Characterization of Time-Consistent Sets of Measures in Finite Trees

Monika Bier and Frank Riedel
Institute of Mathematical Economics
Bielefeld University

Preliminary Draft

March 30, 2009

Abstract

In this paper we give an alternative characterization for time-consistent sets of measures in a discrete setting. For each measure $\mathbb{P}$ in a time-consistent set $\mathcal{P}$ we get a distinct set of predictable processes which in return describe the $\mathbb{P}$ uniquely. This implies we get a one-to-one correspondence between time-consistent sets of measures and sets of predictable processes with specific features.

1 Introduction

In 1944 von Neumann and Morgenstern in [vNM] formulated their famous axioms for preferences over uncertain payoffs and showed that these preferences are equivalent to an Expected Utility Representation of preferences. Later on their model was critized because the distributions of their payoffs were exogeneously given and purely objective. Since this is a very restrictive assumption their model was extended by Savage (cp. [Sav]) and Anscombe and Aumann (cp. [AA]). In contrast to the von Neumann and Morgenstern model Savage regarded the distributions of the payoffs to be purely subjective and endogeneous. Anscombe and Aumann then combined both models taking some objective distributions as given and having others arising purely out of the model.

At some point criticism also arose against these models. One of the most mentioned objections came from Ellsberg (cp. [Ell]). He conducted experiments and empirically showed that Expected Utility models do not always mirror reality. One way of explaining these findings is that people behave only boundedly rational. Another way is to distinguish between uncertainty and risk. While in a risky setting the DM is sure of the distributions of the outcomes in an uncertain setting he is unsure of the right distribution and thinks more than one possible. Following this idea Gilboa and Schmeidler developed their Multiple Priors Model in [GS] using Anscombe's and Aumann's model as a basis. They weakened the Independence Axiom and added an additional axiom formalizing Uncertainty Aversion. This lead the DM to maximize $\inf_{\mathbb{P}\in\mathcal{C}} \mathbb{E}^{\mathbb{P}}[u \circ f]$ among all possible acts f, where $\mathcal{C}$ is a nonempty, closed and convex set of probability measures.

Since this is a purely atemporal model in [ES] Epstein and Schneider expanded the Multiple Priors Model to incorporate the factor time. They modified preferences to be not only state but also time-dependent, adjusted the G-S-axioms appropiately and asked for Dynamic Consistency as an additional axiom. This restriction on preferences yields a very specific property of the set of measures in their Utility Representation. They found out that preferences are dynamically consistent if and only if the set of measures in their Recursive Multiple Priors Representation is rectangular. Rectangularity is a restriction on the whole set of measures. It demands that it is possible for the one-step-ahead measures to be mixed arbitrarily. Since for some purposes (e.g. solving concrete optimal stopping problems) this is not a very easy definition but never the less an important one it is very natural to try and find equivalent definitions.

This was done by different authors. In [Rie2] Riedel gives a survey of the different concepts and shows their equivalence. Among these concepts is rectangularity which was introduced by Epstein and Schneider in [ES] and is a property concerning the one-step-ahead measures. They asked that at every point in time all possible one step ahead measures can be added. Another concept is stability. It was introduced by Föllmer and Schied (cp. [FS]). Here for two measures $\mathbb{P}$ and $\mathbb{Q}$ in the set of measures and every stopping time τ the measure that takes $\mathbb{P}$ up to τ and $\mathbb{Q}$ afterwards also lies in the set. The last concept is time-consistency which was introduced by Delbaen in [Del]. This property demands that at every stopping time density processes can be consistently pasted together. A more formal definition of this specific property will be given in the next section.

In the above cited paper Riedel among other things constructed time-consistent sets of measures via their density processes. Consequently the question arose if in this special setting all time-consistent sets of measures can be constructed in this way. That is why we took a closer look at time-consistent sets of measures and found out that not quite all sets are of this kind. However a slight modification of his construction does the trick.

The main content of this paper is this alternative characterization of time-consistent sets. They are described via a set of predictable processes with specific properties. This will be our first and main theorem. In addition to showing how the set of measures can be related to this set of processes we will also show that sets of processes with the assumed properties define sets of time-consistent measures. This will be the content of our second theorem. So altogether we will provide an equivalent formulation for time-consistent sets of measures.

The build-up of this paper will be the following. After pinning down the model framework and specifying the attributes of our sets more precisely in section 2 we will deduct the first theorem in the succeeding section 3. Then in section 4 we will commit ourselves to proving the second theorem. In the following fifth section we will introduce some example setting where our results are applicable an might simplify calculations. After that we discuss possible extensions in section 6 and then conclude in the last and seventh section.

2 Model

To specify the setting we start with a discrete set $\Omega = \{\omega_1, ..., \omega_k\}$. On this state space we have an information structure $\{\mathbf{F}_t\}_{t=0,...,T}$ with $\mathbf{F}_0 = \Omega$ and $\mathbf{F}_T = \{\{\omega_1\}, ..., \{\omega_k\}\}$. This is a sequence of partitions of Ω, which become finer as time progresses, i.e. every set of $\mathbf{F}_{t+1}$ is a subset of some set of $\mathbf{F}_t$ for all t.

Heuristically this concept describes the information of the prevailing state availible at a certain time t. This means for a fixed time t the decision maker (DM) will not necessarily be able to observe the exact state which occurs but merely which subset of $\mathbf{F}_t$ is realized. If the observed subset consists of only a single state then of course the DM has full knowledge of the realization.

If you want to express this in terms of σ-fields and filtrations you just take the power set $\mathcal{P}ot(\Omega)$ for the filtration $\mathcal{F}$ and define the filtration $\{\mathcal{F}_t\}_t$ by setting $\mathcal{F}_t := \sigma(\mathbf{F}_t)$ i.e. $\mathbf{F}_t$ is the set of atoms generating $\mathcal{F}_t$.

For our considerations we assume our information structure to have a constant and finite splitting function with splitting value ν. This implies that if you draw the filtration as an information tree it will have the same finite number of branches at every vertex. Formally the splitting function f of an information structure $\{\mathbf{F}_t\}_t$ is defined in the following way

$$f : \Omega \times [0, \infty) \to \mathbb{N} \quad , \quad f(\omega, t) = \sharp\{A \in \mathbf{F}_{t+1} \mid A \subseteq \mathbf{F}_t(\omega)\}$$

where $\mathbf{F}_t(\omega)$ is the set $B \in \mathbf{F}_t$ with $\omega \in B$. The finiteness of this index will allow us to apply the martingale representation given in Theorem 5.15 in [Dot] and the constancy will result in unique processes in the representation. We will make these two things more precise in the following section.

For now we will also restrict this model to a finite time horizon $[0, T]$. The finite splitting index and the finite time horizon result in a finite Ω.

To complete our probability space we still need to fix a probability measure $\mathbb{P}_0$ as a reference measure which pins down the sets of measure zero. Since we are on a tree like structure any measure which assigns non-zero probability to each branch will do, for simplicity let us choose the uniform distribution.

The set of measures we want to characterize will be denoted by $\mathcal{P}$. In the following we will make some assumptions on this set and justify their plausibility.

Our first assumption will be

Assumption 2.1. *We assume $\mathbb{P}_0 \in \mathcal{P}$ and for all other measures $\mathbb{P} \in \mathcal{P}$ $\mathbb{P}(A) > 0$ for all $A \in \mathbf{F}_T$.*

In this assumption $\mathbb{P}_0$s function as a reference measure becomes clear. One can see that it has no influence on the stochastic structure of the other measures. It simply implies that all measures contained in $\mathcal{P}$ have the same null sets which means that we know what sure and impossible events are.

An economic interpretation of this assumption was given by Epstein and Marinacci in [EM]. They related it to an axiom on preferences first postulated by Kreps in [Kr]. He claimed that if a DM is ambivalent between an act x and $x \cup x'$ then he should also be ambivalent between $x \cup x''$ and $x \cup x' \cup x''$. Meaning if the possiblity of choosing x' in addition to x brings no extra utility compared to just being able to choose x, then also no additional utility should arise from being able to choose x' supplementary to $x \cup x''$.

In our second assumption we claim

Assumption 2.2. *The family of densities for a fixed t*

$$\mathcal{D}_t := \left\{ \frac{d\mathbb{P}}{d\mathbb{P}_0}\Big|_{\mathcal{F}_t} \mid \mathbb{P} \in \mathcal{P} \right\}$$

is compact in $L^1(\Omega, \mathcal{F}, \mathbb{P}_0)$.

First let us remark that in our setting this assumption is always satisfied since our Ω is finite. However since this is an assumption often made in other settings we mention it here. Besides if we extend our model to an infinite horizon, our Ω need no longer be finite and we will need to revert to this assumption.

Technically this assumption ensures that when looking at expressions of the following kind $\inf_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}}[X_\tau]$ the infimum is always attained for bounded stopping times τ. (cp. [Rie2])

Economically this property results from a feature of preferences already claimed by Arrow in [Arr] and related to this assumption by Chateauneuf et al. in [CMMT]. The condition we need to ask of the preferences to obtain this feature is called Monotone Continuity. It means that if an act f is preferred over an act g then a consequence x is never that bad that there is no small p such that x with probability p and f with probability $(1-p)$ is still preferred over g. The same is true for good consequences mixed with g.

Critics tend to object to this assumption by saying that if the probability of dying is added to the better act f then surely the preferences have to be reversed. However if we take f for getting 100 dollars and g for getting nothing then having to drive 60 miles to get the 100 dollars and so adding a small probability of getting killed will normally not reverse the preferences.

Expressed formally this means for acts $f \succ g$, a consequence x and a sequence of events $\{E_n\}_{n\in\mathbb{N}}$ with $E_1 \supseteq E_2 \supseteq ...$ and $\cap_{n\in\mathbb{N}} E_n = \emptyset$ there exists an $\bar{n} \in \mathbb{N}$ such that

$$\begin{bmatrix} x \text{ if } s \in E_{\bar{n}} \\ f(s) \text{ if } s \notin E_{\bar{n}} \end{bmatrix} \succ g \quad \text{and} \quad f \succ \begin{bmatrix} x \text{ if } s \in E_{\bar{n}} \\ g(s) \text{ if } s \notin E_{\bar{n}} \end{bmatrix}.$$

Our third and last assumption on $\mathcal{P}$ will be

Assumption 2.3. *$\mathcal{P}$ is time-consistent. This means for a stopping time τ and densities $p_t := \left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)_t$ and $q_t := \left(\frac{d\mathbb{Q}}{d\mathbb{P}_0}\right)_t$ belonging to $\mathbb{P}, \mathbb{Q} \in \mathcal{P}$ that the measure $\tilde{\mathbb{P}}$ defined by the density*

$$\left(\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}_0}\right)_t = \begin{cases} p_t & \textit{if } t \leq \tau \\ \frac{p_\tau q_t}{q_\tau} & \textit{else} \end{cases}$$

belongs to $\mathcal{P}$ as well.

As mentioned in the introduction this assumption also originates from a feature claimed for preferences. It was introduced by Epstein and Schneider in [ES]. They expanded the Multiple Priors Model (see [GS]) to a dynamic setting and asked the DM to be dynamically consistent in his decisions. With this they meant that if two acts are identical up to some time t but in $t+1$ the one is preferred over the other, then this should already be the case at time t. This implies that a DM will never regret his earlier decisions. In their paper Epstein and Schneider then showed that preferences fulfill this requirement if and only if the utility functional one obtains contains a rectangular set of measures. Rectangularity is equivalent to time-consistency. Time-consistency was introduced by Delbaen in [Del] where he also showed the equivalence to rectangularity. These two features stand for being able to judge each period in time with a different measure. More technically they allow to consistently paste together different densities at different times and still stay in the set. They also make it possible to use backward induction and allow for a Law of Iterated Expectations.

The set used to characterize $\mathcal{P}$ will be denoted by $\mathcal{A}$. It will be shown that it consists of predictable processes, is compact and that the process constant to zero is included in it. Furthermore we will see that it is stable under pasting. This means for a stopping time τ and processes $(\alpha_t)_t, (\beta_t)_t \in \mathcal{A}$ the process defined by

$$\gamma_t := \begin{cases} \alpha_t & \text{if } t \leq \tau \\ \beta_t & \text{else} \end{cases}$$

belongs to $\mathcal{A}$ as well. Later on we will show if we assume these properties for a set $\mathcal{A}$ then we can derive a set of measures $\mathcal{P}$ that features our original characteristics.

3 From $\mathcal{P}$ to $\mathcal{A}$

The goal of this section is to prove the main theorem of this paper, which tells us, that every time-consistent set of measures in our setting can also be described via a set of predictable processes fulfilling certain properties.

Expressed more formally this results in

Theorem 3.1. *For every set of measures $\mathcal{P}$ satisfying assumptions 2.1, 2.2 and 2.3 there is a set of predictable processes $\mathcal{A}$ such that*

$$\mathcal{P} = \left\{ \mathbb{P} \;\middle|\; \left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)_t = \tilde{\mathcal{E}}_t(\alpha) \ , \ \alpha \in \mathcal{A}^t, t \in \{0, ..., T\} \right\} \qquad \textit{where}$$

$$\tilde{\mathcal{E}}_t(\alpha) = \exp\left(\sum_{s=1}^{t} \sum_{h=1}^{\nu-1} \alpha_{hs} \Delta\omega_{hs} - \sum_{s=1}^{t} \ln \mathbb{E}\left[\exp\left(\sum_{h=1}^{\nu-1} \alpha_{hs} \Delta\omega_{hs} \right) \right] \right)$$

The $\mathcal{A}$ resulting from each $\mathcal{P}$ inhabits following features:

- $0 \in \mathcal{A}$
- *$\mathcal{A}$ is compact.*
- *$\mathcal{A}$ is stable under pasting.*

In order to prove this theorem we will derive a set of predictable processes $\mathcal{A}$ for every time-consistent set $\mathcal{P}$ and then show that it inhabits the requested features. One important step along this way will be a martingale representation theorem which we will explain more thoroughly in the next subsection. After that we will show the construction of the processes starting with an arbitrary time-consistent set of measures satisfying the above assumptions. Following this we will show that the constructed processes really are what we asked for.

3.1 Martingale Representation

This important tool which we will use in our proof tells us that in our setting we can find a set of martingales with which we can represent every other martingale in our setting with the help of predictable processes. A set of martingales which has this representation property is called a martingale basis. More formally we define

Definition 1. *A finite set of martingales* $\{m_{1t}\}, ..., \{m_{kt}\}$ *is called a basis iff for every martingale* $\{x_t\}$ *there are predictable processes* $\{\alpha_{1t}\}, ..., \{\alpha_{kt}\}$ *such that for every* $1 \leq t \leq T$

$$x_t = x_0 + \sum_{h=1}^{k} \sum_{s=1}^{T} \alpha_{hs} \Delta m_{hs} \qquad \text{where } \Delta m_{hs} = m_{hs} - m_{h,s-1}$$

If the martingales $\{m_{1t}\}, ..., \{m_{kt}\}$ *are pairwise orthogonal, i.e. for every* $1 \leq j \leq k$, $1 \leq h \leq m$, $j \neq h$ *and every* $0 \leq t \leq T$, $\langle m_j, m_h \rangle_t = 0$, *then the basis* $\{m_{1t}\}, ..., \{m_{kt}\}$ *is called orthogonal.*

For our purposes it would be good to know in which cases such a basis exists especially with unique α's. An answer for this is provided by the following proposition. A slightly different version of this can be found in [Dot] but since we are looking for a unique representation we need to resrict the setting to a constant splitting function of our information structure. The proof is similar to the one in [Dot].

Proposition. (Martingale Representation)
Given a discrete space $\Omega = \{\omega_1, ..., \omega_k\}$ *which is endowed with an information structure* $\{\mathbf{F}_t\}_{t=0,...,T}$ *with* $\mathbf{F}_0 = \Omega$ *and* $\mathbf{F}_T = \{\{\omega_1\}, ..., \{\omega_k\}\}$ *and a constant splitting function with value* ν. *Then there exists an orthogonal martingale basis* $m_{1t}, ..., m_{\nu-1,t}$ *for which the predictable processes* $\{\alpha^x_{1t}\}, ..., \{\alpha^x_{\nu-1,t}\}$ *in the representation of every* $\{x_t\}$ *are unique.*

Since under the assumption of "no arbitrage" discounted assets are martingales for a martingale measure $\mathbb{P}^*$ this means for a binomial tree setting that there is one asset M_t with which every other asset X_t can be replicated and therefore hedged. More general in an n-nomial tree we can replicate every asset with a set of $n-1$ many assets.

3.2 Exponential form of the densities

The next step we will take is to show that every measure $\mathbb{P} \in \mathcal{P}$ can be uniquely related to predictable processes $\left(\alpha_{1s}^{\mathbb{P}}\right)_s, ..., \left(\alpha_{\nu-1,s}^{\mathbb{P}}\right)_s$.

Remark that this is exactly one process less than our splitting value ν.

The equivalence of the measures in addition to $\mathbb{P}_0 \in \mathcal{P}$ (Ass.2.1) gives us the possibility to identify each $\mathbb{P} \in \mathcal{P}$ uniquely with its density with respect to $\mathbb{P}_0$.

If you define $\left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)_t := \mathbb{E}\left[\left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)\Big|_{\mathcal{F}_t}\right]$ for every $t \leq T$ and every $\mathbb{P} \in \mathcal{P}$ with the expectation taken under $\mathbb{P}_0$ you obtain density processes which are $\mathbb{P}_0$-martingales.

Using Jensen's inequality and Doob's decomposition theorem (cp. [Bau] and [Dot]) each of the above densities can be written in the following form where $(M_t)_t$ is also a $\mathbb{P}_0$-martingale and $(A_t)_t$ is a non-decreasing and predictable process with $A_0 = 0$

$$\left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)_t = \exp(M_t - A_t).$$

Now applying the martingale representation theorem to M_t we obtain an orthogonal martingale basis $(\omega_{1s})_s, ..., (\omega_{\nu-1,s})_s$. This implies that there are predictable processes $\left(\alpha_{1s}^{\mathbb{P}}\right)_s, ..., \left(\alpha_{\nu-1,s}^{\mathbb{P}}\right)_s$ such that our densities can now be written in the following manner where $\Delta\omega_{hs} = \omega_{hs} - \omega_{h,s-1}$

$$\left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)_t = \exp\left(\sum_{s=1}^{t}\sum_{h=1}^{\nu-1}\alpha_{hs}^{\mathbb{P}}\Delta\omega_{hs} - A_t\right).$$

Now we still have to determine the A_t's. Using the martingale property of the densities and the measurability of the A_t's we receive the following recursive relation

$$A_{t+1} - A_t = \ln\mathbb{E}\left[\exp\left(\sum_{h=1}^{\nu-1}\alpha_{h,t+1}^{\mathbb{P}}\Delta\omega_{h,t+1}\right)\ \Big|\ \mathcal{F}_t\right].$$

This results in

$$A_t = \sum_{s=1}^{t}(A_s - A_{s-1}) = \sum_{s=1}^{t}\ln\mathbb{E}\left[\exp\left(\sum_{h=1}^{\nu-1}\alpha_{hs}^{\mathbb{P}}\Delta\omega_{hs}\right)\ \Big|\ \mathcal{F}_{s-1}\right].$$

Additionally thanks to the assumptions on our information structure, we can show that our filtration is generated by our martingale basis and this in addition to the predictability of the α's allows us to drop the conditioning on $\mathcal{F}_{s-1}$.

So for our density $\left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)_t$ we now have following representation

$$\left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)_t = \exp\left(\sum_{s=1}^{t}\sum_{h=1}^{\nu-1}\alpha^{\mathbb{P}}_{hs}\Delta\omega_{hs} - \sum_{s=1}^{t}\ln\mathbb{E}\left[\exp\left(\sum_{h=1}^{\nu-1}\alpha^{\mathbb{P}}_{hs}\Delta\omega_{hs}\right)\right]\right). \quad (1)$$

This construction now allows us to not only identify a measure $\mathbb{P}$ with its density with respect to $\mathbb{P}_0$ and the associated density process but also with the predictable processes in the above representation $\left(\alpha^{\mathbb{P}}_{1s}\right)_s, ..., \left(\alpha^{\mathbb{P}}_{\nu-1,s}\right)_s$. Consequently it gives us a mapping from our density processes to sets of predictable processes.

For notational convenience and in resemblence to a stochastic exponential we will denote the right hand side of (1) as $\tilde{\mathcal{E}}_t(\alpha^{\mathbb{P}})$.

So now if we denote the set of processes generated via this construction and the densities up to time t by

$$\mathcal{A}^t := \left\{\left(\alpha^{\mathbb{P}}_{1,s}, ..., \alpha^{\mathbb{P}}_{\nu-1,s}\right)_{s\in\{0,...,t\}} \mid \mathbb{P}\in\mathcal{P}\right\} \qquad \text{and}$$

$$\mathcal{D}^t := \left\{\left(\left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)_1, ..., \left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)_t\right) \mid \mathbb{P}\in\mathcal{P}\right\}$$

we have constructed a mapping $\tilde{\mathcal{E}}_t^{-1} : \mathcal{D}^t \to \mathcal{A}^t$.

From this construction and from the assumption that $\mathbb{P}_0 \in \mathcal{P}$ we directly conclude that the α's are predictable and that $0 \in \mathcal{A} := \mathcal{A}^T$.

3.3 Compact-valuedness of the α's

One further thing we want to show is that the compactness of the densities resulting from $\mathcal{P}$ implies compactness of $\mathcal{A}$. The compactness on $\mathcal{A}^t$ is defined via the norm $||\alpha||_{t,L^1} := \max_{s\in\{0,...,t\}} ||\alpha_s||_{L^1}$.

This is a straight forward consequence of our assumptions and the preceding construction. In the construction of the α's every step was unique thanks to our assumptions. A density with respect to a designated measure uniquely characterizes a measure, the same is true for the construction of our density processes. Doob's decompostion is also unique and since we assumed a finite and constant splitting function the martingale representation also delivers unique predictable processes once the martingale basis is fixed. All in all the set of α's that belongs to one $\mathbb{P}$ is unique. Additionally a set of α's provides exactly one density and through that uniquely one measure. For this reason our $\tilde{\mathcal{E}}_t$ gives us a bijective mapping from the set of predictable processes $\mathcal{A}^t$ to our set of densities $\mathcal{D}^t$. This mapping is also continuous since the elements of our martingale basis are bounded thanks to the finite splitting index.

Since this also implies a continuous mapping between the densities and the predictable processes, the compactness on one side carries over to the other.

3.4 Stability under Pasting

The final property we claimed for our processes is stability under pasting. This property however follows directly from the assumption that $\mathcal{P}$ is time-consistent. To make this more clear define for $(\alpha_t^{\mathbb{P}})_t, (\alpha_t^{\mathbb{Q}})_t \in \mathcal{A}$ and a stopping time $\tau \leq T$

$$\beta_t := \begin{cases} \alpha_t^{\mathbb{P}} & \text{if } t \leq \tau \\ \alpha_t^{\mathbb{Q}} & \text{else.} \end{cases}$$

Our aim now is to show that this process lies in $\mathcal{A}$, i.e. that there exists a $\mathbb{P}^* \in \mathcal{P}$ such that $\left(\frac{d\mathbb{P}^*}{d\mathbb{P}_0}\right)_t = \tilde{\mathcal{E}}_t(\beta)$. If we plug β into equation (1) and define $\mathbb{P}^*$ by

$$\left(\frac{d\mathbb{P}^*}{d\mathbb{P}_0}\right)_t := \begin{cases} \tilde{\mathcal{E}}_t(\alpha^{\mathbb{P}}) & \text{if } t \leq \tau \\ \frac{\tilde{\mathcal{E}}_t(\alpha^{\mathbb{Q}})\tilde{\mathcal{E}}_\tau(\alpha^{\mathbb{P}})}{\tilde{\mathcal{E}}_\tau(\alpha^{\mathbb{Q}})} & \text{else} \end{cases}$$

we notice that $\beta \in \mathcal{A}$ is equivalent to $\mathbb{P}^* \in \mathcal{P}$. The fact that $\mathbb{P}^* \in \mathcal{P}$ however follows directly from our assumption of time-consistency.

If we now combine the above propositions we have shown our first theorem.

4 Necessity

Now let us look at the conversion of the theorem above. The goal of this section will be to show that every $\mathcal{A}$ with the above properties defines a time-consistent set of measures. So we see that the properties of $\mathcal{A}$ are not only sufficient but also necessary. For this purpose we will derive a set of measures $\mathcal{P}$ from a given set $\mathcal{A}$ of predictable processes which are assumed to be compact-valued and stable under pasting. Additionally we claim that

$\mathcal{A}$ contains the process constant to zero. Our goal will be to verify that the derived $\mathcal{P}$ satisfies the assumptions made in the model specifications.

Formally this will lead to following theorem

Theorem 4.1. *For every set of predictable processes $\mathcal{A}$ that satisfies the properties shown in Theorem 3.1 there exists a set of measures $\mathcal{P}$, such that*

$$\mathcal{A} = \left\{ \alpha \mid \left(\frac{d\mathbb{P}}{d\mathbb{P}_0} \right)_t = \tilde{\mathcal{E}}_t(\alpha) \quad , \quad \mathbb{P} \in \mathcal{P} \right\}$$

Every $\mathcal{P}$ constructed in this way has the following properties:

- $\mathbb{P}_0 \in \mathcal{P}$ *and* $\mathbb{P} \sim_{loc} \mathbb{P}_0$ *for all* $\mathbb{P} \in \mathcal{P}$
- $\mathcal{P}$ *is compact*
- $\mathcal{P}$ *is time-consistent.*

4.1 Construction of $\mathcal{P}$

If we use the same identification as in part 3.2 between the processes $(\alpha_t)_{t\in\{0,\dots,T\}}$ and the densities we are able to construct a density process $\left(\left(\frac{d\mathbb{P}^\alpha}{d\mathbb{P}_0}\right)_t\right)_t$ for every $\alpha \in \mathcal{A}$.

From the construction it follows immediately that the obtained processes are $\mathbb{P}_0$-martingales with expectation 1 and since the processes are clearly strictly larger than zero they are indeed density processes.

Let us define our new set of measures by

$$\mathcal{P} := \left\{ \mathbb{P} \;\middle|\; \frac{d\mathbb{P}}{d\mathbb{P}_0}\Big|_{\mathcal{F}_t} = \tilde{\mathcal{E}}_t(\alpha) \text{ for } \alpha \in \mathcal{A} \right\}.$$

Since the process $\alpha \equiv 0$ is assumed to be an element of $\mathcal{A}$ we get that $\mathbb{P}_0 \in \mathcal{P}$. From the fact that all $\mathbb{P} \in \mathcal{P}$ are constructed via denstiy processes with respect to $\mathbb{P}_0$ we can also directly conclude that our measures are all equivalent to our reference measure.

4.2 Time-Consistency

As when showing that we can derive $\mathcal{A}$ from $\mathcal{P}$ time-consistency in our set $\mathcal{P}$ is equivalent to stability under pasting in our set $\mathcal{A}$ and thus this property follows instantly from our assumptions.

4.3 Compactness of densities

Here again the fact that the $\tilde{\mathcal{E}}$ is a bijective and continuous mapping is the reason why the compactness of the α's implies compactness of the densities.

And again summarizing the above propositions leads us to the proof of the second theorem.

5 Examples

In this section we introduce some examples for which this result is applicable and might simplify calculations.

5.1 Binomial Tree

The most basic example one can think of in this setting is a binomial tree. It has a constant and finite splitting index of two. Here things are still very basic to calculate. One can for instance show that a convex set of priors results in a convex set of processes and vice versa which is in general not true for a higher splitting index. Put more formally we have

Proposition. *On a binomial tree every convex set of measures fulfilling assumptions 2.1, 2.2 and 2.3, i.e.* $\mathcal{P} = \{(p_1, ..., p_T) \ |p_t \in [\underline{p_t}, \overline{p_t}]$ *for all* $t = \{0, ..., T\}\}$*, is equivalent to the respective processes lying in a predictable interval* $[a_t, b_t]$*, where* $p_t = P[X_t = up \ |\mathcal{F}_{t-1}]$.

Proof. For the proof we will work ourselves through the tree succesively for every time period t.

Starting with $t = 1$ the density for a fixed $\mathbb{P}$ takes following form

$$\frac{d\mathbb{P}}{d\mathbb{P}_0}\Big|_{\mathcal{F}_1}(\text{up}) = 2p = \frac{2\exp(\alpha\Delta\omega_1(\text{up}))}{\exp(\alpha\Delta\omega_1(\text{up})) + \exp(\alpha\Delta\omega_1(\text{down}))}$$

this can be transformed to

$$\alpha = \ln\left(\frac{1-p}{p}\right)(\omega_1(\text{down}) - \omega_1(\text{up}))^{-1}$$

which is a function that is monotone and continuous in p. So if $p \in [\underline{p}, \overline{p}]$ then this results in boundaries a, b which are $\mathcal{F}_0$-measurable s.t. $\alpha \in [a, \overline{b}]$.

One can show the conversion by the same argumentation since the above formula can be converted to a function $p(\alpha)$ which is also monotone and continuous in α. Therefore a convex set of α's gives us a convex set of probabilities $[\underline{p_t}, \overline{p_t}]$ where $\underline{p_t} = \inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}[X_t = \text{up } |\mathcal{F}_{t-1}]$.

This can easily be extended to further time periods by just looking at the one step ahead measures or densities in an analogous way. □

5.2 Exponential Families

A further example for expressing time-consistent sets of measures via predictable processes was given by Riedel in [Rie2]. He introduced what he calls dynamic exponential families which is the discrete version of κ-ambiguity in the paper of Epstein and Chen see [EC] but with predictable bounds.

He starts with a probability state space $(S, \mathcal{S}, \nu_0)$ with $S \subset \mathbb{R}^d$. With this he constructs a probability space with $(\Omega, \mathcal{B}, (\mathcal{F}_t)_t, P_0)$, where

- $\Omega = S^{\mathbb{N}}$
- $\mathcal{B} = \bigotimes_{t=1}^{\infty} \mathcal{S}$ σ-field generated by all projections $\epsilon_t : \Omega \to S$
- $(\mathcal{F}_t)$ generated by the sequence (ϵ_t)
- $\mathbb{P}_0 = \bigotimes_{t=1}^{\infty} \nu_0$ probability s.t. ϵ_t iid with distribution ν_0

Now assuming that $\int_S e^{\lambda\cdot x}\nu_0(dx) < \infty$ s.t. the log-Laplace function $L(\lambda) = \log \int_S e^{\lambda\cdot x}\nu_0(dx)$ is well defined. With the help of predictable processes $(\alpha_t)_t$ he then defines densities on $(\Omega, \mathcal{B}, (\mathcal{F}_t)_t, \mathbb{P}_0)$ via

$$\mathcal{D}_t^{\alpha} := \exp\left(\sum_{s=1}^{t} \alpha_s \epsilon_s - \sum_{s=1}^{t} L(\alpha_s)\right).$$

Then for fixed predictable processes $a < b$ one gets a set of densities which defines a time-consistent set of measures by setting

$$\mathcal{P}^{a,b} = \left\{\mathbb{P} \;\Big|\; \left(\frac{d\mathbb{P}}{d\mathbb{P}_0}\right)_t = \mathcal{D}_t^{\alpha} \;,\; \alpha \in [a,b]\right\}.$$

5.3 Trinomial Tree

Switching between these two representations does not work too well in general. Starting with a two period trinomial tree which means we have a state space $\Omega = \{s_1, ..., s_9\}$ and the information structure $\mathbf{F}_0 = \Omega$, $\mathbf{F}_1 = \{\{s_1, s_2, s_3\}, \{s_4, s_5, s_6\}, \{s_7, s_8, s_9\}\}$ and $\mathbf{F}_2 = \{\{s_1\}, ..., \{s_9\}\}$ we define the rather simple time-consistent set

$$\mathcal{P} = \left\{ \left(\frac{1}{3} + \epsilon, \frac{1}{3} + \delta, \frac{1}{3} - \epsilon - \delta \right) \,\middle|\, \epsilon, \delta \in \left(-\frac{1}{3}, \frac{1}{3} \right) \text{ and } \epsilon + \delta \neq \frac{1}{3} \right\}.$$

We then construct a martingale basis in this tree with respect to the uniform distribution and then show what this set looks like expressed via predictable processes and our basis.

A martingale basis $\{\omega_t^1\}, \{\omega_t^2\}$ in this case is given by

ω^1	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9
$t = 0$	0	0	0	0	0	0	0	0	0
$t = 1$	1	1	1	0	0	0	-1	-1	-1
$t = 2$	1	-1	3	2	-4	2	1	-1	-3

and

ω^2	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9
$t = 0$	0	0	0	0	0	0	0	0	0
$t = 1$	1	1	1	-2	-2	-2	1	1	1
$t = 2$	-1	2	2	-1	-2	-3	2	-1	2

These tables contain the values the two martingales ω^1 and ω^2 realize at different times $t = 0, 1, 2$ and each possible state of the world $s_1, ...s_9$.

If we now calculate the processes that belong to each of the measures above we obtain
for $t = 1$ and $i = 1, ..., 9$

$$\alpha_1^1(s_i) = \frac{1}{2} \ln \frac{1 + 3\epsilon}{1 - 3\delta - 3\epsilon} \quad \text{and} \quad \alpha_1^2(s_i) = \frac{1}{3} \ln \frac{(1 + 3\epsilon)(1 - 3\epsilon - 3\delta)}{1 + 3\delta}$$

and for $t = 2$

$$\alpha_2^1(s_i) = \begin{cases} \frac{1}{4} \ln \frac{1-3\epsilon-3\delta}{1+3\delta} & \text{for } i = 1, 2, 3 \\ \frac{1}{6} \ln \frac{\sqrt{(1-3\epsilon-3\delta)(1+3\epsilon)}}{1+3\delta} & \text{for } i = 4, 5, 6 \\ \frac{1}{4} \ln \frac{1+3\epsilon}{1-3\epsilon-3\delta} & \text{for } i = 7, 8, 9 \end{cases}$$

$$\alpha_2^2(s_i) = \begin{cases} \frac{1}{3}\ln\frac{\sqrt{(1-3\epsilon-3\delta)(1+3\delta)}}{1+3\epsilon} & \text{for } i=1,2,3 \\ \frac{1}{2}\ln\frac{1+3\epsilon}{1-3\epsilon-3\delta}1+3\delta & \text{for } i=4,5,6 \\ \frac{1}{3}\ln\frac{\sqrt{(1+3\epsilon)(1-3\epsilon-3\delta)}}{1+3\delta} & \text{for } i=7,8,9 \end{cases}$$

As one can see a comparably simple set in the one representation can become relatively complicated in the other.

5.4 DTV@R

Another important area in which time-consistent sets of measures have been studied are risk measures. Artzner et al. in [ADEH] showed that every coherent risk measure ρ_t has a robust representation involving a set of measures $\mathcal{P}$.

$$\rho_t(X) = \operatorname*{essinf}_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}\left[X \mid \mathcal{F}_t\right]$$

Then in [Rie] Riedel showed that the dynamic risk measure $\rho = (\rho_t)_{t=0,\dots T}$ is dynamically consistent iff the set $\mathcal{P}$ is time-consistent. Roorda and Schumacher in [RS] introduce dynamically consistent tail value at risk to a risk level λ(DTV@R$_\lambda$) as one of these time-consistent risk measures.

As the set $\mathcal{P}$ they take all measures $\mathbb{P}$ for which the one step ahead densities w.r.t. the reference measure $\mathbb{P}^*$ are bounded by $\frac{1}{\lambda}$ where λ is the usual risk level. If we want to describe this in our characterization it gives us

$$\frac{\tilde{\mathcal{E}}_t(\alpha)}{\tilde{\mathcal{E}}_{t-1}(\alpha)} = \exp\left(\alpha_t \cdot \Delta\omega_t - \ln\mathbb{E}\left[\exp\left(\alpha_t \cdot \Delta\omega_t\right)\right]\right) \leq \frac{1}{\lambda}$$

for all $t = 1, \dots T$ and all $\alpha \in \mathcal{A}$.

6 Possible Extensions

In this section we will discuss extensions.

6.1 Convexity

Since time-consistent sets are often used in optimization problems convexity of the sets is often assumed. It would be nice if this feature would carry over to the processes. Unforunately this is not the case in general, as can be seen in the following counterexample.

Take for example a trinomial tree with states s_1, s_2 and s_3 and just one time period. As a reference measure we will fix

$$\mathbb{P}_0(s_1) = \frac{1}{2}\ ,\ \mathbb{P}_0(s_2) = \frac{1}{4} \quad \text{and} \quad \mathbb{P}_0(s_3) = \frac{1}{4}.$$

A second measure will be given by

$$\mathbb{Q}(s_1) = \frac{1}{2}\ ,\ \mathbb{Q}(s_2) = \frac{1}{8} \quad \text{and} \quad \mathbb{Q}(s_3) = \frac{3}{8}.$$

The density of $\mathbb{Q}$ with respect to $\mathbb{P}_0$ will then be

$$\frac{d\mathbb{Q}}{d\mathbb{P}_0}(s_1) = 1\ ,\ \frac{d\mathbb{Q}}{d\mathbb{P}_0}(s_2) = \frac{1}{2} \quad \text{and} \quad \frac{d\mathbb{Q}}{d\mathbb{P}_0}(s_3) = \frac{3}{2}.$$

Since we want to show that from a convex set of measures a non-convex set of processes can arise, let us define our set of measures via

$$\mathcal{P} := \text{convH}\left\{\mathbb{P}, \mathbb{Q}\right\}.$$

Then let us look at the set of processes $\mathcal{A}$ arising from this convex set, especially $\alpha^{\mathbb{P}_0}$ and $\alpha^{\mathbb{Q}}$. Now if $\mathcal{A}$ were a convex set, then every convex combination of $\alpha^{\mathbb{P}_0}$ and $\alpha^{\mathbb{Q}}$ has to be an element of $\mathcal{A}$. Since $\alpha^{\mathbb{P}_0}$ is zero, because we chose $\mathbb{P}_0$ as our reference measure we look at $\frac{1}{2}\alpha^{\mathbb{Q}}$. If we now calculate the associated density to this process, we see that it can never originate from a convex combination of our original measures and therefore $\frac{1}{2}\alpha^{\mathbb{Q}} \notin \mathcal{A}$ and hence $\mathcal{A}$ is not convex.

6.2 Infinite Horizon

When extending our statements to an infinite time horizon let us first remark that our model assumptions can all be transferred without complications. The construction of the processes can also be maintained, since they are always constructed for a fixed time horizon up to a time t. That is also the reason why the mapping from our densities to our processes still inhabits the same features, i.e. it is continuous and bijective. Therefore in this case the compactness also carries over from one side to the other. It is also clear that stability under pasting is equivalent to time-consistency for an infinite horizon as well. So altogether our statements can smoothly be converted from a finite to an infinite time horizon.

6.3 Looser Assumptions on Splitting Function

Since our assumptions on the filtration are very restrictive, it would be nice if they could be relaxed in one way or another.

One way would be to give up the assumption of a constant splitting function. In this case however you run into the problem that the α's that arise from the martingale representation are no longer unique and with that the mapping no longer distinct.

A second way is allowing for the splitting value to become infinite. This however has the consequence that the martingale representation will not necessarily exist anymore.

7 Conclusions

For our special setting, i.e. discrete and with special assumptions on the information structure, we have constructed an alternative characterization for time-consistent sets of measures. We have shown that all sets of time-consistent sets of measures can be expressed by predictable processes and vice versa.

As can be seen in the extensions standard generalizations fail to work. So as far as I am concerned this is the most general this characterization can be formulated in this setting.

For practical applications we have shown that for problems which can be modeled in the form of decision trees (with a constant number of branches e.g. trinomial trees) we now know what a time-consistent set of measures must look like expressed via predictable processes which might simplify calculations. So hopefully our construction will be helpful in the future e.g. for solving Optimal Stopping Problems which can be modeled in this framework.

References

[AA] ANSCOMBE, F.J. and R.J. AUMANN (1963): " *A Definition of Subjective Probability* ", Annals of Mathematical Statistics, Vol. 34, pp. 199-205

[Arr] ARROW, K. (1971): Essays in the Theory of Risk Bearing. Markham Publishing Co., Chicago.

[ADEH] ARTZNER, P., F. DELBAEN, J.-M. EBER, and D. HEATH (1999): *Coherent Measures of Risk*, Mathematical Finance Vol. 9, pp. 203228.

[Bau] BAUER, H. (1974): *"Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie"*, 2. Auflage, de Gruyter, Berlin, New York.

[CMMT] CHATEAUNEUF, A., F. MACCHERONI, M. MARINACCI, and J.-M. TALLON (2005): *" Monotone Continuous Multiple Priors "*, Economic Theory Vol. 26 (4), pp. 973-982

[Del] DELBAEN, F. (2003): *"The Structure of m-Stable Sets and in Particular of the Set of Risk Neutral Measures"*, In Yor, M., Emery, M. (eds.): In Memoriam Paul-André Meyer - Séminaire de Probabilités XXXIX. Berlin Heidelberg New York:Springer, pp. 215-258.

[Dot] DOTHAN, M. (1990): *"Prices in Financial Markets"*, 1. edition, Oxford University Press.

[Ell] ELLSBERG, D. (1961): *" Risk, Ambiguity, and the Savage Axioms "*, Quarterly Journal of Economics, Vol. 75, pp. 643-669

[EC] EPSTEIN, L. and Z. CHEN (2002): *" Ambiguity, Risk and Asset Returns in Continuous Time"*, Econometrica, Vol. 70, pp. 1403-1443.

[EM] EPSTEIN, L. and M. MARINACCI (2006): *" Mutual Absolute Continuity of Multiple Priors "* , Working Paper

[ES] EPSTEIN, L. and M. SCHNEIDER, (2003) : *"Recursive Multiple Priors"*, Journal of Economic Theory, Vol. 113, pp.1-31

[FS] FÖLLMER, H. and A. SCHIED (2002): *" Stochastic Finance "*, Studies in Mathematics. de Gruyter, Berlin, New York

[GS] GILBOA, I. and D. SCHMEIDLER, (1989): *"Maxmin Expected Utility with Non-unique Prior"*, Journal of Mathematical Economics, Vol. 18, pp. 141-153.

[Kr] KREPS, D. (1979): *"A Representation Theorem for 'Preference for Flexibility' "*, Econometrica, Vol. 47, pp. 565-577

[vNM] VON NEUMANN, J. and O. MORGENSTERN (1944): *" Theory of Games and Economic Behavior "*, Second Edition, Princeton University Press, Princeton, NJ, 1947

[Rie] RIEDEL, F. (2004): *" Dynamic Coherent Risk Measures"*, Stochastic Processes and their Applications, Vol. 112, pp. 185-200.

[Rie2] RIEDEL, F. (2009): *"Optimal Stopping with Multiple Priors"* , Econometrica. forthcoming

[RS] ROORDA, B. and H. SCHUMACHER (2007): *"Time Consistency Conditions for Acceptability Measures, with an Application to Tail Value at Risk"*, Insurance: Mathematics and Economics, Vol. 40(2), pp. 209-230.

[Sav] SAVAGE, L.J. (1954): *" The Foundation of Statistics "*, John Wiley and Sons, New York. Revised and enlarged edition, Dover, New York, 1972