Abstract Time Consistency
and
Decomposition

Henri Gerard
J.-P. Chancelier, M. De Lara, P. Carpentier

ENPC - CERMICS
Labex Bézout - LIGM

June 03, 2016
Time consistency in a nutshell

Given two processes \((X_0, \cdots, X_T)\) and \((Y_0, \cdots, Y_T)\),

- we look for numerical evaluations (risk measures) of the tails \((X_t, \cdots, X_T)\) of the process
- that satisfy time consistency, in the same way that the mathetical expectation does in

\[
\mathbb{E}_P[X_0 + \cdots + X_T] = \mathbb{E}_P[X_0 + \cdots + X_t + \mathbb{E}_P[X_{t+1} + \cdots + X_T | \mathcal{F}_t]]
\]

Such a property is essential to establish a dynamic programming equation in dynamic optimization
Example of time consistency

Example

- $(\mathcal{F}_t)_{t \in [0;T]}$ filtration of $(\Omega, \mathcal{F}, P)$
- $X_t = \mathcal{L}^p(\Omega, \mathcal{F}_t, P)$
- $A = X_0 \times \ldots \times X_{t_0}$
- $B = X_{t_0+1} \times \ldots \times X_T$
We focus on the risk averse case

- There is a strong literature on the subject:

- We want to study time consistency for
  - a general criterion (not necessarily time-additive and with dynamic risk measures)
  \[
  F_0[X_0, \cdots, X_T] = F_0[X_0, \cdots, X_t, F_{t+1}[X_{t+1}, \cdots, X_T | F_t]]
  \]
  - with general structure of information (not necessarily filtration, to account for decentralized information among agents)
1. Abstract notion of time consistency and functional representation

2. Revisiting classical examples of the literature
   - Ruszczyński (2010)

3. Perspectives for optimization under risk and conclusion
Outline of the section

1. Abstract notion of time consistency and functional representation
We introduce the following notations

- \( A \) and \( B \) are two sets
- \( \preccurlyeq_B \) is a preorder on \( B \)
- \( \preccurlyeq_{A \times B} \) is a preorder on \( A \times B \)

**Definition (Time consistency)**

\[ b \preccurlyeq_B b' \Rightarrow (a, b) \preccurlyeq_{A \times B} (a, b') \]
Variations around time consistency

Definition (Strong time consistency)

\[ b \preceq_B b' \Rightarrow (a, b) \preceq_{A \times B} (a', b') , \ \forall (a, a') \in A^2 , \ \forall (b, b') \in B^2 \]

Definition (Time consistency)

\[ b \preceq_B b' \Rightarrow (a, b) \preceq_{A \times B} (a, b') , \ \forall a \in A , \ \forall (b, b') \in B^2 \]

Definition (Weak time consistency)

\[ b \sim_A b' \Rightarrow (a, b) \sim_{A \times B} (a, b') , \ \forall a \in A , \ \forall (b, b') \in B^2 \]
We will focus on weak time consistency

**Proposition**

*We have the following implications*

\[\text{Strong time consistency} \Downarrow \text{Time consistency} \Downarrow \text{Weak time consistency}\]
Functional representation

Definition (Representation of a preorder)
Let $\preceq_B$ be a preorder on the set $B$
Let $f : B \to Y$ be a mapping, where $Y$ is equipped with a preorder $\preceq_Y$
We say that $(f, \preceq_Y)$ is a representation of $(B, \preceq_B)$ if

$$b \preceq_B b' \iff f(b) \preceq_Y f(b')$$

Remark
$(f, \leq_R)$ is called a numerical representation

Remark
Let $f : B \to Y$ be a mapping, where $Y$ is equipped with a preorder $\preceq_Y$
Then $f(b) \preceq_Y f(b')$ induces a preorder $\preceq_B$ on $B$
Abstract notion of time consistency and functional representation

Functional definition of time consistency

Definition (Time consistency for mappings (or mapping induced orders))
Consider two mappings $g : A \times B \rightarrow X$ and $f : B \rightarrow Y$, where the sets $X$ and $Y$ are respectively equipped with the preorders $\preceq_X$ and $\preceq_Y$. The quadruplet $(g, \preceq_X, f, \preceq_Y)$ is said to satisfy time consistency when

$$f(b) \preceq_Y f(b') \Rightarrow g(a, b) \preceq_X g(a, b')$$

Remark
We say that $f$ is a factor and that $g$ is an aggregator

Example
$$f((b_{t_0+1}, \ldots, b_T)) = \mathbb{E}_{\mathbb{P}}[b_{t_0+1} + \cdots + b_T \mid \mathcal{F}_{t_0}]$$
$$g((a_0, \ldots, a_{t_0}, b_{t_0+1}, \ldots, b_T)) = \mathbb{E}_{\mathbb{P}}[a_0 + \cdots + a_{t_0} + b_{t_0+1} + \cdots + b_T]$$
We introduce a set-valued mapping

Given an aggregator $g : A \times B \to X$ and a factor $f : B \to Y$, we introduce a set-valued mapping, called subaggregator

**Definition**

We denote by $\phi^{f,g} : A \times Y \Rightarrow X$ the set-valued mapping

$$\phi^{f,g}(a, y) = \{g(a, b) : b \in f^{-1}(y)\}$$

If $y \notin \text{Im}(f)$ then $\phi^{f,g}(a, y) = \emptyset$

**Example**

$$f(b) = \mathbb{E}_P[b_{t_0+1} + \cdots + b_T | \mathcal{F}_{t_0}]$$
Abstract notion of time consistency and functional representation

Nested decomposition of time consistent mappings

Theorem (Weak nested decomposition)

*The aggregator $g$ and factor $f$ are weakly time consistent if and only if the set-valued function $\phi^{f,g}$ is a mapping*

Remark

We then have a nested formula $g(a, b) = \phi^{f,g}(a, f(b))$

Example

\[
\mathbb{E}_P \left[ a_0 + \cdots + a_{t_0} + b_{t_0+1} + \cdots + b_T \right] = \mathbb{E}_P \left[ a_0 + \cdots + a_{t_0} + \mathbb{E}_P \left[ b_{t_0+1} + \cdots + b_T \mid \mathcal{F}_{t_0} \right] \right]
\]

\[
\phi^{f,g}(a, y) = \mathbb{E}_P \left[ a_0 + \cdots + a_{t_0} + y \right], \quad y \in \mathcal{X}_{t_0}
\]

\[
g(a, b) = \phi^{f,g}(a, f(b))
\]
## Functional characterization of three notions of time consistency

<table>
<thead>
<tr>
<th>Weak</th>
<th>Usual</th>
<th>Strong</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b \sim_B b'$</td>
<td>$b \preceq_B b'$</td>
<td>$a \preceq_A a', b \preceq_B b'$</td>
</tr>
<tr>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
<td>$\downarrow$</td>
</tr>
<tr>
<td>$(a, b) \sim_{A \times B} (a, b')$</td>
<td>$(a, b) \preceq_{A \times B} (a, b')$</td>
<td>$(a, b) \preceq_{A \times B} (a, b')$</td>
</tr>
<tr>
<td>$\phi^{f,g}$ is a mapping</td>
<td>$\phi^{f,g}$ is a mapping</td>
<td>$\phi^{f,g}$ is a mapping</td>
</tr>
<tr>
<td>increasing in its second argument</td>
<td>increasing</td>
<td>increasing in both arguments</td>
</tr>
</tbody>
</table>
We have developed an abstract framework to deal with time consistency.
We have illustrated the notions on a (simple) running example.
We hope now to cover other cases with our abstract framework.
We now show how to apply this framework to examples of the literature.
2 Revisiting classical examples of the literature
   - Ruszczyński (2010)
Revisiting classical examples of the literature

- Ruszczyński (2010)
Let \( A = X_0 \times \cdots \times X_{t_0} \) and \( B = X_{t_0+1} \times \cdots \times X_T \)

\[
g : A \times B \rightarrow \mathbb{R} \\
(a, b) \rightarrow \sup_{Q \in \mathcal{Q}} E_Q \left[ a_0 + \cdots + a_{t_0} + b_{t_0+1} + \cdots + b_T \right]
\]

\[
f : B \rightarrow L^p(\Omega, \mathcal{F}_{t_0}, \mathbb{P}) \\
b \rightarrow \sup_{Q \in \mathcal{Q}} E_Q \left[ b_{t_0+1} + \cdots + b_T \mid \mathcal{F}_{t_0} \right]
\]

where \( \mathcal{Q} \) is a (closed convex) set of probability distributions on \( \Omega = \mathbb{R}^{t_0+1} \times \mathbb{R}^{T-t_0} \)
A probability distribution $\mathbb{Q}$ on the product space $\Omega = \mathbb{R}^{t_0+1} \times \mathbb{R}^{T-t_0}$ can be naturally decomposed into

- a marginal distribution $m_\mathbb{Q}$
- a stochastic kernel $k_\mathbb{Q}$ conditional to the $\sigma$–field $\mathcal{F}_{t_0}$
Rectangularity

Definition (Epstein and Schneider (2003))

We say that a set $\mathcal{Q}$ of probability distributions is rectangular if the image of $\mathcal{Q}$ by the mapping $\mathcal{Q} \mapsto (m_\mathcal{Q}, k_\mathcal{Q})$ is a rectangle. By an abuse of notation, we will write

$$\mathcal{Q} = \mathcal{M} \times \mathcal{K}$$

where $\mathcal{M}$ is a set of marginal distributions and $\mathcal{K}$ is a set of stochastic kernels.

Theorem

If $\mathcal{Q}$ is a rectangular set of probability distributions, then $f$ and $g$ are time consistent and we have

$$g(a, b) = \phi^f \cdot g(a, f(b)),$$

$$\phi^f \cdot g(a, y) = \sup_{\mathcal{Q} \in \mathcal{Q}} \mathbb{E}_\mathcal{Q}[a_0 + \cdots + a_{t_0} + y]$$
Sketch of proof

- First, we use a tower formula
  \[
  \sup_{Q \in \Omega} \mathbb{E}_Q \left[ a_0 + \cdots + a_{t_0} + \mathbb{E}_Q \left[ b_{t_0} + \cdots + b_T \mid \mathcal{F}_{t_0} \right] \right]
  \]

- Second, we use the property that \( \Omega \) is rectangular and that \( a_0 + \cdots + a_{t_0} \) is \( \mathcal{F}_{t_0} \) measurable
  \[
  \sup_{(m,k) \in \mathcal{M} \times \mathcal{K}} \mathbb{E}_m \left[ a_0 + \cdots + a_{t_0} + \mathbb{E}_k \left[ b_{t_0} + \cdots + b_T \mid \mathcal{F}_{t_0} \right] \right]
  \]

- Third, we take the supremum over the complete sup semilattice of \( \mathcal{F}_{t_0} \)-measurable random variables
  \[
  \sup_{m \in \mathcal{M}} \mathbb{E}_m \left[ a_0 + \cdots + a_{t_0} + \sup_{k \in \mathcal{K}} \mathbb{E}_k \left[ b_{t_0} + \cdots + b_T \mid \mathcal{F}_{t_0} \right] \right]
  \]
Revisiting classical examples of the literature

- Ruszczyński (2010)
We consider

- the spaces $X_t = L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $X_{t,T} = X_t \times X_{t+1} \times \ldots \times X_T$
- a sequence of conditional risk measures $\rho_{t,T} = X_{t,T} \to X_t$ (with the monotonicity property) called dynamic risk measure

**Definition**

A dynamic risk measure $\{\rho_{t,T}\}_{t=1}^T$ is called time consistent if, for all $1 \leq \tau < \theta \leq T$, and all sequences $X, Y \in X_{\tau,T}$,

$$X_k = Y_k, \ k \in [\tau; \theta - 1]$$

$$\rho_{\theta,T}(X_\theta, \ldots, X_T) \leq \rho_{\theta,T}(Y_\theta, \ldots, Y_T)$$

$$\Downarrow$$

$$\rho_{\tau,T}(X_\tau, \ldots, X_T) \leq \rho_{\tau,T}(Y_\tau, \ldots, Y_T)$$
Theorem (Ruszczyński (2010))

Suppose a dynamic risk measure \( \{ \rho_{t,T} \} \) satisfies, for all \((X_t, ..., X_T)\), the conditions

\[
\rho_{t,T}(X_t, ..., X_T) = X_t + \rho_{t,T}(0, X_{t+1}, ..., X_T)
\]

\[
\rho_{t,T}(0, ..., 0) = 0
\]

Then \( \{ \rho_{t,T} \} \) is time consistent if and only if, for all \(1 \leq \tau \leq \theta \leq T\) and all \((X_\tau, ..., X_T)\), the following identity is true

\[
\rho_{t,T}(X_\tau, ..., X_{\theta-1}, X_{\theta}, ..., X_T) = \rho_{\tau,\theta}(X_\tau, ..., X_{\theta-1}, \rho_{\theta,T}(X_{\theta}, ..., X_T))
\]
Links between our framework and Ruszczyński’s one (1)

Result

The dynamic risk measure \( \{\rho_t, T\}_{t=1}^T \) is weakly time consistent if and only if \( \phi^{\rho_\theta, T, \rho_\tau, T}(X_\tau, \cdots, X_{\theta-1}, \cdot) \) is a mapping, for all \( 1 \leq \tau \leq \theta \leq T \) and for all \( (X_\tau, \cdots, X_T) \)

Then we have

\[
\rho_\tau, T(X_\tau, \cdots, X_{\theta-1}, X_\theta, \cdots, X_T) = \underbrace{\phi^{\rho_\theta, T, \rho_\tau, T}(X_\tau, \cdots, X_{\theta-1}, \rho_\theta, T(X_\theta, \cdots, X_T))}_{\text{subaggregator}}
\]

We put \( A = X_\tau \times \cdots \times X_\theta \) and \( B = X_{\theta+1} \times \cdots \times X_T \), then use our abstract result
Revisiting classical examples of the literature

Ruszczyński (2010)

Links between our framework and Ruszczyński’s one (2)

Suppose that, in addition, the dynamic risk measure \( \{ \rho_t, T \}_{t=1}^T \) satisfies, for all \((X_\tau, \cdots, X_T)\), \( t \in [1, T] \), the conditions

\[
\rho_t, T(X_t, \cdots, X_T) = X_t + \rho_t, T(0, X_{t+1}, \cdots, X_T)
\]

\[
\rho_t, T(0, \cdots, 0) = 0
\]

Then

\[
\phi^{\rho_\theta, T, \rho_\tau, T} = \rho_\tau, \theta
\]

that is

\[
\rho_\tau, T(X_\tau, \cdots, X_{\theta-1}, X_\theta, \cdots, X_T) = \rho_\tau, \theta(X_\tau, \cdots, X_{\theta-1}, \rho_\theta, T(X_\theta, \cdots, X_T))
\]
<table>
<thead>
<tr>
<th>Ruszczyński</th>
<th>Us</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Usual time consistency</strong></td>
<td><strong>Weak time consistency</strong></td>
</tr>
<tr>
<td>$b \preceq_B b' \Rightarrow (a, b) \preceq_{A \times B} (a, b')$</td>
<td>$b \sim_B b' \Rightarrow (a, b) \sim_{A \times B} (a, b')$</td>
</tr>
<tr>
<td><strong>Monotonicity property</strong></td>
<td><strong>$\emptyset$</strong></td>
</tr>
<tr>
<td><strong>Additive criterion</strong></td>
<td><strong>Any criterion</strong></td>
</tr>
<tr>
<td><strong>Explicit subaggregator</strong></td>
<td><strong>Existence</strong></td>
</tr>
<tr>
<td>$\rho_{T, \theta}$</td>
<td>$\phi^{\rho_{\theta, T}, \rho_{T, T}}$</td>
</tr>
</tbody>
</table>
Conclusion of the section and perspectives

So far, we have

- revisited different examples with one framework

We want to

- extend the results established for time additive cases to other time aggregators (multiplicative, maximum...)
- study more general Fenchel transforms than $\sup_{Q \in \mathcal{Q}} \mathbb{E}_Q [X]$
3 Perspectives for optimization under risk and conclusion
Towards dynamic programming

- We want to mix optimization with our framework to obtain dynamic programming equations of the form

\[
\inf_{a \in A, b \in B} g(a, b) = \inf_{a \in A} \phi^f \circ g(a, \inf_{b \in B} f(b))
\]

- For this purpose, we establish results useful for optimization
Inheritance of properties

We assume that factor $f$ and aggregator $g$ are weakly time consistent.

**Theorem (Monotonicity)**

*If the aggregator $g$ is monotonous in its second argument, then the subaggregator $\phi^{f,g}$ is monotonous in its second argument.*

**Theorem (Continuity)**

*If the aggregator $g$ is continuous with a compact image, if the factor $f$ is continuous with compact domain and image, then the subaggregator $\phi^{f,g}$ is continuous.*

**Theorem (Convexity)**

*If there exists $\bar{B} \subset B$ such that $f(\bar{B}) = Y$ and such that $f|_{\bar{B}}$ is linear, and if the aggregator $g$ is convex, then the subaggregator $\phi^{f,g}$ is convex.*
Conclusion and ongoing work

Conclusion

- We have developed a general abstract framework for time consistency and have applied it to classic examples of the literature.
- We have established inheritance properties that are useful for optimization.

Ongoing work

- How can we identify factors $f$ that yield to time consistency, for a given aggregator $g$?
- Switching from time consistency to non nested consistency (using multi-agent framework “à la Witsenhausen”)


