ISSN 2279-9362

Collegio Carlo Alberto

On time consistency for mean-variance portfolio selection

Elena Vigna

No. 476
December 2016

Carlo Alberto Notebooks

www.carloalberto.org/research/working-papers

On time consistency for mean-variance portfolio selection

Elena Vigna

December 22, 2016

Abstract

This paper adds to the debate on what is the "appropriate" approach to time inconsistency by making two contributions. As a first contribution, we define a suitable intertemporal preferences-driven reward and use it to compare the three possible approaches to time inconsistency for the mean-variance portfolio selection problem over $[t_0, T]$: precommitment approach (Zhou & Li (2000)), game theoretical approach (Basak & Chabakauri (2010), Björk & Murgoci (2010)), and dynamic approach (Pedersen & Peskir (2016)). We find that the precommitment strategy beats the other strategies if the investor only cares at the view point at time t_0 and is not concerned to be time-inconsistent in (t_0, T); the Nash-equilibrium strategy dominates the dynamic strategy until a time point $t^* \in (t_0, T)$ and is dominated by the dynamic strategy from t^* onwards. As a second contribution, we further elaborate on this matter, formulate the notion of *preferences consistency* and shed light on the price to be paid in terms of optimality and preferences consistency with each of the three approaches currently available for time inconsistency.

Keywords. Time consistency, dynamic programming, Bellman's optimality principle, time inconsistency, precommitment approach, Nash perfect equilibrium, mean-variance portfolio selection.

JEL classification: C61, D81, G11.

1 Introduction

The notion of time inconsistency for optimization problems dates back to Strotz (1956). Broadly speaking, time inconsistency arises in an intertemporal optimization problem when the optimal strategy selected at some time t is no longer optimal at time $s > t$. In other words, a strategy is time-inconsistent when the individual at future time $s > t$ is tempted to deviate from the strategy decided at time t. Clearly, an optimization problem gives raise to time-inconsistent strategies when the Bellman's principle does not hold and dynamic programming cannot be applied. In finance, a notable example of problem which is time-inconsistent is the mean-variance problem, where the time inconsistency is due to the fact that there is a non-linear function of the expectation of final wealth in the optimization criterium (due to the presence of the variance of final wealth). Another important problem which produces a time-inconsistent behaviour is the investment-consumption problem with non-exponential discounting. This was the case studied by Strotz (1956), and time inconsistency arises because the initial point in time enters in an essential way the objective criterium. For a clarifying formalization of the possible sources of time inconsistency in intertemporal optimization problems, see Björk & Murgoci (2010).

In the last two decades there has been a renewed interest in time inconsistency for financial and economic problems. According to Strotz (1956) there are two possible ways to deal with time-inconsistent problems: (i) precommitment approach; (ii) consistent planning approach. In the precommitment approach, the controller fixes an initial point (t_0, x_0) and finds the optimal control law $\hat{u}$ that maximizes the objective functional at time t_0 with wealth x_0, $J(t_0, x_0, u)$, disregarding the fact that at future time $t > t_0$ the control law $\hat{u}$ will *not* be the maximizer of the objective functional at time t with wealth x_t, $J(t, x_t, u)$; therefore, he precommits to follow the initial strategy $\hat{u}$, despite the fact that at future dates he will no longer be optimal according to his preferences. In the consistent planning approach, one tries to avoid time inconsistency by selecting the "best plan among those that he will actually follow". This approach translates into the search of a Nash subgame perfect equilibrium point. Intuitively, sitting at time t the future time interval $[t, T]$ can be seen as a continuum of players, each player $s \geq t$ being the "reincarnation" at time s of the player who sits at time t. With this approach, a time-consistent equilibrium policy is the collection of all optimal decisions $\hat{u}(s, \cdot)$ taken by any player $s \in [t, T]$, such that if player t knows that all players coming after him (in (t, T)) will use the control $\hat{u}$, then it is optimal to him, too, to play control $\hat{u}$.

The literature is full of examples of applications of the two approaches outlined. For conciseness reasons, we here report only a few of them. For instance, the mean-variance portfolio selection problem has been solved with the precommitment approach by Richardson (1989), Bajeux-Besnainou & Portait (1998), Zhou & Li (2000) and Li & Ng (2000), the first two with the martingale method, the last two with an embedding technique that transforms the mean-variance problem into a standard linear-quadratic control problem. The game theoretical solution to the mean-variance problem has been found originally by Basak & Chabakauri (2010), then extended to a more general class of time-inconsistent problems by Björk & Murgoci (2010). Other papers on the consistent planning approach for the mean-variance problem are Björk, Murgoci & Zhou (2014), Czichowsky (2013). The problem of non-exponential discounting, firstly introduced by Strotz (1956), has been treated with the game theoretical approach by Ekeland & Pirvu (2008), Ekeland, Mbodji & Pirvu (2012).

The precommitment strategy and the game theoretical approach are not the only ways to attack a problem that gives raise to time inconsistency. An alternative approach has been introduced recently by Pedersen & Peskir (2016) for the mean-variance portfolio selection problem, namely, the dynamically optimal strategy. The dynamic solution to the mean-variance problem in continuous-time introduced by Pedersen & Peskir (2016) is a novel approach to time inconsistency, although related work can be found in a recent paper by Karnam, Ma & Zhang (2016). The strategy proposed by Pedersen & Peskir (2016) is time-consistent in the sense that it does not depend on initial time and initial state variable, but differs from the subgame perfect equilibrium strategy. Moreover, their dynamic approach is intuitive and formalizes a quite natural approach to time inconsistency: it represents the behaviour of an optimizer who continuously revaluates his position and solves infinitely many problems in an instantaneously optimal way. The dynamically optimal individual is similar to the continuous version of the naive individual described by Pollak (1968). However, while the naive individual of Pollak (1968) at each revaluation time assumes that he will precommit his future behaviour – despite the evidence that he keeps deviating – the dynamically optimal individual does know that he will continuously deviate, so he does not fall into any contradiction. The dynamically optimal individual turns out to be at each time t the "reincarnation" of the precommitted investor and plays the strategy that the time-t precommitted investor would play at time t, deviating from it immediately after, by wearing the clothes of the time t^+ precommitted investor.

This paper adds to the debate on what is the "appropriate" approach to time inconsistency by making two contributions. As a first contribution, we define a suitable intertemporal preferences-driven reward and use it to compare the three outlined approaches to time inconsistency for the mean-variance portfolio selection problem: precommitment, consistent planning and dynamic optimality. A preview of the results is the following. Expectedly, the precommitment strategy beats the other strategies if the investor only cares at the view point at time t_0 and is not concerned to be time-inconsistent in (t_0, T); for the comparison between the two time-consistent strategies, we find that the Nash-equilibrium strategy dominates the dynamically optimal strategy until a time point $t^* \in (t_0, T)$ and is dominated by the dynamically optimal strategy from t^* onwards. We prove existence and uniqueness of the break even point t^* and provide a closed form for it. Interestingly, the break even point t^* does not depend on wealth, while it increases with the market price of risk and the time horizon T. These results are in line with the results in Pedersen & Peskir (2016), who also address the comparison among the two time-consistent strategies. Differently from them, we make the comparison at any time $t \in [t_0, T]$ and not only at initial time t_0 and final time T.

The main message that one can get from this analysis is that when there is a problem that gives rise to time inconsistency there is no clear-cut answer to the issue "what is the optimal plan". When the optimization criterium fails to satisfy conditions under which the Bellman's optimality principle holds, the concept itself of optimality becomes unclear and one should use a surrogate of optimality. A normative approach that pretends to be universal fails to provide convincing arguments, for the appropriate behaviour is dictated not only by the non-linear optimizing criterium but also by other subjective factors, such as the attitude towards time consistency, and the importance given to different time intervals and singular points in time. Instead, we consider a philosophical approach more appropriate. As a second contribution of this paper, we further elaborate on this matter, formulate the notion of *preferences consistency* and shed light on the price to be paid in terms of optimality and preferences consistency with each of the three approaches to time inconsistency outlined above.

The remainder of the paper is as follows. In Section 2, we formulate the mean-variance optimization problem, specify the financial market and list the strategies corresponding to each of the three approaches to time inconsistency. In Section 3, we define a suitable intertemporal preferences-driven reward and compare the three strategies according to it at all times between initial time and final time. In Section 4, we extend the analysis to general time-inconsistent problems, starting from the standard linear case where dynamic programming is applicable and, taking the linear case as a benchmark, highlighting pro and contra of each approach to time inconsistency. Section 5 concludes.

2 The mean-variance portfolio selection problem

2.1 Statement of the problem

An investor, endowed with a wealth $x_0 > 0$ at time $t_0 \geq 0$, is faced with a portfolio selection problem on the time horizon $[t_0, T]$. The financial market available for the portfolio allocation problem is the Black-Scholes model (see e.g. Björk (1998)). This consists of two assets, a riskless one, whose price $B(t)$ follows the dynamics:

$$dB(t) = rB(t)dt,$$

where $r > 0$, and a risky asset, whose price dynamics $S(t)$ follows a geometric Brownian motion with drift $\lambda \geq r$ and volatility $\sigma > 0$:

$$dS(t) = \lambda S(t)dt + \sigma S(t)dW(t),$$

where $W(t)$ is a standard Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbf{P})$, with $\mathcal{F}_t = \sigma\{W(s) : s \leq t\}$ the natural filtration. The proportion of portfolio invested in the risky asset at time t is denoted by $u(t)$. The fund at time t under control u, $X^u(t)$, grows according to the following SDE:

$$
\begin{aligned}
&dX^u(t) = X^u(t)\left[u(t)(\lambda - r) + r\right]dt + X^u(t)u(t)\sigma dW(t),\\
&X^u(t_0) = x_0 \geq 0.
\end{aligned} \tag{1}
$$

The investor is a mean-variance optimizer and his aim is to solve the problem

$$
\sup_{u\in\mathcal{U}} J(t_0, x_0, u) \equiv [\mathbb{E}_{t_0,x_0}(X^u(T)) - \alpha\mathbb{V}_{t_0,x_0}(X^u(T))], \tag{2}
$$

where $\alpha > 0$ is a measure of risk aversion, and $\mathcal{U}$ is some set of admissible strategies.

2.2 Three possible approaches

Because Problem (2) is time-inconsistent, the investor can adopt three alternative investment strategies, that are optimal according to different perspectives.

If the investor cares *only* at being time-consistent at time t_0, and does not care of being time-inconsistent at future time $t > t_0$, then he could adopt the precommitment strategy u^p (see Zhou & Li (2000)):

1. precommitment policy $$u^p_{t_0,x_0}(t,x) = \frac{\delta}{\sigma x}\left[x_0 e^{r(t-t_0)} - x + \frac{1}{2\alpha}e^{\delta^2(T-t_0)-r(T-t)}\right], \tag{3}$$

where $\delta = (\lambda - r)/\sigma$ is the market price of risk, or Sharpe ratio. If the investor cares *also* at being time-consistent at future time $t > t_0$, then he can adopt either the time-consistent Nash equilibrium policy u^e (see Basak & Chabakauri (2010) and Björk & Murgoci (2010)):

2. Nash equilibrium policy $$u^e_{t_0,x_0}(t,x) = \frac{\delta}{\sigma x}\frac{1}{2\alpha}e^{-r(T-t)}, \tag{4}$$

or the time-consistent dynamically optimal policy u^d (see Pedersen & Peskir (2016)):

3. dynamically optimal policy $$u^d_{t_0,x_0}(t,x) = \frac{\delta}{\sigma x}\frac{1}{2\alpha}e^{(\delta^2-r)(T-t)}. \tag{5}$$

If the investor selects the precommitment policy, his optimal wealth follows the dynamics:

$$X_t^p = x_0 e^{r(t-t_0)} + \frac{1}{2\alpha} e^{(\delta^2 - r)(T-t)} \left[e^{\delta^2 (t-t_0)} - e^{-\delta(W_t - W_{t_0}) - \frac{\delta^2}{2}(t-t_0)} \right]; \tag{6}$$

if he selects the Nash-equilibrium policy, his optimal wealth follows the dynamics:

$$X_t^e = x_0 e^{r(t-t_0)} + \frac{\delta}{2\alpha} e^{-r(T-t)} \left[\delta(t-t_0) + W_t - W_{t_0} \right]; \tag{7}$$

if he selects the dynamically optimal policy, his optimal wealth follows the dynamics:

$$X_t^d = x_0 e^{r(t-t_0)} + \frac{1}{2\alpha} e^{(\delta^2 - r)(T-t)} \left[e^{\delta^2 (t-t_0)} - 1 + \delta \int_{t_0}^{t} e^{\delta^2 (t-s)} dW_s \right]. \tag{8}$$

In all cases, the value function associated to the adoption of strategy u is

$$\begin{aligned} V^u : \ & [t_0, T] \times \mathbb{R} \ \to \ \mathbb{R} \\ V^u(t,x) \ = \ & \mathbb{E}_{t,x}(X^u(T)) - \alpha \mathbb{V}_{t,x}(X^u(T)). \end{aligned} \tag{9}$$

3 Comparison among different strategies

In this section, we show how to make a comparison among the three approaches illustrated in Section 2.2. The comparison will be performed both at initial time t_0 and also at future time $t > t_0$, by defining proper stochastic reward functions at time t for each strategy.

In order to make a comparison of the three strategies, we will henceforth imagine that we have three investors: the precommitted static one (P-investor), the Nash-equilibrium one (E-investor) and the dynamically optimal one (D-investor), who will adopt the policies (3), (4) and (5), respectively.

3.1 Comparison among strategies at initial time t_0

The comparison among the three strategies at initial time t_0 is straightforward. It suffices to compare the three value functions at t_0.

Adopting the precommitment strategy (3), the value function at time t_0 with wealth x_0 is

$$V^{u^p}(t_0, x_0) = \mathbb{E}_{t_0,x_0}(X^{u^p}(T)) - \alpha \mathbb{V}_{t_0,x_0}(X^{u^p}(T)) = x_0 e^{r(T-t_0)} + \frac{e^{\delta^2(T-t_0)} - 1}{4\alpha}. \tag{10}$$

Adopting the Nash equilibrium strategy (4), the value function at time t_0 with wealth x_0 is

$$V^{u^e}(t_0, x_0) = \mathbb{E}_{t_0,x_0}(X^{u^e}(T)) - \alpha \mathbb{V}_{t_0,x_0}(X^{u^e}(T)) = x_0 e^{r(T-t_0)} + \frac{\delta^2(T-t_0)}{4\alpha}. \tag{11}$$

Adopting the dynamically optimal strategy (5), the value function at time t_0 with wealth x_0 is

$$V^{u^d}(t_0, x_0) = \mathbb{E}_{t_0,x_0}(X^{u^d}(T)) - \alpha \mathbb{V}_{t_0,x_0}(X^{u^d}(T)) = x_0 e^{r(T-t_0)} + \frac{4e^{\delta^2(T-t_0)} - e^{2\delta^2(T-t_0)} - 3}{8\alpha}. \tag{12}$$

The following results hold:

Proposition 3.1. *For all $t_0 \leq T$ and $x_0 \in \mathbb{R}$*

$$V^{u^p}(t_0, x_0) \geq V^{u^e}(t_0, x_0) \geq V^{u^d}(t_0, x_0). \tag{13}$$

The equalities hold if and only if $t_0 = T$, or $\delta = 0$.

Proof. For the first of the two inequalities:

$$V^{u^p}(t_0, x_0) - V^{u^e}(t_0, x_0) = \frac{e^{\delta^2(T-t_0)} - 1 - \delta^2(T-t_0)}{4\alpha} \geq 0,$$

where the inequality holds as an equality if and only if $\delta^2(T-t_0) = 0$, that holds if and only if $t_0 = T$ or $\delta = 0$.
For the second of the two inequalities:

$$V^{u^e}(t_0, x_0) - V^{u^d}(t_0, x_0) = \frac{e^{2\delta^2(T-t_0)} + 2\delta^2(T-t_0) + 3 - 4e^{\delta^2(T-t_0)}}{8\alpha} \geq 0,$$

where the inequality is due to the facts that (i) $\delta^2(T-t_0) \geq 0$, (ii) the function $f(x) = e^{2x} + 2x + 3 - 4e^x$ is strictly increasing, and (iii) $f(0) = 0$. (ii) and (iii) imply that the inequality holds as an equality if and only if $t_0 = T$ or $\delta = 0$. □

Remark 1. We see from Proposition 3.1 that, considering the reward only at time t_0, the P-investor receives a higher value function than the E-investor and the D-investor. This is obvious, because the precommitment strategy by definition maximizes the objective criterium at initial time. The second inequality, already found in Pedersen & Peskir (2016), shows that at initial time t_0 the Nash-equilibrium strategy provides a higher value function than the dynamically optimal strategy. This is also expected and consistent with the fact that, by construction, the Nash-equilibrium strategy is the best among all the time-consistent strategies.

3.2 Comparison among strategies at time t: Reward functions and expected reward functions

The comparison cannot be done only at time t_0, otherwise the obvious answer to the question *what is the best strategy to be adopted* is "the precommitment strategy", that beats all the others from the point of view at t_0. Indeed, if the investor only cares of being mean-variance at time t_0, he will select the precommitment strategy, and will not care of being time-inconsistent after t_0.

Suppose instead that the investor is concerned of being time-consistent at every $t \in [t_0, T]$. Then, he will not change his mean-variance preferences, so that his criterion at every time $t \in [t_0, T]$ will still be to maximize the mean of final wealth while minimizing its variance. Therefore, it is reasonable to assume that the **reward for the mean-variance investor** at time t with wealth x_t adopting strategy u is:

$$J^u(t, x_t) = \mathbb{E}_{t,x_t}(X_T^u) - \alpha \mathbb{V}_{t,x_t}(X_T^u). \tag{14}$$

Let us notice that the comparison among the three investors at time $t > t_0$ is delicate, because, while at time t_0 they have the same wealth x_0, at time $t > t_0$ they have different wealths, because they have been following three different investment strategies from t_0 to t. The P-investor will have wealth X_t^p, the E-investor will have wealth X_t^e, the D-investor will have wealth X_t^d, and in general these levels of wealth will be different from each other. Nevertheless, considering (14), their degree of happiness can be measured by their rewards:

$$J^p(t, X_t^p) = \mathbb{E}_{t,X_t^p}(X_T^p) - \alpha \mathbb{V}_{t,X_t^p}(X_T^p) \qquad \text{reward for the P-investor at time } t$$

$$J^e(t, X_t^e) = \mathbb{E}_{t,X_t^e}(X_T^e) - \alpha \mathbb{V}_{t,X_t^e}(X_T^e) = V^e(t, X_t^e) \qquad \text{reward for the E-investor at time } t$$

$$J^d(t, X_t^d) = \mathbb{E}_{t,X_t^d}(X_T^d) - \alpha \mathbb{V}_{t,X_t^d}(X_T^d) = V^d(t, X_t^d) \qquad \text{reward for the D-investor at time } t\ .$$

We notice that in the last two cases, the reward at time t coincides with the value function at time t, because the Nash-equilibrium and the dynamically optimal strategies are time-consistent. In the first case, the reward at time t does not coincide with the value function, because the investor follows the time-inconsistent strategy decided at time t_0 and therefore does not optimize the mean-variance criterion at any time $t > t_0$.

Notice that, standing at time t_0, the rewards $J^p(t, X_t^p)$, $V^e(t, X_t^e)$ and $V^d(t, X_t^d)$ that refer to time $t > t_0$ are random variables. However, it is possible to compare them standing at time t_0, by comparing their time-t_0 expectations. We thus define the following expected reward functions:

$$R(t, u^p; t_0, x_0) = \mathbb{E}_{t_0,x_0}\left(J^p(t, X_t^p)\right) \tag{15}$$

$$R(t, u^e; t_0, x_0) = \mathbb{E}_{t_0,x_0}\left(V^e(t, X_t^e)\right) \tag{16}$$

$$R(t, u^d; t_0, x_0) = \mathbb{E}_{t_0,x_0}\left(V^d(t, X_t^d)\right). \tag{17}$$

In general, the expected value in t_0 of the reward at time t of the investor who follows the strategy u is

$$R(t, u; t_0, x_0) = \mathbb{E}_{t_0,x_0}\left[\mathbb{E}_{t,X_t^u}(X_T^u) - \alpha \mathbb{V}_{t,X_t^u}(X_T^u)\right]. \tag{18}$$

3.2.1 Comparison of expected reward functions at times t_0 and T

For the comparison among the expected rewards of the three strategies at times t_0 and T, the following results hold:

Proposition 3.2. *If $t_0 < T$, $\delta \neq 0$, and $x_0 \in \mathbb{R}$, then*

$$R(t_0, u^p; t_0, x_0) > R(t_0, u^e; t_0, x_0) > R(t_0, u^d; t_0, x_0) \tag{19}$$

and

$$R(T, u^p; t_0, x_0) = R(T, u^d; t_0, x_0) > R(T, u^e; t_0, x_0) \tag{20}$$

Proof. For every strategy u

$$R(t_0, u; t_0, x_0) = \mathbb{E}_{t_0,x_0}\left[\mathbb{E}_{t_0,x_0}(X_T^u) - \alpha \mathbb{V}_{t_0,x_0}(X_T^u)\right] = V^u(t_0, x_0).$$

Claim (19) follows by Proposition 3.1.
For every strategy u,

$$\mathbb{E}_{T,X_T^u}(X_T^u) - \alpha \mathbb{V}_{T,X_T^u}(X_T^u) = X_T^u \quad \Rightarrow \quad R(T, u; t_0, x_0) = \mathbb{E}_{t_0,x_0}\left(X_T^u\right).$$

Using the dynamics (6), (7) and (8), we get:

$$\mathbb{E}_{t_0,x_0}\left(X_T^p\right) = \mathbb{E}_{t_0,x_0}\left(X_T^d\right) = x_0 e^{r(T-t_0)} + \frac{1}{2\alpha}\left(e^{\delta^2(T-t_0)} - 1\right) \tag{21}$$

and

$$\mathbb{E}_{t_0,x_0}\left(X_T^d\right) = x_0 e^{r(T-t_0)} + \frac{\delta^2(T-t_0)}{2\alpha}.$$

If $\delta^2(T-t_0) \neq 0$, then $e^{\delta^2(T-t_0)} - 1 > \delta^2(T-t_0)$. Therefore, claim (20) is obtained. $\square$

Remark 2. Result (20) was already found by Pedersen & Peskir (2016), in a first attempt of comparison among the two time-consistent strategies for the mean-variance problem. In their comparative analysis, they did not provide detailed argumentation for the reasonableness of the method used for the comparison.

Remark 3. Proposition 3.2 indicates that the static precommitment strategy is apparently never inferior to the other strategies. Similarly to Remark 1, this is due to the fact that the criterion $R(t, u; t_0, x_0)$ illustrates the degree of happiness at time t of the mean-variance optimizer *as measured at time* t_0. No surprise that the precommitment strategy gives an outcome at least as good as the other strategies. This will be confirmed by Theorem 3.4 in Section 3.2.3. The comparison between the two time-consistent strategies is commented in the next section.

3.2.2 Comparison of expected reward functions at time t: comparison between time-consistent strategies

From Proposition 3.2, we see that the Nash-equilibrium strategy provides a higher reward than the dynamic strategy at initial time t_0, and a lower reward than the dynamic strategy at final time T, suggesting the occurrence of a swap between the two strategies at some time $t^* \in (t_0, T)$. The next theorem shows the existence and uniqueness of such a break even point.

Theorem 3.3. *If $t_0 < T$, $\delta \neq 0$, and $x_0 \in \mathbb{R}$, there exists one and only one point $t^* \in (t_0, T)$ such that*

$$R(t^*, u^e; t_0, x_0) = R(t^*, u^d; t_0, x_0). \tag{22}$$

The break even point t^* is the unique solution of the equation

$$e^{2\delta^2(T-t)} - \left(4e^{\delta^2(T-t_0)} + 4\delta^2 t_0 - 3\right) + 2\delta^2(T-t) = 0. \tag{23}$$

Proof. By defining the function

$$\Delta R^{ed}(t) = R(t, u^e; t_0, x_0) - R(t, u^d; t_0, x_0), \tag{24}$$

claim (22) is equivalent to prove the existence and the uniqueness of a root of the function $\Delta R^{ed}(t)$ in the interval (t_0, T). Proposition 3.2 yields:

$$\Delta R^{ed}(t_0) > 0 \qquad \text{and} \qquad \Delta R^{ed}(T) < 0. \tag{25}$$

Recalling that both the Nash-equilibrium strategy and the dynamic strategy are time-consistent, we can obtain $V^e(t, X_t^e)$ and $V^d(t, X_t^d)$ just by replacing t_0 with t, and x_0 with X_t^e and X_t^d in (11) and (12), respectively.

$$V^e(t, X_t^e) = X_t^e e^{r(T-t)} + \frac{\delta^2 (T-t)}{4\alpha}. \tag{26}$$

$$V^d(t, X_t^d) = X_t^d e^{r(T-t)} + \frac{4e^{\delta^2(T-t)} - e^{2\delta^2(T-t)} - 3}{8\alpha}. \tag{27}$$

Using (16), (17), (26) and (27), we have

$$R(t, u^e; t_0, x_0) = \mathbb{E}_{t_0,x_0}\left(X_t^e\right) e^{r(T-t)} + \frac{\delta^2(T-t)}{4\alpha}, \tag{28}$$

and

$$R(t, u^d; t_0, x_0) = \mathbb{E}_{t_0,x_0}\left(X_t^d\right) e^{r(T-t)} + \frac{4e^{\delta^2(T-t)} - e^{2\delta^2(T-t)} - 3}{8\alpha}. \tag{29}$$

Using the closed-form expressions (7) and (8), we get

$$\mathbb{E}_{t_0,x_0}\left(X_t^e\right) = x_0 e^{r(t-t_0)} + \frac{\delta^2}{2\alpha}(t-t_0)e^{-r(T-t)}, \tag{30}$$

and

$$\mathbb{E}_{t_0,x_0}\left(X_t^d\right) = x_0 e^{r(t-t_0)} + \frac{1}{2\alpha} e^{(\delta^2 - r)(T-t)}\left[e^{\delta^2(t-t_0)} - 1\right]. \tag{31}$$

By plugging (30) and (31) into (28) and (29), respectively, and using (24), after some simplifications we get

$$\Delta R^{ed}(t) = \frac{e^{2\delta^2(T-t)} + (3 - 4e^{\delta^2(T-t_0)} - 4\delta^2 t_0) + 2\delta^2(T-t)}{8\alpha}. \tag{32}$$

The function $\Delta R^{ed}(t)$ is continuous and, due to (25), it takes different signs at the extremes of $[t_0, T]$. Moreover,

$$\frac{d}{dt}\left(\Delta R^{ed}(t)\right) = -2\delta^2(e^{2\delta^2(T-t)} + 1) < 0,$$

that implies that $\Delta R^{ed}(t)$ is also strictly decreasing. Therefore, there exists a unique root of $\Delta R^{ed}(t)$ in (t_0, T) and is given by the unique t^* that nullifies the numerator of (32). This concludes the proof. □

Remark 4. Let us notice that Theorem 3.3 implies that

$$R(t, u^e; t_0, x_0) > R(t, u^d; t_0, x_0) \qquad \forall\, t \in [t_0, t^*)$$

and

$$R(t, u^e; t_0, x_0) < R(t, u^d; t_0, x_0) \qquad \forall\, t \in (t^*, T],$$

meaning that, among the time-consistent strategies for the mean-variance problem, the Nash-equilibrium strategy provides on average a higher reward until time t^*, while the dynamic strategy provides on average a higher reward from time t^* onwards. This result is meaningful and suggests that the importance allocated by the decision-maker to different points in time should affect his attitude toward time inconsistency, and should play a part in the entire decision-making process. The first inequality is also consistent with the fact that the Nash equilibrium strategy is the *best* plan among those that are time-consistent: the criterium "best" here indicates the best at initial time t_0, and indeed $R(t_0, u^e; t_0, x_0) > R(t_0, u^d; t_0, x_0)$. Clearly, selecting the best time-consistent plan only from the view point at time t_0 can be considered insufficient for a decision-maker who is intrinsically an intertemporal optimizer.

While a detailed analysis of the break even point is beyond the scope of this paper, we notice that t^* does not depend neither on wealth nor on risk aversion α. It depends only on the market price of risk δ, on the time horizon T and on initial time t_0. Table 1 reports the value of t^* with some typical values of δ and T, when the initial time is $t_0 = 0$: $\delta = 0.1, 0.2, 0.3, 0.4, 0.5$ and $T = 10, 20, 30, 40$ (see Vigna (2014)).

The break even point increases with δ and with T. This is expected, because when $\delta = 0$ the three strategies collapse into the riskless strategy (where all the fund is invested continuously in the riskless asset) and there is no difference between portfolios, value functions and reward functions. The value and reward functions are trivially the same also in the degenerate case $T = 0$. We notice that t^* varies between one week for 10 years and small δ ($\delta = 0.1$) and 17 years for 40 years and high δ ($\delta = 0.5$). In relative terms, t^* ranges between 0.01% and 2% of T for small δ, and between 25% and 43% of T for high δ.

Sharpe ratio	Time horizon T			
δ	10	20	30	40
0.1	0.02	0.12	0.4	0.91
0.2	0.23	1.43	3.73	6.87
0.3	0.85	3.85	7.99	12.61
0.4	1.72	5.92	10.73	15.68
0.5	2.53	7.26	12.23	17.23

Table 1: Break even point t^* with different Sharpe ratios δ and time horizons T (in years).

The intuition behind the fact that t^* increases with δ and T can be the following. The dynamically optimal strategy is more aggressive than the Nash-equilibrium strategy by a factor $e^{\delta^2(T-t)}$ (see (4) and (5)). When δ or T increase, the dynamically optimal strategy becomes even more aggressive than the Nash-equilibrium strategy. A more aggressive strategy increases to a larger extent the variance of final wealth, and to a lower extent the mean of final wealth, hence the global effect is a reduction of the intertemporal reward, and this happens with higher severity with the dynamically optimal strategy than with the Nash-equilibrium strategy. Therefore, an increase in δ or T moves the break even point (that is the time point when the expected rewards of the two strategies switch) more far in the future, i.e., t^* increases.

3.2.3 Comparison of expected reward functions at time t: comparison among time-inconsistent and time-consistent strategies

We now intend to compare the precommitment time-inconsistent strategy with the two time-consistent strategies considered above. In order to do this, we need to calculate the expected value in t_0 of the reward function for the precommitment strategy (see (15)):

$$R(t, u^p; t_0, x_0) = \mathbb{E}_{t_0,x_0}\left[J^p(t, X_t^p)\right] = \mathbb{E}_{t_0,x_0}\left[\mathbb{E}_{t,X_t^p}(X_T^p) - \alpha \mathbb{V}_{t,X_t^p}(X_T^p)\right].$$

As we noticed in Section 3.2, and differently from the other two cases, in the precommitment case the value of the reward $J^p(t, X_t^p)$ *does not* coincide with the value function $V^p(t, X_t^p)$ calculated at time t, because the precommitment strategy is *not* time-consistent. Indeed, if the precommitted investor finds himself at time t with fund X_t^p, he will still apply the precommitted strategy (3). Therefore, the value of his reward $J^p(t, X_t^p)$ will *not* be the supremum of all possible values, the value function at (t, X_t^p): the value function could be reached only by applying a *new* precommitment strategy with starting point t and initial wealth X_t^p. In other words, for the static precommitted investor the "value function" has meaning only at time t_0 and has no meaning after t_0. However, inspired by (14), we can still assume that the reward for the precommitted investor at time t with wealth X_t^p is given by $J^p(t, X_t^p)$, and we can calculate its expectation at t_0, $\mathbb{E}_{t_0,x_0}\left[J^p(t, X_t^p)\right]$.

If the static investor has wealth x_t^p at time $t > t_0$, and he adopts the investment strategy (3), then his future wealth at time $\tau > t$ follows the dynamics given by the SDE:

$$\begin{cases} dX_\tau^p = X_\tau^p \left[r + u^p(\tau, X_\tau^p)(\lambda - r)\right] d\tau + X_\tau^p\, u^p(\tau, X_\tau^p)\, \sigma dW_\tau \\ X_t^p = x_t^p \end{cases} \tag{33}$$

where

$$u^p(\tau, x) = \frac{\delta}{\sigma}\frac{1}{x}\left(Ke^{-r(T-\tau)} - x\right)$$

with

$$K = x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0)}. \tag{34}$$

Let us define the new stochastic process

$$Z_\tau = K^{-r(T-t_0)} - X_\tau^p e^{-r(\tau-t_0)}, \tag{35}$$

with K given by (34). By applying Ito's lemma to Z_τ, we get its dynamics for $\tau > t$:

$$\begin{cases} dZ_\tau = -\delta^2 Z_\tau d\tau - \delta Z_\tau dW_\tau \\ Z_t = Ke^{-r(T-t_0)} - X_t^p e^{-r(t-t_0)}. \end{cases} \tag{36}$$

Therefore

$$Z_\tau = Z_t\, e^{-\frac{3}{2}\delta^2(\tau-t)-\delta(W_\tau-W_t)} \tag{37}$$

where Z_t is given by (36). Plugging (37) into (35), after some simplifications, we get the solution to (33), i.e. the dynamics of X^p_τ for $\tau > t$:

$$X^p_\tau = \left[x_0 e^{r(\tau-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0)-r(T-\tau)}\right] + \\ + \left[x^p_t - x_0 e^{r(t-t_0)} - \frac{1}{2\alpha} e^{\delta^2(T-t_0)-r(T-t)}\right] e^{(r-\frac{3}{2}\delta^2)(\tau-t)} \cdot e^{-\delta(W_\tau - W_t)}. \quad (38)$$

Therefore

$$X^p_T = \left[x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0)}\right] + \\ + \left[x^p_t - x_0 e^{r(t-t_0)} - \frac{1}{2\alpha} e^{\delta^2(T-t_0)-r(T-t)}\right] e^{(r-\frac{3}{2}\delta^2)(T-t)} \cdot e^{-\delta(W_T - W_t)}. \quad (39)$$

Thus

$$\mathbb{E}_{t,x^p_t}\left(X^p_T\right) = \left(1 - e^{-\delta^2(T-t)}\right)\left(x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0)}\right) + x^p_t e^{(r-\delta^2)(T-t)}, \quad (40)$$

and

$$\mathbb{V}_{t,x^p_t}\left(X^p_T\right) = \left[x^p_t - x_0 e^{r(t-t_0)} - \frac{1}{2\alpha} e^{\delta^2(T-t_0)-r(T-t)}\right]^2 \left(e^{(2r-\delta^2)(T-t)} - e^{2(r-\delta^2)(T-t)}\right). \quad (41)$$

Therefore, the reward at time t for the precommitted investor is

$$J^p(t, X^p_t) = \mathbb{E}_{t,X^p_t}(X^p_T) - \alpha \mathbb{V}_{t,X^p_t}(X^p_T) = \\ \left(1 - e^{-\delta^2(T-t)}\right)\left(x_0 e^{r(T-t_0)} + \frac{1}{2\alpha} e^{\delta^2(T-t_0)}\right) + X^p_t e^{(r-\delta^2)(T-t)} + \\ - \alpha \left[X^p_t - x_0 e^{r(t-t_0)} - \frac{1}{2\alpha} e^{\delta^2(T-t_0)-r(T-t)}\right]^2 \left(e^{(2r-\delta^2)(T-t)} - e^{2(r-\delta^2)(T-t)}\right). \quad (42)$$

Taking expectation at time t_0, we get:

$$\mathbb{E}_{t_0,x_0}\left(J^p(t,X_t^p)\right) = \mathbb{E}_{t_0,x_0}\left[\mathbb{E}_{t,X_t^p}(X_T^p) - \alpha\mathbb{V}_{t,X_t^p}(X_T^p)\right] =$$
$$\left(1-e^{-\delta^2(T-t)}\right)\left(x_0e^{r(T-t_0)}+\frac{1}{2\alpha}e^{\delta^2(T-t_0)}\right)+\mathbb{E}_{t_0,x_0}\left(X_t^p\right)e^{(r-\delta^2)(T-t)}+$$
$$-\alpha\left[\mathbb{E}_{t_0,x_0}\left(X_t^p\right)-x_0e^{r(t-t_0)}-\frac{1}{2\alpha}e^{\delta^2(T-t_0)-r(T-t)}\right]^2\left(e^{(2r-\delta^2)(T-t)}-e^{2(r-\delta^2)(T-t)}\right). \quad (43)$$

Using the dynamics (6), we get

$$\mathbb{E}_{t_0,x_0}(X_t^p) = x_0e^{r(t-t_0)}+\frac{1}{2\alpha}e^{(\delta^2-r)(T-t)}\left[e^{\delta^2(t-t_0)}-1\right]. \quad (44)$$

Plugging (44) into (43), after some simplifications, we get the expected value in t_0 of the reward function at time t for the precommitted investor:

$$R(t,u^p;t_0,x_0) = \mathbb{E}_{t_0,x_0}\left(J^p(t,X_t^p)\right) = x_0e^{r(T-t_0)}+\frac{1}{4\alpha}\left[2e^{\delta^2(T-t_0)}-e^{\delta^2(T-t)}-1\right]. \quad (45)$$

Remark 5. Notice that using (45) to calculate $R(t,u^p;t_0,x_0)$ in t_0 and T, we get

$$R(t_0,u^p;t_0,x_0) = x_0e^{r(T-t_0)}+\frac{1}{4\alpha}\left[e^{\delta^2(T-t_0)}-1\right],$$

which, as expected, coincides with (10), and

$$R(T,u^p;t_0,x_0) = x_0e^{r(T-t_0)}+\frac{1}{2\alpha}\left[e^{\delta^2(T-t_0)}-1\right],$$

which, as expected, coincides with (21).

We are now ready to compare the static precommitment strategy with the time-consistent strategies. The following results hold:

Theorem 3.4. *If $t_0 < T$, $\delta \neq 0$, and $x_0 \in \mathbb{R}$,*

$$R(t,u^p;t_0,x_0) > R(t,u^e;t_0,x_0) \qquad \textit{for all } t\in[t_0,T], \quad (46)$$

$$R(t, u^p; t_0, x_0) > R(t, u^d; t_0, x_0) \qquad \textit{for all } t \in [t_0, T), \tag{47}$$

and

$$R(T, u^p; t_0, x_0) = R(T, u^d; t_0, x_0). \tag{48}$$

Proof. By defining the function

$$\Delta R^{pe}(t) = R(t, u^p; t_0, x_0) - R(t, u^e; t_0, x_0), \tag{49}$$

claim (46) is equivalent to the strict positivity of $\Delta R^{pe}(t)$ over $[t_0, T]$. By plugging (45) and (28) into (49), we get:

$$\Delta R^{pe}(t) = \frac{1}{4\alpha}\left[2e^{\delta^2(T-t_0)} - e^{\delta^2(T-t)} - 1 - \delta^2(T + t - 2t_0)\right]. \tag{50}$$

The first derivatives of $\Delta R^{pe}(t)$ is

$$\frac{d\left(\Delta R^{pe}(t)\right)}{dt} = \frac{\delta^2}{4\alpha}\left(e^{\delta^2(T-t)} - 1\right) > 0 \qquad \text{for } t \in (t_0, T),$$

implying that $\Delta R^{pe}(t)$ is increasing over (t_0, T). Note that $\Delta R^{pe}(t_0) > 0$ by Proposition 3.2. Then, claim (46) follows.

By defining the function

$$\Delta R^{pd}(t) = R(t, u^p; t_0, x_0) - R(t, u^d; t_0, x_0), \tag{51}$$

claim (47) is equivalent to the strict positivity of $\Delta R^{pd}(t)$ over $[t_0, T)$, and claim (48) is equivalent to $\Delta R^{pd}(T) = 0$. By plugging (45) and (29) into (51), we get:

$$\Delta R^{pd}(t) = \frac{1}{8\alpha}\left[e^{\delta^2(T-t)} - 1\right]^2. \tag{52}$$

Then, claims (47) and (48) follow easily. □

From Theorem 3.4 we see that the static precommitment strategy provides a higher expected reward for any time $t \in [t_0, T]$ than the two time-consistent strategies. As we observed in Remark 1 and Remark 3, this is due to the fact that the expectation of the future reward is done at time t_0. From Theorem 3.4 we also see that the difference $\Delta R^{pe}(t)$ between precommitment and Nash-equilibrium strategies is *increasing* over $[t_0, T]$, and reaches its maximum in T: this means that the positive difference between expected rewards becomes larger when time passes. On the contrary, the difference $\Delta R^{pd}(t)$ between precommitment and dynamic strategies is *decreasing* over $[t_0, T]$, and is null in T: this means that the positive difference between expected rewards becomes smaller when time passes, and disappears at terminal time T.

4 On "optimality" for time-inconsistent problems

In this section, we elaborate on the meaning of optimality and preferences consistency for problems that give rise to time inconsistency. First, we start by revisiting the standard linear optimization problem where dynamic programming can be applied. With this kind of problems the optimal solution has all the desirable properties that one wishes to have. Then, we introduce non-linear problems that do not permit application of dynamic programming and we illustrate the current approaches to them. We show that no strategy carries all the desirable features that the optimal solution to a linear problem has, and for each strategy we analyze what are the features in common with the optimal strategy of a linear problem. This might help understanding the essence of each approach, and might help individuals in selecting their most appropriate approach to time inconsistency.

4.1 Linear problems

Consider the following optimization problem that the investor wishes to solve at time t_0 with wealth x_0.

$$\text{Problem } \mathcal{P}^L_{t_0,x_0}: \qquad \sup_u J(t_0, x_0, u) = \sup_u \mathbb{E}_{t_0,x_0}\left[\int_{t_0}^{T} U^1(s, X_s, u_s)ds + U^2(X_T)\right]. \tag{53}$$

According to Björk & Murgoci (2010), if the running utility $U^1(\cdot)$ and the terminal utility $U^2(\cdot)$ **do not** depend on the initial point (t_0, x_0), then Problem $\mathcal{P}^L_{t_0,x_0}$ is "linear" and dynamic programming is applicable. By dynamic programming, in order to approach Problem $\mathcal{P}^L_{t_0,x_0}$ one should consider the more general problem to be solved at time t with wealth x:

$$\text{Problem } \mathcal{P}^L_{t,x}: \qquad \sup_u J(t,x,u) = \sup_u \mathbb{E}_{t,x}\left[\int_t^T U^1(s,X_s,u_s)ds + U^2(X_T)\right], \tag{54}$$

write and solve the associated Hamilton-Jacobi-Bellman (HJB) equation to find the value function

$$V(t,x) = \sup_u J(t,x,u),$$

and the optimal control law $u^*(s,y;t,x)$ ($s \geq t, y \in \mathbb{R}$) as the maximizing control of the HJB equation. Notice that the optimal control is a function of time $s \geq t$ and wealth $y \in \mathbb{R}$, but also a function of the initial point (t,x). Once the Problem $\mathcal{P}^L_{t,x}$ is solved, the initial problem $\mathcal{P}^L_{t_0,x_0}$ is solved as a special case, by replacing (t,x) with (t_0,x_0). In this case, the Bellman's optimality principle holds: the optimal control law $u^*(s,y;t_0,x_0)$ is optimal not only on $[t_0,T]$ but also on every subinterval $[\tau,T]$ with $\tau > t_0$. This means that the optimal strategy for the new problem $\mathcal{P}^L_{\tau,x_\tau}$ at time τ with initial wealth x_τ coincides with the restriction on $[\tau,T]$ of the optimal strategy found at initial time t_0:

$$\text{argmax}_u J(\tau,x_\tau,u) = u^*(s,y;\tau,x_\tau) = u^*(s,y;t_0,x_0)\big|_{[\tau,T]}. \tag{55}$$

Since this happens for every $\tau \in (t,T]$ and every $x_\tau \in \mathbb{R}$, the optimal control law is the same no matter what the initial time-wealth are:

$$u^*(s,y;\tau,x_\tau) = u^*(s,y;t_0,x_0)\big|_{[\tau,T]} = u^*(s,y).$$

The optimality of the right queue of the strategy $u^*(s,y;t,x)$ is crucial for the meaning of time consistency and preferences consistency. Indeed, at each instant of time $t > t_0$ the investor that applies the optimal strategy decided at initial time t_0 is also playing what would be the optimal strategy for the same problem at time t. In other words, by obeying to the initial plan, *the investor is always consistent with his own preferences*: he is preferences-consistent. Indeed, he does not change the optimization problem and keeps on solving it, with revaluated conditions, hence he does not change his preferences.

In this case, we will say that the optimal control law $u^*(s,y)$ is the *time-consistent solution to Problem* $\mathcal{P}^L_{t_0,x_0}$, and we will also say that the family of problems $\left\{\mathcal{P}^L_{t,x}\right\}_{t\in[t_0,T],x\in\mathbb{R}}$ is *time-consistent* with optimal policy $u^*(s,y)$. Indeed, when talking about time consistency, one cannot avoid to talk about an intertemporal family of problems and the related optimal control law. This analysis makes clear that for an intertemporal family of problems, the notion of time consistency is equivalent to the Bellman's optimality principle.

The situation becomes more complicated when the running utility depends on initial time or wealth, or the bequest function includes also a non-linear function of expected final wealth. This is the case considered in the next section.

4.2 Non-linear problems

Adopting the framework introduced by Björk & Murgoci (2010), let us now suppose that an investor wants to solve the following non-linear problem[1]

$$\text{Problem } \mathcal{P}^{NL}_{t_0,x_0}: \quad \sup_u J(t_0,x_0,u) = \sup_u \left\{\mathbb{E}_{t_0,x_0}\left[\int_{t_0}^T U^1(s,X_s,u_s)ds + U^2(X_T)\right] + U^3\left[\mathbb{E}_{t_0,x_0}(X_T)\right]\right\} \tag{56}$$

where $U^3(\cdot)$ is a non-linear utility function. The presence of the non-linear term $U^3\left[\mathbb{E}_{t_0,x_0}(X_T)\right]$ prevents the straight use of dynamic programming.

Possible approaches to the non-linear problem (56) are the precommitment approach, the dynamic approach and the Nash-equilibrium approach.

Precommitment

One fixes the initial point (t_0,x_0) and finds, if it exists, the control law $\hat{u}$ that maximizes only $J(t_0,x_0,u)$, i.e., the precommitment strategy:

Definition 4.1. *Given the non-linear Problem* $\mathcal{P}^{NL}_{t_0,x_0}$ *as in* (56)*, the strategy* $\hat{u}$ *that maximizes* $J(t_0,x_0,u)$*, i.e., the control* $\hat{u}$ *such that*

$$J(t_0,x_0,\hat{u}) = \sup_u J(t_0,x_0,u)$$

[1] For simplicity, we consider only the case when the bequest function includes a non-linear function of expected final wealth. The case in which the running utility depends on initial time or wealth is similar.

if it exists, is called the precommitment strategy *and is indicated as*

$$\hat{u}(s, y; t_0, x_0). \tag{57}$$

Because in this kind of problems dynamic programming cannot be applied and the Bellman's principle does not hold, by adopting $\hat{u}$, one disregards the fact that at a later point in time $\tau \in (t_0, T]$ with wealth x_τ the control law $\hat{u}(s, y; t_0, x_0)$ is not optimal for the criterion $J(\tau, x_\tau, u)$. In other words,

$$\text{argmax}_u J(\tau, x_\tau, u) = \hat{u}(s, y; \tau, x_\tau) \neq \hat{u}(s, y; t_0, x_0)\big|_{[\tau,T]},$$

while there would be equality with validity of the Bellman's principle (see equation (55)). In other words, the precommitment strategy *depends essentially on the initial point* (t_0, x_0).

This is the reason why the strategy is named precommitment strategy: the decision-maker standing at time t_0 should precommit to follow the strategy $\hat{u}(s, y; t_0, x_0)$ even if he knows that at later time he is still solving the original problem $\mathcal{P}^{NL}_{t_0,x_0}$, but *not* the translated problem at time τ, $\mathcal{P}^{NL}_{\tau,x_\tau}$. Therefore, the investor over the time interval $[t_0, T]$ is solving the original problem, but at any time $\tau \in (t_0, T]$ he is not consistent to his own preferences, because the strategy adopted on the interval $[\tau, T]$ is not the optimal solution to the problem $\mathcal{P}^{NL}_{\tau,x_\tau}$. Therefore, he is preferences-consistent at time t_0 but is not preferences-consistent at any time $t > t_0$. Clearly, the precommitment strategy is the **best** strategy standing at time t_0. The problem of precommitment is about time consistency: the precommitted decision-maker only cares about initial time t_0 and final time T, disregarding the time interval (t_0, T). In other words, the precommitment strategy is closer in spirit to the single-period Markovitz framework than to the continuous-time intertemporal setup: only t_0 and T matter, what happens at any time $t \in (t_0, T)$ does not matter. The interval (t_0, T) goes into a black box and the investor is consistent with his own preferences only at initial time t_0. In this respect, the name "static" given by some authors to identify the precommitment strategy (Pedersen & Peskir (2016)) or the optimization problem as defined in (t_0, x_0) only (Karnam et al. (2016)), could not be more appropriate.

Dynamic optimality

We illustrate the construction of the dynamically optimal strategy in 4 steps.
Step 1. A family of non-linear problems $\{\mathcal{P}_{t,x}^{NL}\}_{t\in[t_0,T],x\in\mathbb{R}}$ as in (56) is given.
Step 2. Assume that for fixed initial point (t_0, x_0) the precommitment optimal solution maximizing the criterion $J(t_0, x_0, u)$ exists and is given by (see (57)):

$$\hat{u}(s, y; t_0, x_0). \tag{58}$$

Step 3. Define the new control map

$$\tilde{u}(s, y) = \hat{u}(s, y; s, y), \tag{59}$$

where the right hand side of (59) is the function (58) by replacing t_0 with s and x_0 with y.
Step 4. The strategy $\tilde{u}(s, y)$ is called the *dynamically optimal strategy*.

At each $t \in [t_0, T]$ the dynamic strategy coincides with the *first* control of the precommitment strategy solution to $\{\mathcal{P}_{t,x}^{NL}\}$, but deviates from it immediately after, at time t^+. Therefore, the dynamically optimal investor can be seen as the *continuous reincarnation* of the precommitted investor. Notice that over $[t_0, T]$ the control map $\tilde{u}(s, y)$ is not optimal for any specific problem, but is **instantaneously optimal** at each $t \in [t_0, T]$, so it is instantaneously optimal for infinitely many non-linear problems. Therefore, unlike the precommitted investor who is optimal over $[t_0, T]$ (and only over $[t_0, T]$), the dynamic investor is never optimal over a specified time interval; however, unlike the precommitted investor who is consistent with his own preferences only at initial time t_0, the dynamic investor is continuously consistent with his own preferences. Therefore, the dynamically optimal investor is never optimal over any time interval, but is always instantaneously preferences-consistent.

Nash equilibrium

According to consistent planning one should choose "the best plan among those that he will actually follow". The construction of this strategy is based on the game theoretic interpretation that to each point in time t is associated a player who can choose the control at time t. At time $s > t$ there is another player who chooses the control at time s. The key of this approach is to search a Nash subgame perfect equilibrium among the continuum of players $[t_0, T]$. A strategy $\hat{u}$ is an equilibrium strategy if, given that all players in $(t, T]$ will play $\hat{u}$ then also player t finds it optimal to play $\hat{u}$. The equilibrium strategy is found by solving an extended Hamilton-Jacobi-Bellman equation for the value function, see Björk & Murgoci (2010), Proposition 4.1 and Theorem 4.1. Björk & Murgoci (2010) also prove that it is possible to associate to each time-inconsistent problem a standard time-consistent problem such that (i) the optimal value function of the standard problem is equal to the equilibrium value function of the time-inconsistent problem; (ii) the optimal control law of the standard problem is equal to the equilibrium strategy of the time-inconsistent problem (see Björk & Murgoci (2010), Proposition 5.1). This implies that the Nash-equilibrium strategy associated to the non-linear problem (56) coincides with the optimal solution to the standard linear problem

$$\text{Problem } \mathcal{P}^{ass-NL}_{t_0,x_0} : \quad \sup_u \left\{ \mathbb{E}_{t_0,x_0} \left[\int_{t_0}^T U^4(s, X_s, u_s) ds + U^5(X_T) \right] \right\} \tag{60}$$

where $U^4(\cdot)$ and $U^5(\cdot)$ are some functions which are not necessarily easy to find. For instance, in the case of the mean-variance preferences, $U^4(\cdot) \equiv 0$ while $U^5(\cdot)$ is the CARA utility function. Indeed, the optimal solution to the standard linear optimization problem

$$\text{Problem } \mathcal{P}^{ass-MV}_{t_0,x_0} : \quad \sup_u \mathbb{E}_{t_0,x_0} \left[-\frac{1}{2\alpha} e^{-2\alpha X_T} \right]$$

coincides with the Nash-equilibrium strategy (4) (see also Basak & Chabakauri (2010), Remark 1). Björk & Murgoci (2010) comment that there is no gain by enlarging the class of consumer behaviour to time-inconsistent preferences, because every time-inconsistent strategy can be replicated by some time-consistent utility function. Instead, we comment this result from a different angle. For a non-linear problem (56) the Nash-equilibrium approach is equivalent to apply the solution to the associated linear problem (60). This means that in order to be time-consistent, the investor has to choose a different objective functional, in other words, different preferences. For the mean-variance problem, the investor who chooses the Nash-equilibrium approach applies a strategy that is optimal according to a *different*

criterium than the mean-variance one, namely the exponential preferences. The price to be paid in order to be time-consistent consists in changing preferences.
Therefore, the Nash-equilibrium individual is optimal over every subinterval of $[t_0, T]$ but with different preferences than the original ones. He is never preferences-consistent, not even at time t_0.

To sum up, the non-applicability of dynamic programming and the violation of the Bellman's optimality principle imposes an unavoidable price to be paid to decision-makers. The price is different depending on the approach selected. By adopting the precommitment approach the investor solves a kind of static Markovitz problem over $[t_0, T]$, but is not preferences-consistent after t_0. By adopting the dynamically optimal approach the investor should give up the idea of optimality over a time interval. By adopting the Nash-equilibrium strategy the investor should give up his original preferences.

5 Concluding remarks

When an intertemporal optimization problem over a time period $[t_0, T]$ can be solved using dynamic programming, then the optimal solution presents two important desirable features.

First, the optimal strategy is *optimal* over $[t_0, T]$ and over every subinterval $[t, T]$; second, the decision-maker is *preferences-consistent* at any time, i.e., he is always consistent with his original preferences, because at each instant of time t by adopting the optimal strategy found at t_0 he is still solving the same initial problem translated at time t with revaluated conditions at time t.

When an intertemporal optimization problem does not permit application of dynamic programming, then the two features described above cannot hold simultaneously. According to the existing literature, we say that the problem gives raise to time inconsistency. This kind of problem and the resulting time inconsistency can be addressed in different ways. With the precommitment approach, one keeps the first property, so he is optimal over $[t_0, T]$, but gives up the second feature, i.e., he is not preferences-consistent, apart at initial time t_0. With the dynamically optimal approach, the investor gives up optimality over every time interval, but is instantaneously preferences-consistent at any time. With the Nash-equilibrium approach, the decision-maker is optimal over $[t_0, T]$ and every subinterval, but is not preferences-consistent, because he is solving a linear time-consistent optimization problem but with *different* preferences, so he has to give up his original preferences.

In the special case of the mean-variance portfolio selection problem over $[t_0, T]$, an intertemporal criterium based on the original mean-variance preferences and defined as the expectation at initial time t_0 of future reward at time t suggests that the precommitment approach should be preferable if the individual is static and does not care of being time-inconsistent. Among the two time-consistent strategies currently available, the same criterium indicates that the Nash-equilibrium approach should be preferable until a time point $t^* \in (t_0, T)$, and the dynamic approach should be preferable from t^* onwards.

In general, it seems quite hard to argue that one of the three approaches to time inconsistency currently available should be unambiguously preferable to the others for all individuals and for all non-linear optimization problems. Rather the opposite: each approach has its own pro and contra, and the appropriate strategy depends not only on the non-linear optimizing criterium but also on other subjective factors, such as the attitude towards time consistency, the importance given to different time intervals and singular points in time, as well as the importance given to one's own original preferences. A normative approach that pretends to be universal fails to provide convincing arguments that hold for all individuals, while we believe that a philosophical approach to discuss appropriateness of each approach is more suitable.

References

Bajeux-Besnainou, I. & Portait, R. (1998). Dynamic asset allocation in a mean-variance framework, *Management Science* **44**: S79–S95.

Basak, S. & Chabakauri, G. (2010). Dynamic mean-variance asset allocation, *Review of Financial Studies* **23**: 2970–3016.

Björk, T. (1998). *Arbitrage Theory in Continuous Time*, Oxford University Press.

Björk, T. & Murgoci, A. (2010). A general theory of markovian time inconsistent stochastic control problems, *Preprint, SSRN: http://ssrn.com/abstract=1694759*.

Björk, T., Murgoci, A. & Zhou, X. (2014). Mean variance portfolio optimization with state dependent risk aversion, *Mathematical Finance* **24**: 1–24.

Czichowsky, C. (2013). Time consistent mean variance portfolio selection in discrete and continuous time, *Finance and Stochastics* **17**: 227–271.

Ekeland, I., Mbodji, O. & Pirvu, T. (2012). Time-consistent portfolio management, *SIAM Journal on Financial Mathematics* **3**: 1–32.

Ekeland, I. & Pirvu, T. (2008). Investment and consumption without commitment, *Mathematics and Financial Economics* **2**: 57–86.

Karnam, C., Ma, J. & Zhang, J. (2016). Dynamic approaches for some time inconsistent problems, *arXiv:1604.03913 [math.OC]*.

Li, D. & Ng, W.-L. (2000). Optimal dynamic portfolio selection: multiperiod mean-variance formulation, *Mathematical Finance* **10(3)**: 387–406.

Pedersen, J. & Peskir, G. (2016). Optimal mean-variance portfolio selection, *to appear in Mathematics and Financial Economics* .

Pollak, R. A. (1968). Consistent planning, *The Review of Economic Studies* **35**: 201–208.

Richardson, H. (1989). A minimum variance result in continuous trading portfolio optimization, *Management Science* **35**: 1045–1055.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization, *The Review of Economic Studies* **23**: 165–180.

Vigna, E. (2014). On efficiency of mean variance based portfolio selection in defined contribution pension schemes, *Quantitative Finance* **14**: 237–258.

Zhou, X. & Li, D. (2000). Continuous-time mean-variance portfolio selection: A stochastic LQ framework, *Applied Mathematics and Optimization* **42**: 19–33.