as the following example shows. We take $\Omega = [0,1] \times [0,1]$, $\mathcal{F}_2$ is the product sigma algebra of the Borel sigma algebas on $[0,1]$, the measure $\mathbb{P}$ is the product measure of the usual Lebesgue measures. $\mathcal{F}_0$ is the trivial sigma algebra and $\mathcal{F}_1$ is generated by the first coordinate mapping. For $\xi \in L^\infty(\mathcal{F}_2), \xi \geq 0$ we define

$$u_{0,2}(\xi) = \int_0^1 d\alpha \int_0^\infty dx\, \mathbb{P}[\xi(\alpha,.) \geq x]^{1+\alpha}.$$

For $0 \leq \xi \in L^\infty(\mathcal{F}_2)$ the utility function $u_{1,2}$ is then given by

$$u_{1,2}(\xi)(\alpha) = \int_0^\infty \mathbb{P}[\xi(\alpha,.) > x]^{1+\alpha}\, dx.$$

Such expressions are known as distortions or Choquet integrals. They are standard examples of commonotonic utility functions, see [7]. We need a little bit less than commonotonicity, in fact we only need that for ξ, η we have $u_{1,2}(\xi+\eta) = u_{1,2}(\xi) + u_{1,2}(\eta)$ as soon as for each α the random variables $\xi(\alpha,.), \eta(\alpha,.)$ are commonotonic. To see that $u_{0.2}$ is not linear let us calculate the outcomes for $\xi(\alpha,y) = \mathbf{1}_{[0,1/2]}(y)$ and $\eta(\alpha,y) = \mathbf{1}_{[1/2,1]}(y)$. For both random variables we find $\frac{1}{4\log(2)}$ which do not sum up to $u_{0,2}(\xi+\eta) = u_{0,2}(1) = 1$.

7. A continuous time result

In this section we use a filtration indexed by the time interval $[0,T]$. This filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ does not necessarily fulfil the usual assumptions. The only assumption is that $\mathcal{F}_T$ is generated by $\cup_{0\leq t<T}\mathcal{F}_t$. We also suppose that a family of coherent utility functions $u_{t,s}, 0 \leq t \leq s \leq T$, $u_{t,s}\colon L^\infty(\mathcal{F}_s) \to L^\infty(\mathcal{F}_t)$ is given. We assume the following time consistency: for $t \leq s \leq v$ we have $u_{t,v} = u_{t,s} \circ u_{s,v}$.

Theorem 8. *We assume the notation introduced in this section. We suppose that for $0 \leq t < T$, the sigma algebra $\mathcal{F}_T$ is atomless conditionally to $\mathcal{F}_t$. If $u_{0,T}$ is relevant, Lebesgue continuous and commonotonic then there is a probability $\mathbb{Q} \sim \mathbb{P}$ such that for all $\xi \in L^\infty(\mathcal{F}_T)$: $u_{0,T}(\xi) = \mathbb{E}_\mathbb{Q}[\xi]$.*

Proof The results of the previous section show that on each $L^\infty(\mathcal{F}_t)$, the utility function $u_{0,T}$ is linear. The utility function $u_{0,T}$ is therefore linear on the vector space $\cup_{t<T}L^\infty(\mathcal{F}_t)$. This space is sequentially dense in $L^\infty(\mathcal{F}_T)$ for the Mackey topology (simply use the martingale convergence theorem). Because of Lebesgue continuity, the utility function $u_{0,T}$ is therefore linear on $L^\infty(\mathcal{F}_T)$. It is therefore given by a probability measure $\mathbb{Q} \ll \mathbb{P}$. But since the utility function is relevant we find that $\mathbb{Q} \sim \mathbb{P}$.

Remark 11. The previous results can be applied for most filtrations used in finance and insurance. This is for instance true for filtrations of Brownian Motion in one or several dimensions, filtrations generated by most Lévy processes and so on. In other words *commonotonicity and time consistency are not good friends.*

References

[1] Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath: Thinking Coherently, *RISK*, November 97, 68–71, (1997)

[2] Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath: Characterisation of Coherent Risk Measures, *Mathematical Finance* **9**, 145–175, (1999)

[3] Bellini,F., Bignozzi, V., Delbaen, F., Ziegel, J.: Risk Measures with the CxLS property, *Finance and Stochastics*, **18**, 433–453 (2014)

[4] Bellini, F. Koch-Medina, P., Munari, C., Svindland, G.: Law-invariant insurance pricing and its limitations. *arxiv 1808.00821* (2018)

[5] Castagnoli, E., Maccheroni, M., Marinacci, M.: Choquet Insurance Pricing: a caveat. *Mathematical Finance,* **14**, 481–485 (2004)

[6] Delbaen, F: Coherent Risk Measures, *Lectures given at the Cattedra Galileiana at the Scuola Normale Superiore di Pisa, March 2000*, Published by the *Scuola Normale Superiore di Pisa*, (2002)

[7] Delbaen, F: Monetary Utility Functions, *Lectures held in 2008 and published in the series "Lecture Notes of the University of Osaka"*, (2011)

[8] Delbaen, F.: Conditionally Atomless Extensions of Sigma Algebras, *arxiv 2003.09254, (2020)*

[9] Denneberg, D.: Verzerrte Wahrscheinlichkeiten in der Versicherungsmathematik, quantilabhängige Prämienprinzipen. *Mathematik-Arbeitspapiere*, **34**, 1989, Universität Bremen

[10] Dellacherie, C., Meyer, P.-A.: Probabilités et potentiel, Chapitres I à IV, Chapitres V à VIII, 1980, *Hermann, Paris*

[11] Kupper, M., Schachermayer, W.: Representation results for law invariant time consistent functions, *Mathematics and Financial Economics* **2** (3), 189–210

[12] Koopmans, T. C.: Stationary ordinal utility and impatience. *Econometrica*, **28**, 287 – 309, (1960)

[13] Lorenc, P., Witula, R.: Darboux Property of the Nonatomic sigma-additive Positive and Finite dimensional Vector Measures, Zeszyty Naukowe Politechniki Šlskiej, Serie Matematyka Stosowana 3, 25–36 (2013)

[14] Schmeidler, D.: Integral Representation without Additivity, *Proc. Amer. Math. Soc.*, **97**, 255–261, (1986)

[15] Sierpiński, W.: Sur les fonctions d'ensemble additives et continues, Fund. Math. 3, 240–246 (1922)

[16] Shen, J., Shen, Y., Wang, B., and Wang, R.: *Distributional Compatibility for Change of Measures*, https://arxiv.org/abs/1706.01168, to be published in *Finance and Stochastics*

[17] Wang, S.S., Young, V.R. and Panjer, H.: Axiomatic Characterization of Insurance Prices. *Insurance: Mathematics and Economics* 21:173–183 (1997)

[18] Wang, T.: A Class of Multi-Prior Preferences. *working paper*, (2002), http://www.ccfr.org.cn/english/N&E&R/research.php

[19] Wang, R., Ziegel,J.: Elicitable distortion risk measures: A concise proof. *Statistics and Probability Letters,* **100** 172–175 (2015)

[20] Weber, S.: Distribution-invariant risk measures, information and dynamic consistency. *Mathematical Finance,* **16**, 419–441 (2006)

[21] Yaari, M. E.: The dual theory of choice under risk. *Econometrica*, **55**, 95–115, (1987)

[22] Ziegel, J.: Coherence and Elicitability *Mathematical Finance,* **26**, 901 – 918 (2016)

Departement für Mathematik, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland