A Theory of Markovian Time Inconsistent Stochastic Control in Discrete Time *

Tomas Björk
Department of Finance,
Stockholm School of Economics
tomas.bjork@hhs.se

Agatha Murgoci
Department of Economics
Aarhus University
agatha.murgoci@econ.au.dk

Published in

Finance and Stochastics **18**, No 3, 545–592, (2013)

Abstract

We develop a theory for a general class of discrete time stochastic control problems which, in various ways, are time inconsistent in the sense that they do not admit a Bellman optimality principle. We attach these problems by viewing them within a game theoretic framework, and we look for subgame perfect Nash equilibrium points. For a general controlled Markov process and a fairly general objective functional we derive an extension of the standard Bellman equation, in the form of a system of non-linear equations, for the determination for the equilibrium strategy as well as the equilibrium value function. Most known examples of time inconsistent stochastic control problems in the literature are easily seen to be special cases of the present theory. We also prove that for every time inconsistent problem, there exists an associated time consistent problem such that the optimal control and the optimal value function for the consistent problem coincides with the equilibrium control and value function respectively for the time inconsistent problem. To exemplify the theory we study some concrete examples, such as hyperbolic discounting and mean variance control.

Keywords: Time consistency, time inconsistency, time inconsistent control, dynamic programming, stochastic control, Bellman equation,hyperbolic discounting,mean-variance
AMS code: 49L, 60J, 91A, 91G
JEL Code C61, C72, C73, D5, G11, G12

*The authors are greatly indebted to the associate editor, two anonymous referees, Ivar Ekeland, Ali Lazrak, Martin Schweizer, Traian Pirvu, Suleyman Basak, Mogens Steffensen, and Eric Böse-Wolf for very helpful comments.

Contents

1 Introduction

In a standard discrete time time stochastic optimal control problem the object is that of maximizing (or minimizing) a functional of the form

$$E\left[\sum_{n=0}^{T} C\left(n, X_n, u_n\right) + F(X_T)\right].$$

where X is some controlled Markov process, u_n is the control applied at time n, and F, C are given real valued functions. A typical example is when X is a controlled scalar stochastic equation of the form

$$X_{n+1} = \mu(X_n, u_n, Y_{n+1}),$$

where Y is the stochastic noise process, and we have some initial condition $X_0 = x_0$. Later on in the paper we will allow for more general dynamics than those of a difference equation, but in this informal section we restrict ourselves to this case and for simplicity we assume that there are no constraints on the scalar control u_n.

The object of the present paper is to study problems which are similar to the one stated above, but where there is also an element of "time inconsistency". In order to understand exactly why and how our problems are different from the standard one above, and what the term "time inconsistency" really means, we need to recapitulate, very briefly, the main ideas of dynamic programming.

1.1 Dynamic programming and time consistency

A standard way of attacking a problem like the one above is by using Dynamic Programming (henceforth DynP), so we now give a brief recapitulation of the main ideas. We restrict ourselves to **control laws**, i.e., the control at time k, given that $X_k = y$, is of the form $\mathbf{u}(k, y)$ where the control law $\mathbf{u}$ is a deterministic function of the variables (k, y). We then embed the problem above in a family of problems indexed by the initial point. More precisely we consider, for every (n, x), the problem $\mathcal{P}_{n,x}$ of maximizing the reward functional

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{T} C\left(k, X_k, u_k\right) + F(X_T)\right].$$

given the initial condition $X_n = x$. Denoting the optimal control law for $\mathcal{P}_{n,x}$ by $\hat{\mathbf{u}}_{nx}(k,y)$ (where $n \le k \le T-1$) and the corresponding optimal value function by $V_n(x)$ we see that the original problem corresponds to the problem $\mathcal{P}_{0,x_0}$.

We note that *ex ante* the optimal control law $\hat{\mathbf{u}}_{nx}(k,y)$ for the problem $\mathcal{P}_{n,x}$ must be indexed by the initial point (n,x) but, as is well known, problems of the kind described above turn out to be **time consistent** in the sense that we have the **Bellman optimality principle**, which roughly says that the optimal control is independent of the initial point. More precisely: if a control law is optimal on the time interval $\{n,...,T\}$, then it is also optimal for any subinterval $\{m,...,T\}$ where $n \le m$, or more formally

$$\hat{\mathbf{u}}_{nx}(k,y) = \hat{\mathbf{u}}_{mz}(k,y),$$

for all states x, y, z and for all times $n \le m \le k$.

Given the Bellman principle, it is easy to derive the Bellman equation

$$\begin{aligned} V_n(x) &= \sup_{u \in R} \left\{ C(n,x,u) + E_{n,x}\left[V_{n+1}(X^u_{n+1}) \right] \right\}, \\ V_T(x) &= F(x), \end{aligned}$$

for the determination of V.

We end this section by listing some important conditions concerning time consistency, and in the next section we will see some, seemingly quite natural, problems where these conditions do not hold, thus giving rise to time inconsistency.

Remark 1.1 *The main reasons for the time consistency of the indexed family* $\{\mathcal{P}_{n,x} :\ x \in R, n = 0,1,2,\ldots\}$ *of problems above are as follows.*

- *The term* $C(k, X_k, u_k)$ *in the problem* $\mathcal{P}_{n,x}$ *is allowed to depend on* k*,* X_k *and* u_k*. It is* **not** *allowed to depend on the initial point* (n,x)*.*
- *The terminal evaluation term is allowed to be of the form* $E_{n,x}[F(X_T)]$*, i.e the expected value of a non-linear function of the terminal value* X_T*. We are* **not** *allowed to have a term of the form* $G(E_{n,x}[X_T])$*, which is a non-linear function of the expected value.*
- *We are* **not** *allowed to let the terminal evaluation function* F *depend on the initial point* (n,x)*.*

1.2 Three disturbing examples

We will now consider three seemingly simple examples from financial economics, where time consistency fail to hold. In all these cases we consider a financial market with a risky asset as well as a risk free asset with rate of return r. We denote by X the market value of a self financing portfolio, and by c the consumption process. We now consider three indexed families of optimization problems. In all cases the (naive) objective is to maximize the objective functional $J_n(x, \mathbf{u})$, where (n,x) is the initial point and $\mathbf{u}$ a shorthand expression for the control strategy, consisting of consumption and portfolio weights.

1. **Non-exponential discounting**

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{T-1} \varphi(k-n)h(c_k) + \varphi(T-n)F(X_T)\right]$$

In this problem h is the local utility of consumption, F is the utility of terminal wealth, and φ is the **discounting function**. This problem differs from a standard problem by the fact that the initial point in time n enters in the discounting function (see Remark 1.1). Obviously; if φ is a power function so $\varphi(k-n) = \delta^{k-n}$ then we can factor out δ^{-n} and convert the problem into a standard problem with objective functional

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{T-1} \delta^k h(c_k) + \delta^T F(X_T)\right]$$

One can show, however, that **every** choice of the discounting function φ, apart from the the power case, will lead to a time inconsistent problem. More precisely, the Bellman optimality principle will not hold.

2. **Mean variance utility**

$$J_n(x, \mathbf{u}) = E_{n,x}\left[X_T\right] - \frac{\gamma}{2} Var_{n,x}\left(X_T\right)$$

This case is a dynamic version of a standard Markowitz investment problem where we want to maximize utility of final wealth. The utility of final wealth is basically linear in wealth, as given by the term $E_{n,x}\left[X_T\right]$, but we penalize the risk by the conditional variance $\frac{\gamma}{2} Var_{n,x}\left(X_T\right)$. This looks innocent enough, but we recall the elementary formula

$$Var[X] = E[X^2] - \left(E[X]\right)^2.$$

Now, in a standard time consistent problem we are allowed to have terms like $E_{n,x}\left[F(X_T)\right]$ in the objective functional, i.e. we are allowed to have the expected value of a non-linear function of terminal wealth. In the present case, however we have the term $\left(E_{n,x}\left[X\right]\right)^2$. This is not an expected value of a non-linear function of terminal wealth, but instead a non-linear function of the expected value of terminal wealth, and we thus have a time inconsistent problem (see Remark 1.1).

3. **Endogenous habit formation**

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\ln\left(X_T - x + \beta\right)\right], \quad \beta > 0.$$

In this particular example we basically want to maximize log utility of terminal wealth. In a standard problem we would have the objective $E_{n,x}\left[\ln\left(X_T - d\right)\right]$ where $d > 0$ is the lowest acceptable level of terminal wealth. In our problem, however, the lowest acceptable level of terminal wealth is given by $x - \beta$ and it thus depends on your wealth $X_n = x$ at time n. This again leads to a time inconsistent problem. (We remark in passing that there are other examples of endogenous habit formation which are indeed time consistent.)

1.3 Approaches to handle time inconsistency

In all the three examples of the previous subsection we are faced with a time inconsistent family of problems, in the sense that if for some fixed initial point (n, x) we determine the control law $\hat{\mathbf{u}}$ which maximizes $J_n(x, \mathbf{u})$, then at some later point (k, X_k) the control law $\hat{\mathbf{u}}$ (restricted to the interval $[k, T]$) will no longer be optimal for the functional $J_k(X_k, \mathbf{u})$. It is thus conceptually unclear what we mean by "optimality" and even more unclear what we mean by "an optimal control law", so our first task is to specify more precisely exactly which problem we are trying to solve. There are then at least three different ways of handling a family of time inconsistent problems, like the ones above

- We dismiss the entire problem as being silly.
- We fix **one** initial point, like for example $(0, x_0)$, and then try to find the control law $\hat{\mathbf{u}}$ which maximizes $J_0(x_0, \mathbf{u})$. We then simply disregard the fact that at a later points in time such as (n, X_n) the control law $\hat{\mathbf{u}}$ will not be optimal for the functional $J_n(X_n, \mathbf{u})$. In the economics literature, this is known as **pre-commitment**.
- We take the time inconsistency seriously and formulate the problem in game theoretic terms.

All of the three strategies above may in different situations be perfectly reasonable, but in the present paper we choose the last one. The basic idea is then that when we decide on a control action at time t we should explicitly take into account that at future times we will have a different objective functional or, in more loose terms, "our tastes are changing over time". We can then view the entire problem as a non-cooperative game, with one player for each time n, where player n can be viewed as the future incarnation of ourselves (or rather of our preferences) at time n. Player n chooses the control law $\mathbf{u}(n, \cdot)$ so, given this point of view, it is natural to look for Nash equilibria for the game, and this is exactly our approach. For the case of a finite time horizon, the approach works roughly as follows. See Section 2 for precise definitions.

1. Given that $X_{T-1} = x$, player $T-1$ has as standard optimization problem to solve, namely that of maximizing
$$J_n(x, u_{T-1})$$
over u_{T-1}. We denote the optimal u by $\hat{u}_{T-1}(x)$.
2. Given that $X_{T-2} = x$, and that player $T-1$ is using $\hat{u}_{T-1}$, player $T-2$ now maximizes
$$J_n(x, u_{T-2}, \hat{u}_{T-1})$$
over u_{T-2}. We denote the optimal u by $\hat{u}_{T-2}(x)$.
3. We then proceed by induction.

1.4 Previous literature

The game theoretic approach to time inconsistency using Nash equilibrium points as above has a long history starting with [16] where a deterministic Ramsay problem with non-exponential discounting is studied. Further work along this line in continuous and discrete time is provided in [1], [9], [11], [14], [15], and [17].

Recently there has been renewed interest in these problems in continuous time. In the interesting, and mathematically very advanced, papers [7] and [8], the authors consider optimal consumption and investment under hyperbolic discounting (Problem 1 in our list above) in deterministic and stochastic models from the above game theoretic point of view.

In [2] the authors undertake a deep study of the mean variance problem within a Wiener driven framework. This is basically Problem 2 in the list above, but the authors also consider the case of multiple assets, as well as the case of a hidden Markov process driving the parameters of the asset price dynamics. Their methodology is based on using an "iterated variance formula". This fits the mean variance framework very well, but it is hard to see how to extend the methodology to more complicated objective functionals.

In [6] the author develops a very complete and impressive theory of the mean variance problem within in a non Markovian general semi martingale framework, thus extending the results of [2] considerably. The technique in [6] is different from that in [2], but it is closely related to the case of a mean variance problem, and it is not clear that it can be extended to other objective functionals.

In all the cited papers above, the various authors have studied particular models and/or objective functionals, each author deriving the relevant equilibrium conditions for his or her model. What has been lacking in the literature so far, is a (reasonably) **general** theory of time inconsistent stochastic control, and the purpose of the present paper is precisely to present such a theory.

To our knowledge, the present paper, which is the discrete time part of our working paper [3], is the first attempt to derive a general (albeit Markovian) theory of time inconsistent control. The corresponding continuous time theory (which is technically more complicated) depends heavily on the discrete time results of the present paper and can be found in the working paper [3]. It will appear separately as [4].

In the working paper [12] the authors use the theory of [3] to study several interesting new applications.

1.5 Contributions of the present paper

The object of the present paper is to undertake a rigorous study of time inconsistent control problems in a reasonably general Markovian framework, and in particular we do not want to tie ourselves down to a particular applied problem. We have therefore chosen a setup of the following form.

- We consider a **general controlled Markov process** X, living on some suitable space (details are given below). It is important to notice that we do not make any structural assumptions whatsoever about X, and we note that the setup obviously includes the case when X is determined by a system of stochastic difference equations.

- We consider a general reward functional of the form

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{T-1} C_{n,k}\left(x, X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}})\right) + F_n(x, X_T^{\mathbf{u}})\right] + G_n\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right),$$

where we also allow the case $T = \infty$ (see Section 5.1).

Referring to the discussion in Remark 1.1 we see that with the choice of functional above, time inconsistency will enter at several points:

- The shape of the utility functional depends explicitly on the initial position (n, x) in time-space, as can be seen in the appearance of n and x in the expression $F_n(x, X_T)$ and similarly for the other terms. In other words, as the X process moves around, our utility function changes, so at time k this part of the utility function will have the form $F_k(X_k, X_T)$.
- For a standard time consistent control problem we are allowed to have expressions like $E_{n,x}[G(X_T)]$ in the utility function, i.e. we are allowed to have the expected value of a non linear function G of the future process value. Time consistency is then a relatively simple consequence of the law of iterated expectations. In our problem above, however, we have an expression of the form $G_n\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right)$ which, even apart from the appearance of n and x in the function G, is not the expectation of a non linear function, but a nonlinear function of the expected value. We thus do not have access to iterated expectations, so the problem becomes time inconsistent. On top of this we also have the appearance of n and x in the expression $G_n\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right)$.

This setup is studied in some detail and our main results are as follows.

- We derive an extension of the standard Bellman equation to a non standard system of equations for the determination of the equilibrium value function V and the equilibrium control $\hat{\mathbf{u}}$.
- We prove that to every time inconsistent problem of the form above, there exists an associated **standard, time consistent**, control problem with the following properties:
  - The optimal value function for the standard problem coincides with the equilibrium value function for the time inconsistent problem.
  - The optimal control law for the standard problem coincides with the equilibrium strategy for the time inconsistent problem.

For the case of a Ramsay problem with non exponential discounting, a related equivalence result can be found in [1] but our result is more general and also structurally different from that of [1].

- We solve some specific test examples. In particular we study non-exponential discounting in some detail, and we also study mean variance optimal portfolios.

We thus extend the existing literature substantially by allowing for a considerably more general utility functional, and a completely general Markovian structure.

1.6 Structure of the paper

We develop the general discrete time theory in Section 2 and the main result is Theorem 3.2. In Section 4 we prove that for each time inconsistent problem there exists an equivalent standard time consistent problem, such that the optimal control for the standard problem coincides with the equilibrium control for the time inconsistent problem. We discuss existence and uniqueness questions in Section 6. In Section 7 we exemplify the general theory by studying the special case of non exponential discounting in some detail, and in Section 8 we specialize further to quasi-hyperbolic discounting. Section 9.1 is devoted to mean variance portfolio analysis, and in Section 10 we conclude and give directions for future research.

2 General theory I: Setup

In this section we present the setup and in the next section we derive the main theoretical results.

2.1 Setup

We consider a given **controlled Markov process** X, evolving on a measurable state space $\{\mathcal{X}, \mathcal{G}_X\}$, with controls taking values in a measurable control space $\{\mathcal{U}, \mathcal{G}_U\}$. The action is in discrete time, indexed by the set $\mathbf{N}$ of natural numbers. The intuitive idea is that if $X_n = x$, then we can choose a control $u_n \in \mathcal{U}$, and this control will affect the transition probabilities from X_n to X_{n+1}. This idea is formalized by specifying a family of **transition probabilities**,

$$\{p_n^u(dz; x) : n \in \mathbf{N}, x \in \mathcal{X}, u \in \mathcal{U}\}.$$

For every fixed $n \in \mathbf{N}$, $x \in \mathcal{X}$ and $u \in \mathcal{U}$, we assume that $p_n^u(\cdot; x)$ is a probability measure on $\mathcal{X}$, and for each $A \in \mathcal{G}_X$, the probability $p_n^u(A; x)$ is jointly measurable in (x, u). The interpretation of this is that $p_n^u(dz; x)$ is the probability distribution of X_{n+1}, given that $X_n = x$, and that we at time n apply the control u, i.e.,

$$p_n^u(dz; x) = P\left(X_{n+1} \in dz \,|X_n = x, u_n = u\right).$$

To obtain a Markov structure, we restrict the controls to be **feedback control laws**, i.e. at time n, the control u_n is allowed to depend on time n and state X_n. We can thus write

$$u_n = \mathbf{u}_n(X_n),$$

where the mapping $\mathbf{u} : \mathbf{N} \times \mathcal{X} \to \mathcal{U}$ is measurable. Note the boldface notation for the mapping $\mathbf{u}$. In order to distinguish between functions and function values, we will always denote a control law (i.e. a mapping) by using boldface, like $\mathbf{u}_n$, whereas a possible value of the mapping will be denoted without boldface, like, $u \in \mathcal{U}$.

Remark 2.1 *It is natural to ask whether our analysis can be extended to the class of adapted control strategies, rather than the more restricted class of feedback laws considered in the present paper. For Markovian standard (i.e. time consistent) stochastic control problems, it is well known that it is sufficient to consider the class of feedback laws, in the sense that it can be proved that the optimal adapted policy is in fact of feedback form. It would thus be natural to investigate whether we would have the corresponding result also for our class of time inconsistent problems. One would for example hope to be able to prove that all adapted equilibrium controls are in fact also feedback laws. We have tried to study also this question but it turns out to be technically (and notationally) quite complicated, so it has to be postponed to a separate paper.*

Given the family of transition probabilities we may define a corresponding family of operators, operating on function sequences.

Definition 2.1 *A* **function sequence** *is a mapping $f : \mathbf{N} \times \mathcal{X} \to R$, where we use the notation $(n, x) \longmapsto f_n(x)$.*

- *For each $u \in \mathcal{U}$, the operator $\mathbf{P}^u$, acting on the set of integrable function sequences, is defined by*

$$\left(\mathbf{P}^u f\right)_n(x) = \int_{\mathcal{X}} f_{n+1}(z) p_n^u(dz, x).$$

  The corresponding discrete time "infinitesimal operator" $\mathbf{A}^u$ is defined by

$$\mathbf{A}^u = \mathbf{P}^u - \mathbf{I},$$

  where $\mathbf{I}$ is the identity operator.

- *For each control law $\mathbf{u}$ the operator $\mathbf{P}^{\mathbf{u}}$ is defined by*

$$\left(\mathbf{P}^{\mathbf{u}} f\right)_n(x) = \int_{\mathcal{X}} f_{n+1}(z) p_n^{\mathbf{u}_n(x)}(dz, x),$$

  and $\mathbf{A}^{\mathbf{u}}$ is defined correspondingly as

$$\mathbf{A}^{\mathbf{u}} = \mathbf{P}^{\mathbf{u}} - \mathbf{I},$$

In more probabilistic terms we have the interpretation.

$$(\mathbf{P}^u f)_n(x) = E\left[f_{n+1}(X_{n+1}) \middle| X_n = x, u_n = u \right],$$

or, as we often will write,

$$(\mathbf{P}^u f)_n(x) = E_{n,x}\left[f_{n+1}(X^u_{n+1}) \right],$$

and $\mathbf{A}^u$ is the discrete time version of the continuous time infinitesimal operator. We immediately have the following result.

Proposition 2.1 *Consider a real valued, function sequence $\{f_n(x)\}$, and a control law $\mathbf{u}$. The process $f_n(X^{\mathbf{u}}_n)$ is then a martingale under the measure induced by $\mathbf{u}$ if and only if the following conditions are satisfied*

- *The process $f_n(X^{\mathbf{u}}_n)$ is integrable.*
- *The sequence $\{f_n\}$ satisfies the equation*

$$(\mathbf{A}^{\mathbf{u}} f)_n(x) = 0, \quad n = 0, 1, \ldots, T-1.$$

Proof. Obvious from the definition of $\mathbf{A}^{\mathbf{u}}$. ∎

It is clear that for a fixed initial point (n, x) and a fixed control law $\mathbf{u}$ we may in the obvious way define a Markov process denoted by $X^{n,x,\mathbf{u}}$, where for notational simplicity we often drop the upper index n, x and use the notation $X^{\mathbf{u}}$. The corresponding expectation operator is denoted by $E^{\mathbf{u}}_{n,x}[\cdot]$, and we often drop the upper index $\mathbf{u}$, and instead use the notation $E_{n,x}[\cdot]$. A typical example of an expectation will thus have the form $E_{n,x}[F(X^{\mathbf{u}}_k)]$ for some real valued function F and some point in time k.

2.2 Basic problem formulation

For a fixed $(n, x) \in \mathbf{N} \times \mathcal{X}$, a fixed control law $\mathbf{u} = \{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{T-1}\}$, and a fixed time horizon T, we consider a functional of the basic form

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{T-1} C\left(x, X^{\mathbf{u}}_k, \mathbf{u}_k(X^{\mathbf{u}}_k)\right) + F(x, X^{\mathbf{u}}_T) \right] + G\left(x, E_{n,x}\left[X^{\mathbf{u}}_T\right]\right), \quad (1)$$

Later on we will in fact allow an even more general functional, but for the present purposes, the form above is general enough.

Obviously, the functional J does not depend on the entire control law $\mathbf{u} = \{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{T-1}\}$ but only on the restriction of $\mathbf{u}$ to the interval $[n, T]$, i.e. on $\{\mathbf{u}_n, \mathbf{u}_{n+1}, \ldots, \mathbf{u}_{T-1}\}$.

The intuitive idea is that we are standing at (n, x) and that we would like to choose a control law $\mathbf{u}$ which maximizes J. We can thus define an indexed family of problems $\{\mathcal{P}_{n,x}\}$ by

$$\mathcal{P}_{n,x}: \qquad \max_{\mathbf{u}} J_n(x, \mathbf{u}),$$

where *max* is shorthand for the imperative "maximize!".

Remark 2.2

- *For simplicity we assume that there are no constraints on the control, so at time n we are allowed to choose any $u_n \in \mathcal{U}$. State and time dependent control constraints can, however, easily be incorporated. See Section 3.6.*

- *We can easily extend the theory to the case when the term $G\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right)$ is replaced by $G\left(x, E_{n,x}\left[h(X_T^{\mathbf{u}})\right]\right)$ for some real valued function h. See Section 3.7.*

As we have seen in Remark 1.1 above, the complicating factor with our indexed family of problems is that the family $\{\mathcal{P}_{n,x}\}$ is time inconsistent in the sense that if $\hat{\mathbf{u}}$ is optimal for $\mathcal{P}_{n,x}$, then the restriction of $\hat{\mathbf{u}}$ to the time set $k, k+1, \dots, T$ (for $k > n$) is not necessarily optimal for the problem $\mathcal{P}_{k,X_k^{\mathbf{u}}}$. Thus, if we at some point (n, x) decide on a feedback law $\hat{\mathbf{u}}$ which is optimal from the point of view of (n, x) then as time goes by, we will no longer consider $\hat{\mathbf{u}}$ to be optimal. To handle this problem we will use a game theoretic approach and we now go on to describe this in some detail.

2.3 The game theoretic formulation

The idea, which appears already in [16], is to view the setup above in game theoretic terms. More precisely we view it as a non-cooperative game where we have one player at each point n in time. We refer to this player as "player number n" and the rule is that player number n can only choose the control u_n, or more precisely the control law $\mathbf{u}_n(\cdot)$. One interpretation is that these players are different future incarnations of yourself (or rather incarnations of your future preferences), but conceptually it is perhaps easier to think of it as one separate player at each n.

Given the data (n, x), player number n would, in principle, like to maximize $J_n(x, \mathbf{u})$ over the class of feedback controls $\mathbf{u}$ restricted to $[n, T]$, i.e. he would like to maximize $J_n(x, \mathbf{u})$ over $\{\mathbf{u}_n, \mathbf{u}_{n+1}, \dots, \mathbf{u}_{T-1}\}$, but since he can only choose the control $\mathbf{u}_n$, this is not possible. Instead of looking for "optimal" feedback laws, we take the game theoretic point of view and study so called subgame perfect Nash equilibrium strategies. The formal definition is as follows.

Definition 2.2 *We consider a fixed control law $\hat{\mathbf{u}}$ and make the following construction.*

1. *Fix an arbitrary point (n, x) where $n < T$, and choose an arbitrary control value $u \in \mathcal{U}$.*

2. *Now define the control law $\mathbf{u}^{u,n}$ on the time set $n, n+1, \dots, T-1$ by setting, for any $y \in \mathcal{X}$,*

$$\mathbf{u}_k^{u,n}(y) = \begin{cases} \hat{\mathbf{u}}_k(y), & \textit{for} \quad k = n+1, \dots, T-1, \\ u, & \textit{for} \quad k = n. \end{cases}$$

We say that $\hat{\mathbf{u}}$ *is a subgame perfect Nash* **equilibrium** *strategy if, for every fixed* (n, x)*, the following condition hold*

$$\sup_{u \in \mathcal{U}} J_n(x, \mathbf{u}^{u,n}) = J_n(x, \hat{\mathbf{u}}_n).$$

If an equlibrium control $\hat{\mathbf{u}}$ *exists, we define the* **equilibrium value function** V *by*

$$V_n(x) = J_n(x, \hat{\mathbf{u}}).$$

In more pedestrian terms this means that if player number n knows that player number k will choose the control $\hat{\mathbf{u}}_k$ for all $k > n$, then it is optimal for player number n to choose $\hat{\mathbf{u}}_n$.

Remark 2.3 *An equivalent, and perhaps more concrete, way of describing an equilibrium strategy is as follows.*

- *The equilibrium control* $\hat{\mathbf{u}}_{T-1}(x)$ *is obtained by letting player* $T-1$ *optimize* $J_{T-1}(x, \mathbf{u})$ *over* u_{T-1} *for all* $x \in \mathcal{X}$*. This is a standard optimization problem without any game theoretic components.*
- *The equilibrium control* $\hat{\mathbf{u}}_{T-2}$ *is obtained by letting player* $T-2$ *choose* u_{T-2} *to optimize* J_{T-2}*, given the knowledge that player number* $T-1$ *will use* $\hat{\mathbf{u}}_{T-1}$*.*
- *Proceed recursively by backward induction.*

We thus see that, in discrete time and for a finite horizon, the equilibrium control is determined by backward induction. Note, however, that for the discrete time infinite horizon case, as well as for the continuous time case, the situation is much more complicated.

Obviously; for a standard time consistent control problem, the game theoretic aspect becomes trivial and the equilibrium control law coincides with the standard (time consistent) optimal law. The equilibrium value function V will coincide with the optimal value function and, using dynamic programming arguments, V is seen to satisfy a standard Bellman equation.

The main result of the present paper is that in the time inconsistent case, the equilibrium value function V will satisfy a system of non linear equations. This system of equations extend the standard Bellman equation, and for a time consistent problem they reduce to the Bellman equation.

3 The extended Bellman equation

In this section we assume that there exists an equilibrium control law $\hat{\mathbf{u}}$ (which may not be unique) and we consider the corresponding equilibrium value function V defined above. The goal of this section is to derive an system of equations, extending the standard Bellman equation, for the determination of V. This will be done in the following two steps:

- For an arbitrarily chosen control law $\mathbf{u}$, we will derive a recursive equation for $J_n(x, \mathbf{u})$.
- We will then fix (n, x) and consider two control laws. The first one is the equilibrium law $\hat{\mathbf{u}}$, and the other one is the law $\mathbf{u}$ where we choose $u = \mathbf{u}_n(x)$ arbitrarily, but follow the law $\hat{\mathbf{u}}$ for all k with $k = n+1, \dots T-1$. The trivial observation that

$$\sup_{u \in \mathcal{U}} J_n(x, \mathbf{u}) = J_n(x, \hat{\mathbf{u}}) = V_n(x),$$

will finally give us the extension of the Bellman equation.

The reader with experience from dynamic programming (DynP) will recoginize that the general program above is in fact more or less the same as for standard (time consistent) DynP. However; in the present time inconsistent setting, the derivation of the recursion in the first step is much more tricky than in the corresponding step from DynP, and it also requires some completely new constructions.

3.1 Simplifying the problem

In order to derive the recursion for $J_n(x, \mathbf{u})$ we consider an arbitrary initial point (n, x), and we consider an arbitrarily chosen control law $\mathbf{u}$. The value taken by $\mathbf{u}$ at (n, x) will play a special role in the sequel, and for ease of reading we will use the notation $\mathbf{u}_n(x) = u$.

We now go on to derive a recursion between J_n and J_{n+1}. This is conceptually rather delicate, and sometimes a bit messy. In order to increase readability we therefore carry out a detailed derivation only for the case when the objective functional has the simpler form

$$J_n(x, \mathbf{u}) = E_{n,x}\left[F(x, X_T^{\mathbf{u}})\right] + G\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right). \tag{2}$$

We then provide the result for the general case in Section 3.5. The derivation of this is completely parallel to that of the simplified case.

Since also the derivation for the case (2) is rather intricate we will in fact simplify even further. For pedagogical purposes we will thus consider two special case, namely the case

$$J_n(x, \mathbf{u}) = E_{n,x}\left[F(x, X_T^{\mathbf{u}})\right]. \tag{3}$$

and the case

$$J_n(x, \mathbf{u}) = G\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right). \tag{4}$$

The point is that, by considering these special cases, it is easy to see how we separately handle the two main sources of time inconsistency in our model:

- The occurrence of the present state x in the expression $F(x, X_T^{\mathbf{u}})$, in the (otherwise standard) objective functional $E_{n,x}\left[F(x, X_T^{\mathbf{u}})\right]$.

- The occurrence of the nonstandard term $G\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right)$.

Having understood how to handle these special cases, the extension to the case (2) is very easy. The treatments of the two special cases can be read independently of each other, so the reader reader can (depending on interest) study both of them or any one of them. The reader who wants to be *in medias res* can skip the derivations and go directly to Sections 3.4 and 3.5.

Before going on to these special cases we make a remark on notation. Given an initial point (n, x), the random variable X_{n+1} will only depend on x and on the control value $\mathbf{u}_n(x) = u$ motivating the notation X_{n+1}^u. The distribution of X_k for $k > n+1$ will, on the other hand depend on the entire control law $\mathbf{u}$ (restricted to the interval $[n, k-1]$) so for $k > n+1$ we use the notation $X_k^{\mathbf{u}}$.

3.2 The case $J_n(x, \mathbf{u}) = E_{n,x}\left[F(x, X_T^{\mathbf{u}})\right]$

From the definition of J we have

$$J_{n+1}(X_{n+1}^u, \mathbf{u}) = E_{n+1}\left[F(X_{n+1}^u, X_T^{\mathbf{u}})\right], \tag{5}$$

where for simplicity of notation we write $E_{n+1}[\cdot]$ instead of $E_{n+1, X_{n+1}^u}[\cdot]$. We now make the following definition which will play a central role in the sequel.

Definition 3.1 *For any control law $\mathbf{u}$, we define the function sequence $\{f_n^{\mathbf{u}}\}$, where $f_n^{\mathbf{u}} : \mathcal{X} \times \mathcal{X} \to \mathbf{R}$ by.*

$$f_n^{\mathbf{u}}(x, y) = E_{n,x}\left[F\left(y, X_T^{\mathbf{u}}\right)\right].$$

We also introduce the notation

$$f_n^{\mathbf{u},y}(x) = f_n^{\mathbf{u}}(x, y).$$

The difference between $f_n^{\mathbf{u},y}$ and $f_n^{\mathbf{u}}$, is that we view $f_n^{\mathbf{u}}$ as a function of the two variables x and y, whereas $f_n^{\mathbf{u},y}$ is, for a fixed y, viewed as a function of the single variable x.

From the definitions above it is obvious that, for any fixed y, the process $f_n^{\mathbf{u},y}(X_n^{\mathbf{u}})$ is a martingale under the measure generated by $\mathbf{u}$. We thus have the following result.

Lemma 3.1 *For every fixed control law $\mathbf{u}$ and every fixed choice of $y \in \mathcal{X}$, the function sequence $\{f_n^{\mathbf{u},y}\}$ satisfies the recursion*

$$\begin{aligned} \left(\mathbf{A}^{\mathbf{u}} f^{\mathbf{u},y}\right)_n(x) &= 0, \quad n = 0, 1, \ldots, T-1. \\ f_T^{\mathbf{u},y}(x) &= F(y, x). \end{aligned}$$

We now go on to derive the recursion for $J_n(x, \mathbf{u})$.

3.2.1 The recursion for $J_n(x, \mathbf{u})$

Going back to (5) we note that, from the Markovian structure and Definition 3.1, we have

$$E_{n+1}\left[F(X_{n+1}^u, X_T^{\mathbf{u}})\right] = f_{n+1}^{\mathbf{u}}(X_{n+1}^u, X_{n+1}^u).$$

We can now write (5) as

$$J_{n+1}(X_{n+1}^u, \mathbf{u}) = f_{n+1}^{\mathbf{u}}(X_{n+1}^u, X_{n+1}^u).$$

Taking expectations gives us

$$E_{n,x}\left[J_{n+1}(X_{n+1}^u, \mathbf{u})\right] = E_{n,x}\left[f_{n+1}^{\mathbf{u}}(X_{n+1}^u, X_{n+1}^u)\right],$$

and, going back to the definition of $J_n(x, \mathbf{u})$, we can write this as

$$\begin{aligned} E_{n,x}\left[J_{n+1}(X_{n+1}^u, \mathbf{u})\right] &= J_n(x, \mathbf{u}) \\ &+ E_{n,x}\left[f_{n+1}^{\mathbf{u}}(X_{n+1}^u, X_{n+1}^u)\right] - E_{n,x}\left[F(x, X_T^{\mathbf{u}})\right]. \end{aligned}$$

At this point it may seem natural to use the identity $E_{n,x}\left[F(x, X_T^{\mathbf{u}})\right] = f_n^{\mathbf{u}}(x, x)$, but for various reasons this is not a good idea. [1]

Instead we note that

$$E_{n,x}\left[F(x, X_T^{\mathbf{u}})\right] = E_{n,x}\left[E_{n+1}\left[F(x, X_T^{\mathbf{u}})\right]\right] = E_{n,x}\left[f_{n+1}^{\mathbf{u}}(X_{n+1}^u, x)\right],$$

Substituting this into the recursion above, we can collect the findings so far.

Lemma 3.2 *The value function J satisfies the following recursion.*

$$\begin{aligned} J_n(x, \mathbf{u}) &= E_{n,x}\left[J_{n+1}(X_{n+1}^u, \mathbf{u})\right] \\ &- \left\{E_{n,x}\left[f_{n+1}^{\mathbf{u}}(X_{n+1}^u, X_{n+1}^u)\right] - E_{n,x}\left[f_{n+1}^{\mathbf{u}}(X_{n+1}^u, x)\right]\right\}. \end{aligned}$$

3.2.2 The recursion for $V_n(x)$

We will now derive the recursive equation for the equilibrium function $V_n(x)$. In order to do this we assume that there exists an equilibrium control $\hat{\mathbf{u}}$. We then fix an arbitrarily chosen initial point (n, x) and consider two strategies (control laws).

1. The first control law is simply the equilibrium law $\hat{\mathbf{u}}$.
2. The second control law $\mathbf{u}$ is slightly more complicated. We choose an arbitrary point $u \in \mathcal{U}$ and then defined the control law $\mathbf{u}$ as follows

$$\mathbf{u}_k(y) = \begin{cases} u, & \text{for } k = n, \\ \hat{\mathbf{u}}_k(y), & \text{for } k = n+1, \ldots, T-1. \end{cases}$$

[1] The main reason is that, in order to get a good recursion, we need to express the right hand side of the equation above as $E\,[]$-expectations of objects involving X_{n+1}^u.

We now compare the objective function J_n for these two control laws. Firstly, and by definition, we have

$$J_n(x, \hat{\mathbf{u}}) = V_n(x),$$

where V is the equilibrium value function defined earlier. Secondly, and also by definition, we have

$$J_n(x, \mathbf{u}) \leq J_n(x, \hat{\mathbf{u}}),$$

for all choices of $u \in \mathcal{U}$. We thus have the inequality

$$J_n(x, \mathbf{u}) \leq V_n(x),$$

for all $u \in \mathcal{U}$, with equality if $u = \hat{\mathbf{u}}_n(x)$. We thus have the basic relation

$$\sup_{u \in \mathcal{U}} J_n(x, \mathbf{u}) = V_n(x). \tag{6}$$

We now make a small variation of Definition 5.

Definition 3.2 *We define the function sequence $\{f_n\}_{n=0}^T$ where $f_n : \mathcal{X} \times \mathcal{X} \to$ $\mathbf{R}$, by*

$$f_n(x, y) = E_{n,x}\left[F\left(y, X_T^{\hat{\mathbf{u}}}\right)\right].$$

We also introduce the notation

$$f_n^y(x) = f_n(x, y),$$

where we view f_n^y as a function of x with y as a fixed parameter.

Using Lemma 3.2, the basic relation (6) now reads

$$\sup_{u \in \mathcal{U}} \Big\{ E_{n,x}\left[J_{n+1}(X_{n+1}^u, \mathbf{u})\right] - V_n(x)$$
$$- \left(E_{n,x}\left[f_{n+1}^{\mathbf{u}}(X_{n+1}^u, X_{n+1}^u)\right] - E_{n,x}\left[f_{n+1}^{\mathbf{u}}(X_{n+1}^u, x)\right]\right)\Big\} = 0.$$

We now observe that, since the control law $\mathbf{u}$ coincides with the equilibrium law $\hat{\mathbf{u}}$ on $[n+1, T-1]$, we have the following equalities

$$\begin{aligned} J_{n+1}(X_{n+1}^u, \mathbf{u}) &= V_{n+1}\left(X_{n+1}^u\right), \\ f_{n+1}^{\mathbf{u}}(X_{n+1}^u, x) &= f_{n+1}(X_{n+1}^u, x). \end{aligned}$$

We can thus write the recursion as

$$\sup_{u \in \mathcal{U}} \Big\{ E_{n,x}\left[V_{n+1}\left(X_{n+1}^u\right)\right] - V_n(x)$$
$$- \left(E_{n,x}\left[f_{n+1}(X_{n+1}^u, X_{n+1}^u)\right] - E_{n,x}\left[f_{n+1}(X_{n+1}^u, x)\right]\right)\Big\} = 0.$$

The first line in this equation can be rewritten as

$$E_{n,x}\left[V_{n+1}\left(X_{n+1}^u\right)\right] - V_n(x) = \left(\mathbf{A}^u V\right)_n(x).$$

The second line can be written as

$$
\begin{aligned}
& E_{n,x}\left[f_{n+1}(X_{n+1}^u, X_{n+1}^u)\right] - E_{n,x}\left[f_{n+1}(X_{n+1}^u, x)\right] \\
= \;& E_{n,x}\left[f_{n+1}(X_{n+1}^u, X_{n+1}^u)\right] - f_n(x,x) - \left(E_{n,x}\left[f_{n+1}(X_{n+1}^u, x) - f_n(x,x)\right]\right) \\
= \;& \left(\mathbf{A}^u f\right)_n(x,x) - \left(\mathbf{A}^u f^x\right)_n(x).
\end{aligned}
$$

To avoid misunderstandings: The first term $\left(\mathbf{A}^u f\right)_n(x,x)$, can be viewed as the operator $\mathbf{A}^u$ operating on the function sequence $\{h\}_n$ defined by $h_n(x) = f_n(x,x)$. In the second term, $\mathbf{A}^u$ is operating on the function sequence $f_n^x(\cdot)$ where the upper index x is viewed as a fixed parameter.

We can now state the main result for the case under study.

Proposition 3.1 *Consider a functional of the form*

$$J_n(x, \mathbf{u}) = E_{n,x}\left[F(x, X_T^{\mathbf{u}})\right],$$

and assume that an equilibrium control law $\hat{\mathbf{u}}$ *exists. Then, with notation as above, the equilibrium value function* V *satisfies the equation.*

$$\sup_{u \in \mathcal{U}} \left\{ \left(\mathbf{A}^u V\right)_n(x) - \left(\mathbf{A}^u f\right)_n(x,x) + \left(\mathbf{A}^u f^x\right)_n(x) \right\} = 0, \tag{7}$$

$$V_T(x) = F(x,x), \tag{8}$$

where the supremum above is realized by $u = \hat{\mathbf{u}}_n(x)$. *Furthermore, the following hold.*

1. *For every fixed* $y \in \mathcal{X}$ *the function sequence* $f_n^y(x)$ *is determined by the recursion*

$$\mathbf{A}^{\hat{\mathbf{u}}} f_n^y(x) = 0, \quad n = 0, \dots, T-1, \tag{9}$$

$$f_T^y(x) = F(y, x), \tag{10}$$

   and $f_n(x,x)$ *is given by*

$$f_n(x,x) = f_n^x(x).$$

2. *The probabilistic interpretation of* f *is, as before, given by*

$$f_n^y(x) = E_{n,x}\left[F\left(y, X_T^{\hat{\mathbf{u}}}\right)\right].$$

3.3 The case $J_n(x, \mathbf{u}) = G\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right)$

To derive a recursion for J_n we start by noting that from the definition of J we have

$$J_{n+1}(X_{n+1}^u, \mathbf{u}) = G\left(X_{n+1}^u, E_{n+1}\left[X_T^{\mathbf{u}}\right]\right), \tag{11}$$

where for simplicity of notation we write $E_{n+1}[\cdot]$ instead of $E_{n+1, X_{n+1}^u}[\cdot]$. We now make the following definition which will play a central role in the sequel.

Definition 3.3 *For any control law* $\mathbf{u}$*, we define the function sequence* $\{g_n^{\mathbf{u}}\}$*, where* $g_n^{\mathbf{u}} : \mathcal{X} \times \mathcal{X} \to \mathbf{R}$ *by.*

$$g_n^{\mathbf{u}}(x) = E_{n,x}\left[X_T^{\mathbf{u}}\right].$$

From the definition above it is obvious that $g_n^{\mathbf{u}}(X_n^{\mathbf{u}})$ is a martingale under the measure generated by $\mathbf{u}$. We thus have the following result.

Lemma 3.3 *For every fixed control law* $\mathbf{u}$ *the function sequence* $\{g_n^{\mathbf{u}}\}$ *satisfies the recursion*

$$\begin{aligned} \left(\mathbf{A}^{\mathbf{u}} g^{\mathbf{u}}\right)_n(x) &= 0, \quad n = 0, 1, \ldots, T-1. \\ g_T^{\mathbf{u}}(x) &= x. \end{aligned}$$

3.3.1 The recursion for $J_n(x, \mathbf{u})$

Going back to (11) we note that, from the Markovian structure and the definitions above, we have

$$E_{n+1}\left[X_T^{\mathbf{u}}\right] = g_{n+1}^{\mathbf{u}}(X_{n+1}^u),$$

so we can write (11) as

$$J_{n+1}(X_{n+1}^u, \mathbf{u}) = G\left(X_{n+1}^u, g_{n+1}^{\mathbf{u}}(X_{n+1}^u)\right).$$

Taking expectations gives us

$$E_{n,x}\left[J_{n+1}(X_{n+1}^u, \mathbf{u})\right] = E_{n,x}\left[G\left(X_{n+1}^u, g_{n+1}^{\mathbf{u}}(X_{n+1}^u)\right)\right],$$

and, going back to the definition of $J_n(x, \mathbf{u})$, we can write this as

$$E_{n,x}\left[J_{n+1}(X_{n+1}^u, \mathbf{u})\right] = J_n(x, \mathbf{u}) + E_{n,x}\left[G\left(X_{n+1}^u, g_{n+1}^{\mathbf{u}}(X_{n+1}^u)\right)\right] - G\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right).$$

We now note that

$$E_{n,x}\left[X_T^{\mathbf{u}}\right] = E_{n,x}\left[E_{n+1}\left[X_T^{\mathbf{u}}\right]\right] = E_{n,x}\left[g_{n+1}^{\mathbf{u}}(X_{n+1}^u)\right].$$

Substituting this identity into the recursion above, we can now collect the findings so far.

Lemma 3.4 *The value function* J *satisfies the following recursion.*

$$\begin{aligned} J_n(x, \mathbf{u}) &= E_{n,x}\left[J_{n+1}(X_{n+1}^u, \mathbf{u})\right] \\ &- \left\{E_{n,x}\left[G\left(X_{n+1}^u, g_{n+1}^{\mathbf{u}}(X_{n+1}^u)\right)\right] - G\left(x, E_{n,x}\left[g_{n+1}^{\mathbf{u}}(X_{n+1}^u)\right]\right)\right\}. \end{aligned}$$

3.3.2 The recursion for $V_n(x)$

We will now derive the fundamental equation for the determination of the equlibrium function $V_n(x)$. In order to do this we assume, as in Section 3.2.2, that there exists an equilibrium control $\hat{\mathbf{u}}$. We then fix an arbitrarily chosen initial point (n, x) and consider two strategies (control laws).

1. The first control law is simply the equilibrium law $\hat{\mathbf{u}}$.
2. The second control law $\mathbf{u}$ is slightly more complicated. We choose an arbitrary point $u \in \mathcal{U}$ and then defined the control law $\mathbf{u}$ as follows

$$\mathbf{u}_k(y) = \begin{cases} u, & \text{for } k = n, \\ \hat{\mathbf{u}}_k(y), & \text{for } k = n+1, \dots, T-1. \end{cases}$$

We now compare the objective function J_n for these two control laws. Exactly as in Section 3.2.2 we obtain the inequality

$$J_n(x, \mathbf{u}) \leq V_n(x),$$

for all $u \in \mathcal{U}$, with equality if $u = \hat{\mathbf{u}}_n(x)$. We thus have the basic relation

$$\sup_{u \in \mathcal{U}} J_n(x, \mathbf{u}) = V_n(x). \tag{12}$$

We now make a small variation of Definition 3.3.

Definition 3.4 *We define the function sequence* $\{g_n\}_{n=0}^T$, *where* $g_n : \mathcal{X} \to \mathbf{R}$ *by*

$$g_n(x) = E_{n,x}\left[X_T^{\hat{\mathbf{u}}}\right].$$

Using Lemma 3.4, the basic relation (12) now reads

$$\sup_{u \in \mathcal{U}} \Big\{ E_{n,x}\left[J_{n+1}(X_{n+1}^u, \mathbf{u})\right] - V_n(x)$$
$$- \left(E_{n,x}\left[G\left(X_{n+1}^u, g_{n+1}^{\mathbf{u}}(X_{n+1}^u)\right)\right] - G\left(x, E_{n,x}\left[g_{n+1}^{\mathbf{u}}(X_{n+1}^u)\right]\right)\right)\Big\} = 0.$$

We now observe that, since the control law $\mathbf{u}$ conicides with the equilibrium law $\hat{\mathbf{u}}$ on $[n+1, T-1]$, we have the following equalities

$$\begin{aligned} J_{n+1}(X_{n+1}^u, \mathbf{u}) &= V_{n+1}\left(X_{n+1}^u\right), \\ g_{n+1}^{\mathbf{u}}(X_{n+1}^u) &= g_{n+1}(X_{n+1}^u). \end{aligned}$$

We can thus write the recursion as

$$\sup_{u \in \mathcal{U}} \Big\{ E_{n,x}\left[V_{n+1}\left(X_{n+1}^u\right)\right] - V_n(x)$$
$$- \left(E_{n,x}\left[G\left(X_{n+1}^u, g_{n+1}(X_{n+1}^u)\right)\right] - G\left(x, E_{n,x}\left[g_{n+1}(X_{n+1}^u)\right]\right)\right)\Big\} = 0.$$

The first line in this equation can be rewritten as

$$E_{n,x}\left[V_{n+1}\left(X^u_{n+1}\right)\right] - V_n(x) = \left(\mathbf{A}^u V\right)_n(x).$$

We rewrite the second line of the recursion as

$$\begin{aligned}
&E_{n,x}\left[G\left(X^u_{n+1}, g_{n+1}(X^u_{n+1})\right)\right] - G\left(x, E_{n,x}\left[g_{n+1}(X^u_{n+1})\right]\right)\\
&= E_{n,x}\left[G\left(X^u_{n+1}, g_{n+1}(X^u_{n+1})\right)\right] - G(x, g_n(x))\\
&- \left\{G\left(x, E_{n,x}\left[g_{n+1}(X^u_{n+1})\right]\right) - G(x, g_n(x))\right\}.
\end{aligned}$$

In order to simplify this we need to introduce some new notation.

Definition 3.5 *The function sequence* $\{G \diamond g\}_k$ *and, for a fixed* $z \in \mathcal{X}$*, the mapping* $G^z : \mathcal{X} \to \mathbf{R}$ *are defined by*

$$\begin{aligned}
\left(G \diamond g\right)_k(y) &= G(y, g_k(y)),\\
G^z(y) &= G(z, y).
\end{aligned}$$

With this notation we can write

$$\begin{aligned}
&E_{n,x}\left[G\left(X^u_{n+1}, g_{n+1}(X^u_{n+1})\right)\right] - G\left(x, E_{n,x}\left[g_{n+1}(X^u_{n+1})\right]\right)\\
&= \mathbf{A}^u\left(G \diamond g\right)_n(x) - \left\{G^x(\mathbf{P}^u g_n(x)) - G^x(g_n(x))\right\}.
\end{aligned}$$

We now introduce the last piece of new notation.

Definition 3.6 *With notation as above we define the function sequence* $\left\{\mathbf{H}^u_g G\right\}_k$ *by*

$$\left\{\mathbf{H}^u_g G\right\}_n(x) = G^x(\mathbf{P}^u g_n(x)) - G^x(g_n(x)).$$

Finally, we may state the main result for the present form of J_n.

Proposition 3.2 *Consider a functional of the form*

$$J_n(x, \mathbf{u}) = G\left(x, E_{n,x}\left[X^{\mathbf{u}}_T\right]\right),$$

and assume that an equilibrium control law $\hat{\mathbf{u}}$ *exists. Then, with notation as above, the equilibrium value function* V *satisfies the equation.*

$$\sup_{u \in \mathcal{U}}\left\{\left(\mathbf{A}^u V\right)_n(x) - \mathbf{A}^u\left(G \diamond g\right)_n(x) + \mathbf{H}^u_g G_n(x)\right\} = 0, \quad (13)$$

$$V_T(x) = G(x, x), \quad (14)$$

where the supremum above is realized by $u = \hat{\mathbf{u}}_n(x)$*. Furthermore, the following hold.*

1. *The function sequence* $g_n(x)$ *is determined by the recursion.*

$$\begin{aligned}
\mathbf{A}^{\hat{\mathbf{u}}} g_n(x) &= 0, \quad n = 0, \ldots, T-1, & (15)\\
g_T(x) &= x, & (16)
\end{aligned}$$

2. *The probabilistic interpretation of* g *is , as before, given by*

$$g_n(x) = E_{n,x}\left[X^{\hat{\mathbf{u}}}_T\right].$$

3.4 The case $J_n(x, \mathbf{u}) = E_{n,x}\left[F(x, X_T^{\mathbf{u}})\right] + G\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right)$.

For this form of the objective functional the extended Bellman equation is easily obtained from the results for the special cases discussed in Sections 3.2 and 3.3. With notation as above, we obtain the following result, which is basically a superposition of the results for the special cases.

Theorem 3.1 *Consider a functional of the form (2), and assume that an equilibrium control law* $\hat{\mathbf{u}}$ *exists. Then, with notation as in Section 2.1, the equilibrium value function* V *satisfies the equation.*

$$\begin{aligned} \sup_{u \in \mathcal{U}} \{ (\mathbf{A}^u V)_n(x) - (\mathbf{A}^u f)_n(x,x) + (\mathbf{A}^u f^x)_n(x) & \\ - \mathbf{A}^u (G \diamond g)_n(x) + \mathbf{H}_g^u G_n(x) \} &= 0, \quad (17) \\ V_T(x) = F(x,x) + G(x,x), & \quad (18) \end{aligned}$$

where the supremum above is realized by $u = \hat{\mathbf{u}}_n(x)$*. Furthermore, the following hold.*

1. *For every fixed* $y \in \mathcal{X}$ *the function sequence* $f_n^y(x)$ *is determined by the recursion*

$$\begin{aligned} \mathbf{A}^{\hat{\mathbf{u}}} f_n^y(x) &= 0, \quad n = 0, \dots, T-1, \quad (19) \\ f_T^y(x) &= F(y,x), \quad (20) \end{aligned}$$

and $f_n(x,x)$ *is given by*

$$f_n(x,x) = f_n^x(x).$$

2. *The function sequence* $g_n(x)$ *is determined by the recursion.*

$$\begin{aligned} \mathbf{A}^{\hat{\mathbf{u}}} g_n(x) &= 0, \quad n = 0, \dots, T-1, \quad (21) \\ g_T(x) &= x, \quad (22) \end{aligned}$$

3. *The probabilistic interpretations of* f *and* g *are, as before, given by*

$$\begin{aligned} f_n^y(x) &= E_{n,x}\left[F\left(y, X_T^{\hat{\mathbf{u}}}\right)\right], \\ g_n(x) &= E_{n,x}\left[X_T^{\hat{\mathbf{u}}}\right]. \end{aligned}$$

We now have some comments on this result.

Remark 3.1
- *The first point to notice is that, as opposed to a standard time consistent problem where we would have* **one** *equation for the determination of the optimal value function* V*, we now have a* **system** *of recursion equations (17)-(22) for the simultaneous determination of* V*,* f *and* g*.*

- *To see the recursive structure more clearly we can rewrite the equation for V in the form*

$$\begin{aligned} V_n(x) &= \sup_{u \in \mathcal{U}} \left\{ E_{n,x}\left[V_{n+1}\left(X^u_{n+1} \right) \right] - \left(\mathbf{A}^u f \right)_n (x,x) + \left(\mathbf{A}^u f^x \right)_n (x) \right. \\ &- \left. \mathbf{A}^u \left(G \diamond g \right)_n (x) + \mathbf{H}^u_g G_n(x) \right\} . \end{aligned}$$

  The recursions for f and g can similarly be written as

$$\begin{aligned} f^y_n(x) &= E_{n,x}\left[f^y_{n+1}\left(X^{\hat{\mathbf{u}}}_{n+1} \right) \right] \\ g_n(x) &= E_{n,x}\left[g_{n+1}(X^{\hat{\mathbf{u}}}_{n+1} \right] . \end{aligned}$$

  This is thus the backward recursion scheme discussed in Remark 2.3.

- *In the case when $F(x, y)$ does not depend upon x, and there is no G term, the problem trivializes to a standard time consistent problem. The terms $\left(\mathbf{A}^u f \right)_n (x,x) + \left(\mathbf{A}^u f^x \right)_n (x)$ in the V-equation (17) cancel, and the system reduces to the standard Bellman equation*

$$\begin{aligned} V_n(x) &= \sup_{u \in \mathcal{U}} E_{n,x}\left[V_{n+1}\left(X^u_{n+1} \right) \right] , \\ V_T(x) &= F(x). \end{aligned}$$

- *In order to solve the V-equation (17) we need to know f and g but these are determined by the equilibrium control law $\hat{\mathbf{u}}$, which in turn is determined by the sup-part of (17).*

- *We can view the system as a fixed point problem, where the equilibrium control law $\hat{\mathbf{u}}$ solves an equation of the form $M(\hat{\mathbf{u}}) = \hat{\mathbf{u}}$. The mapping M is defined by the following procedure.*

  - *Start with a control $\mathbf{u}$.*
  - *Generate the functions f and g by the recursions*

$$\begin{aligned} \mathbf{A}^{\mathbf{u}} f^y_n(x) &= 0, \\ \mathbf{A}^{\mathbf{u}} g_n(x) &= 0, \end{aligned}$$

    and the obvious terminal conditions.
  - *Now plug these choices of f and g into the V equation and solve it for V. The control law which realizes the sup-part in (17) is denoted by $M(\mathbf{u})$. The optimal control law is determined by the fixed point problem $M(\hat{\mathbf{u}}) = \hat{\mathbf{u}}$.*

  This fixed point property is rather expected since we are looking for a Nash equilibrium point, and it is well known that such a point is typically determined as fixed points of a mapping. We also note that we can view the system as a fixed point problem for f and g.

3.5 The general case

We finally consider the most general functional form, where J is given by

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{T-1} C_{n,k}\left(x, X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}})\right) + F_n(x, X_T^{\mathbf{u}})\right] + G_n\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right). \tag{23}$$

This case differs from the case above, firstly by the introduction of the sum, and secondly by allowing F and G to depend on current time n. The arguments for the C terms in the sum above are very similar to the previous arguments for the F term, and the occurrence of present time n can be handled very much like the occurrence of the variable x. It is therefore natural to introduce indexed function sequences defined by

$$f_n^{ky}(x) = E_{n,x}\left[F_k\left(y, X_T^{\hat{\mathbf{u}}}\right)\right], \tag{24}$$

$$g_n(x) = E_{n,x}\left[X_T^{\hat{\mathbf{u}}}\right]. \tag{25}$$

$$c_n^{k,m,y}(x) = E_{n,x}\left[C_{k,m}\left(y, X_m^{\hat{\mathbf{u}}}, \hat{\mathbf{u}}_m\left(X_m^{\hat{\mathbf{u}}}\right)\right)\right] \tag{26}$$

where, as usual, $\hat{\mathbf{u}}$ denotes the equilibrium law. Arguing very much like in the simplified cases above we then have the following main result.

Theorem 3.2 *Consider a functional of the form (23), and assume that an equilibrium control law $\hat{\mathbf{u}}$ exists. Then, with notation as in Section 2.1, the the equilibrium value function V satisfies the equation.*

$$\sup_{u\in\mathcal{U}}\Big\{\left(\mathbf{A}^u V\right)_n(x) + C_{nn}(x,x,u) - \sum_{m=n+1}^{T-1}(\mathbf{A}^u c^m)_{nn}(x,x) + \sum_{m=n+1}^{T-1}(\mathbf{A}^u c^{nmx})_n(x)$$

$$-\left(\mathbf{A}^u f\right)_{nn}(x,x) + \left(\mathbf{A}^u f^{nx}\right)_n(x) - \mathbf{A}^u\left(G\diamond g\right)_n(x) + \mathbf{H}_g^u G_n(x)\Big\} = 0, \tag{27}$$

$$V_T(x) = F_T(x,x) + G_T(x,x), \tag{28}$$

where the supremum above is realized by $u = \hat{\mathbf{u}}_n(x)$. Furthermore, the following hold.

1. *For every fixed $k = 0, 1, \ldots, T$ and every $y \in \mathcal{X}$ the function sequence $f_n^{ky}(x)$ is determined by the recursion*

$$\mathbf{A}^{\hat{\mathbf{u}}} f_n^{ky}(x) = 0, \quad n = 0, \ldots, T-1, \tag{29}$$

$$f_T^{ky}(x) = F_k(y, x), \tag{30}$$

and $f_{nn}(x,x)$ is defined by

$$f_{nn}(x,x) = f_n^{nx}(x).$$

2. *The function sequence $g_n(x)$ is determined by the recursion.*

$$\mathbf{A}^{\hat{\mathbf{u}}} g_n(x) = 0, \quad n = 0, \ldots, T-1, \tag{31}$$

$$g_T(x) = x, \tag{32}$$

3. *For every $k, m = 0, 1, \ldots, T$, with $k \leq m$, and $y \in \mathcal{X}$ the function sequence $c_n^{kmy}(x)$ is defined by*

$$(\mathbf{A}^{\hat{\mathbf{u}}} c^{k,m,y})_n(x) = 0, \quad 0 \leq n \leq m-1, \tag{33}$$

$$c_m^{k,m,y}(x) = C_{k,m}(y, x, \hat{\mathbf{u}}_m(x)). \tag{34}$$

   and $c_{nn}^m(x, x)$ is defined by

$$c_{nn}^m(x, x) = c_n^{nmx}(x).$$

4. *The probabilistic interpretations of f, g and c are given by (24)-(26).*

5. *In the expressions above, $\hat{\mathbf{u}}$ always denotes the equilibrium control law.*

6. *Recall that the operators $\mathbf{A}^u$ and $\mathbf{A}^{\hat{\mathbf{u}}}$ only operates on lower case time indices and variables within parentheses. Upper case indices are treated as constant parameters.*

The detailed discussion at the end of Section 3.4 carries over also to this case.

3.6 Control constraints

In the discussions above we have assumed that there are no constraints on the controls, so at time n we are allowed to choose any $u_n \in \mathcal{U}$. The theory above can easily be extended to the case when we have constraints of the form

$$u_n \in D_n(X_n),$$

where, for each n and x, $D_n(x)$ is a subset of $\mathcal{U}$. In the extended Bellman system we only have to change the expression $\sup_{u \in \mathcal{U}}$ to $\sup_{u \in D_n(x)}$.

3.7 A slight extension

We can easily extend the theory above to the case when the term

$$G_n\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right)$$

is replaced by

$$G_n\left(x, E_{n,x}\left[h(X_T^{\mathbf{u}})\right]\right)$$

for some real valued function h. In this case we simply define the g sequence by

$$g_n(x) = E_{n,x}\left[h(X_T^{\hat{\mathbf{u}}})\right].$$

Theorem 3.2 will still hold, apart from the fact that the boundary condition for g will be replaced by

$$g_T(x) = h(x).$$

3.8 A scaling result

In this section we derive a small scaling result, which is sometimes quite useful. Consider the objective functional (23) above and denote, as usual, the equilibrium control and value function by $\hat{\mathbf{u}}$ and V respectively. Let $\varphi : \mathcal{X} \to R_+$ be a fixed real valued function and consider a new objective functional J^φ, defined by,

$$J_n^\varphi(x, \mathbf{u}) = \varphi(x) J_n(x, \mathbf{u}), \quad n = 0, 1, \dots, T$$

and denote the corresponding equilibrium control and value function by $\hat{\mathbf{u}}^\varphi$ and V^φ respectively. Since player No n is (loosely speaking) trying to maximize $J_n^\varphi(x, \mathbf{u})$ over u_n, and $\varphi(x)$ is just a scaling factor which is not affected by u_n the following result is intuitively obvious.

Proposition 3.3 *Assume that $\varphi(x) J_n(x, \mathbf{u})$ is integrable for all $(n, \mathbf{u})$ [2]. With notation as above we then have*

$$\begin{aligned} V_n^\varphi(x) &= \varphi(x) V_n(x), \\ \hat{\mathbf{u}}_n^\varphi(x) &= \hat{\mathbf{u}}_n(x). \end{aligned}$$

Proof. The result follows from an easy (but messy) induction argument. ∎

4 An equivalent time consistent problem

The object of the present section is to provide a surprising equivalence result between time inconsistent and time consistent problems. To this end we go back to the general extended HJB system of equations. The first part of this reads as

$$\sup_{u \in \mathcal{U}} \Big\{ (\mathbf{A}^u V)_n (x) + C_{nn}(x, x, u) - \sum_{m=n+1}^{T-1} (\mathbf{A}^u c^m)_{nn}(x, x) + \sum_{m=n+1}^{T-1} (\mathbf{A}^u c^{nmx})_n(x)$$

$$- (\mathbf{A}^u f)_{nn} (x, x) + (\mathbf{A}^u f^{nx})_n (x) - \mathbf{A}^u (G \diamond g)_n (x) + \mathbf{H}_g^u G_n(x) \Big\} = 0,$$

Now consider the equilibrium control law $\hat{\mathbf{u}}$. Using $\hat{\mathbf{u}}$ we can then construct f, g, and c by solving the equations (24)-(26). We now define the function h by

$$\begin{aligned} h_n(x, u) &= C_{nn}(x, x, u) - \sum_{m=n+1}^{T-1} (\mathbf{A}^u c^m)_{nn}(x, x) + \sum_{m=n+1}^{T-1} (\mathbf{A}^u c^{nmx})_n(x) \\ &- (\mathbf{A}^u f)_{nn} (x, x) + (\mathbf{A}^u f^{nx})_n (x) - \mathbf{A}^u (G \diamond g)_n (x) + \mathbf{H}_g^u G_n(x), \end{aligned}$$

[2] An easy sufficient condition is that φ takes values in $[0, 1)$

where it is important to notice that h does not involve the equilibrium value function V. With this definition of h, the equation for V above and its boundary condition become

$$\begin{aligned} \sup_{u \in \mathcal{U}} \left\{ (\mathbf{A}^u V)_n (x) + h_n(x,u) \right\} &= 0, \\ V(T,x) &= H(x), \end{aligned}$$

where H is defined by $H(x) = F(x,x) + G(x,x)$. We now observe, simply by inspection, that this is a standard HJB equation for the standard time consistent optimal control problem to maximize

$$E_{n,x} \left[\sum_{k=n}^{T} h_k(X_k, u_k) + H(X_T) \right]. \tag{35}$$

We have thus proved the following result.

Proposition 4.1 *For every time inconsistent problem in the present framework there exists a standard, time consistent, optimal control problem with the following properties.*

- *The optimal value function for the standard problem coincides with the equilibrium value function for the time inconsistent problem.*
- *The optimal control for the standard problem coincides with the equilibrium control for the time inconsistent problem.*
- *The objective functional for the standard problem is given by (35).*

We immediately remark that Proposition 4.1 above is mostly of theoretical interest, and of little "practical" value. The reason is of course that in order to formulate the equivalent standard problem we need to know the equilibrium control $\hat{\mathbf{u}}$. In our opinion it is however quite surprising.

Related results can be found in [1], [7], [10] and [13]. In these papers it is proved that, for various models where time inconsistency stems from non-exponential discounting, there exists an equivalent standard problem (with exponential discounting).

Proposition 4.1 differs from the results in the cited references above two ways. Firstly it differs by being quite general and not confined to a particular model. Secondly it differs from the results in the cited references by having a different structure. In other words, for the models studied in the cited papers, the equivalent problem described in Proposition 4.1 is structurally different from the equivalent problems presented in the cited references. See Section 8.3 for a more detailed discussion of issues of this kind.

Furthermore, Proposition 4.1 has modeling consequences for economics. Suppose that you want to model consumer behavior. You have done this using standard time consistent dynamic utility maximization and now you are contemplating to introduce time inconsistent preferences to obtain a richer class of consumer behavior. Proposition 4.1 then tells us that from the point of view of revealed preferences, nothing is gained by introducing time inconsistent preferences: Every kind of behavior that can be generated by time inconsistency can also be generated by time consistent preferences. We immediately remark, however, that even if a concrete model of time inconsistent preferences is, in some sense, "natural", the corresponding time consistent preferences may look extremely "weird".

5 Infinite horizon

In the arguments above we have always assumed that the time horizon T is finite. In many applications, however, it is natural to consider a problem with infinite horizon.

5.1 Generalities

We consider a value functional of the form

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{\infty} C_{n,k}\left(x, X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}})\right)\right], \tag{36}$$

Using exactly the same arguments as above, we then have the following result.

Proposition 5.1 *Consider a functional of the form (36), and assume that an equilibrium control law $\hat{\mathbf{u}}$ exists. Then the the corresponding equilibrium value function V satisfies the equation.*

$$\sup_{u\in\mathcal{U}}\left\{\left(\mathbf{A}^u V\right)_n(x) + C_{nn}(x,x,u) - \sum_{m=n+1}^{\infty}(\mathbf{A}^u c^m)_{nn}(x,x) + \sum_{m=n+1}^{\infty}(\mathbf{A}^u c^{nmx})_n(x)\right\} = 0$$

where the supremum above is realized by $u = \hat{\mathbf{u}}_n(x)$. The function $c_n^{k,m,y}(x)$ is as usual defined by (26).

For the infinite horizon case, existence and uniqueness of an equilibrium control is highly nontrivial. See Section 6 for more details.

5.2 A time invariant problem

The infinite sums in Proposition 5.1 look rather forbidding, but one would expect a simpler formulation of the extended Bellman equation in the case of a time invariant problem. To investigate this we consider the case when the controlled Markov process X is time invariant, and the reward functional has the time invariant form

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{\infty} \delta^{k-n} H\left(x, X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}})\right)\right].$$

for some real valued function H and some δ with $0 < \delta < 1$. In the notation of (36) we thus have

$$C_{n,k}(x,y,u) = \delta^{k-n} H(x,y,u).$$

In the notation of Proposition 5.1 we now have

$$c_n^{k,m,y}(x) = \delta^{m-k} k_n^{m,y}(x),$$

where

$$k_n^{m,y}(x) = E_{n,x}\left[H\left(y, X_m^{\hat{\mathbf{u}}}, \hat{\mathbf{u}}_m\left(X_m^{\hat{\mathbf{u}}}\right)\right)\right]$$

and we also have

$$V_n(x) = \sum_{m=n}^{\infty} \delta^{m-n} k_n^{mx}(x).$$

We now have to compute the term

$$-(\mathbf{A}^u c^m)_{nn}(x,x) + (\mathbf{A}^u c^{n,m,x})_n(x)$$

in the extended Bellman equation. After some calculations we obtain

$$\begin{aligned}
&-(\mathbf{A}^u c^m)_{nn}(x,x) + (\mathbf{A}^u c^{n,m,x})_n(x) \\
&= \delta^{m-n} E_{n,x}\left[k_{n+1}^{m,x}(X_{n+1}^u)\right] - \delta^{m-(n+1)} E_{n,x}\left[k_{n+1}^{m,X_{n+1}^u}(X_{n+1}^u)\right].
\end{aligned}$$

We now use the time invariance property of the problem, so we look for time invariant equilibria. This implies that, with obvious notation, we have

$$\begin{aligned}
\hat{\mathbf{u}}_m(x) &= \hat{\mathbf{u}}(x), \\
k_n^{m,y}(x) &= h_{m-n}(y,x), \\
V_n(x) &= V(x), \\
V(x) &= \sum_{k=0}^{\infty} \delta^k h_k(x,x).
\end{aligned}$$

With this notation we have

$$\begin{aligned}
&-(\mathbf{A}^u c^m)_{nn}(x,x) + (\mathbf{A}^u c^{n,m,x})_n(x) \\
&= \delta^{m-n} E_{n,x}\left[h_{m-(n+1)}(x, X_{n+1}^u)\right] - \delta^{m-(n+1)} E_{n,x}\left[h_{m-(n+1)}(X_{n+1}^u, X_{n+1}^u)\right].
\end{aligned}$$

which gives us

$$\begin{aligned}
&-\sum_{m=n+1}^{\infty} (\mathbf{A}^u c^m)_{nn}(x,x) + \sum_{m=n+1}^{\infty} (\mathbf{A}^u c^{n,m,x})_n(x) \\
&= \sum_{m=0}^{\infty} \delta^{m+1} E_x\left[h_m(x, X_1^u)\right] - \sum_{m=0}^{\infty} \delta^m E_x\left[h_m(X_1^u, X_1^u)\right] \\
&= \delta \sum_{m=0}^{\infty} \delta^m E_x\left[h_m(x, X_1^u)\right] - E_x\left[V(X_1^u)\right],
\end{aligned}$$

where $E_x[\ ] = E_{0,x}[\ ]$.

Putting these terms together gives us the following result.

Proposition 5.2 *For the time invariant problem above we have the following extended Bellman equation.*

$$V(x) = \sup_u \left\{ H(x,x,u) + \delta \sum_{m=0}^{\infty} \delta^m E_x \left[h_m(x, X_1^u) \right] \right\}$$

where

$$\begin{aligned} h_{n+1}(y,x) &= E_x \left[h_n(y, X_1^{\hat{u}(x)}) \right], \\ h_0(y,x) &= H(y,x,\hat{\mathbf{u}}(x)). \end{aligned}$$

We still have an infinite sum in the Bellman equation, but we can formally get rid of that by defining $K(y,x)$ by

$$K(y,x) = \sum_{m=0}^{\infty} \delta^m h_m(y,x),$$

so we have the following corollary, where we have used the recursion for h_n.

Corollary 5.1 *The extended Bellman equation can also be written as*

$$V(x) = \sup_u \left\{ H(x,x,u) + \delta E_x \left[K(x, X_1^u) \right] \right\}$$

The functions $V(x)$ and $K(y,x)$ are jointly determined by the Bellman equation above and the recursion

$$\begin{aligned} E_x \left[K(y, X_1^{\hat{u}(x)}) \right] &= \frac{1}{\delta} \left\{ K(y,x) - H(y,x,\hat{\mathbf{u}}(x)) \right\}, \\ K(x,x) &= V(x). \end{aligned}$$

6 Existence and uniqueness

So far we have not discussed existence and uniqueness of an equilibrium control. These issues are in fact quite complicated and there is also a marked difference between the case $T < \infty$ and the case $T = \infty$.

For the case $T < \infty$, existence and uniqueness does not, at least at first sight, seem to be complicated. Theorem 3.2 does in fact give us a concrete backward recursion for the equilibrium value function, starting at $n = T$ (see Remark 3.1). For a well posed problem with a unique optimizer in the extended Bellman equation for all n, the equilibrium control and the corresponding value function are thus determined recursively.

This simple intuition is, however, quite hard to formalize, in the sense that we have, so far failed in providing good sufficient conditions for the existence of an equilibrium control.

Concerning non-uniqueness, it is obvious that it could happen that for some n there is more than one global optimum in the extended Bellman equation. This would then lead to non unique equilibrium strategies, and it could also lead to non unique-equilibrium value processes. This is in contrast to a time consistent control problem, where we also may have several optimal control strategies, but where the optimal value function (by definition) is unique.

The case $T = \infty$ is even more complicated. For this case we have a recursion, namely the extended Bellman equation, but we have no natural boundary condition. Existence is thus a highly non trivial issue, and we have so far not been able to obtain any general results in this direction, so this is an object of future research. Concerning uniqueness, we conjecture that we may generically have multiple equilibria. See [7] fort an example of non-uniqueness and [17] for a detailed study of the existence of multiple equilibria in a concrete case.

7 General non exponential discounting

Problems with non exponential discounting constitute an important subclass of the family of time inconsistent problems. To see how the general theory works in this more concrete case we now consider a fairly general model class with non-exponential discounting. As a special case we will then study the case of hyperbolic discounting. The general model is specified as follows.

- We assume that the controlled Markov process X is time homogeneous.
- The value functional for player n is given by

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{T-1} \varphi(k-n)H(X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}})) + \varphi(T-n)K(X_T)\right] \quad (37)$$

- In the expression above, the discounting function φ, and the utility functions H and K, are assumed to be given deterministic functions.
- Without loss of generality we assume that

$$\varphi(0) = 1.$$

We note that if the discounting function φ has the form

$$\varphi(k) = \delta^k, \quad k = 0, 1, 2, \ldots$$

then we have a standard time consistent control problem with infinite horizon. The interesting case for us is thus the case when φ is not of exponential form.

7.1 A general discount function

We see that, in the notation of Theorem 3.2 we have

$$\begin{aligned} C_{n,k}(y, x, u) &= \varphi(k-n)H(x, u), \\ F_n(y, x) &= \varphi(T-n)K(x), \end{aligned}$$

so for this model we have no y-variable. The extended Bellman equation for this case is now easily seen to have the form

$$\sup_u \left\{ \left(\mathbf{A}^u V\right)_n(x) + H(x,u) + \sum_{m=n+1}^{T-1} \left[\mathbf{A}^u c_n^{nm}(x) - \mathbf{A}^u c_{nn}^m(x)\right] - \mathbf{A}^u f_{nn}(x) + \mathbf{A}^u f_n^n(x) \right\} = 0$$

where

$$\begin{aligned} c_n^{km}(x) &= \varphi(m-k) E_{n,x}\left[H\left(X_m^{\hat{\mathbf{u}}}, \hat{\mathbf{u}}_m(X_m^{\hat{\mathbf{u}}})\right)\right], \\ c_{nn}^m(x) &= c_n^{nm}(x), \\ f_n^k(x) &= \varphi(T-k) E_{n,x}\left[K\left(X_T^{\hat{\mathbf{u}}}\right)\right], \\ f_{nn}(x) &= f_n^n(x). \end{aligned}$$

We also recall that the operator $\mathbf{A}^u$ only operates on lower case indices and the variable inside the the parenthesis. In this setting it is natural to define $h_n^m(x)$ and $k_n(x)$ by

$$\begin{aligned} h_n^m(x) &= E_{n,x}\left[H\left(X_m^{\hat{\mathbf{u}}}, \hat{\mathbf{u}}_m(X_m^{\hat{\mathbf{u}}})\right)\right], \\ k_n(x) &= E_{n,x}\left[K\left(X_T^{\hat{\mathbf{u}}}\right)\right], \end{aligned}$$

so we have

$$\begin{aligned} c_n^{km}(x) &= \varphi(m-k) h_n^m(x), \\ f_n^k(x) &= \varphi(T-k) k_n(x). \end{aligned}$$

With this notation we obtain

$$\begin{aligned} \mathbf{A}^u c_n^{nm}(x) &= E_{nx}\left[c_{n+1}^{nm}(X_{n+1}^u)\right] - c_n^{nm}(x) \\ &= E_{nx}\left[\varphi(m-n) h_{n+1}^m(X_{n+1}^u)\right] - \varphi(m-n) h_n^m(x) \\ &= \varphi(m-n) \mathbf{A}^u h_n^m(x), \end{aligned}$$

and in the same way

$$\mathbf{A}^u f_n^n(x) = \varphi(T-n) \mathbf{A}^u k_n(x)$$

Furthermore we obtain

$$\begin{aligned} \mathbf{A}^u c_{nn}^m(x) &= E_{nx}\left[c_{n+1}^{n+1,m}(X_{n+1}^u)\right] - c_n^{nm}(x) \\ &= E_{nx}\left[\varphi(m-n-1) h_{n+1}^m(X_{n+1}^u)\right] - \varphi(m-n) h_n^m(x) \\ &= \varphi(m-n-1) \mathbf{A}^u h_n^m(x) - \Delta\varphi(m-n) h_n^m(x), \end{aligned}$$

where we have used the notation

$$\Delta\varphi(k) = \varphi(k) - \varphi(k-1),$$

and in the same way we obtain

$$\mathbf{A}^u f_{nn}(x) = \varphi(T-n-1) \mathbf{A}^u k_n(x) - \Delta\varphi(T-n) k_n(x).$$

For the terms in the sum in the Bellman equation we thus have

$$\begin{aligned}\mathbf{A}^u c_n^{nm}(x) - \mathbf{A}^u c_{nn}^m(x) &= \Delta\varphi(m-n)\mathbf{A}^u h_n^m(x) + \Delta\varphi(m-n)h_n^m(x)\\ &= \Delta\varphi(m-n)E_{nx}\left[h_{n+1}^m(X_{n+1}^u)\right]\\ &= \Delta\varphi(m-n)\mathbf{P}^u h_n^m(x),\end{aligned}$$

and similarly

$$\mathbf{A}^u f_n^n(x) - \mathbf{A}^u f_{nn}(x) = \Delta\varphi(T-n)\mathbf{P}^u k_n(x).$$

We thus have the following main result for non exponential discounting.

Proposition 7.1 *The extended Bellman system for the non exponential discounting problem (37) is given by*

$$\begin{aligned}\sup_u\left\{(\mathbf{A}^u V)_n(x) + H(x,u) + \Delta\varphi(T-n)\mathbf{P}^u k_n(x) + \sum_{m=n+1}^{T-1}\Delta\varphi(m-n)\mathbf{P}^u h_n^m(x)\right\} &= 0,\\ V_T(x) &= K(x),\end{aligned}$$

where the operator $\mathbf{P}^u$ *is defined in Definition 2.1. The function sequences* h *and* k *are defined by*

$$\begin{aligned}h_n^m(x) &= E_{n,x}\left[H\left(X_m^{\hat{\mathbf{u}}}, \hat{\mathbf{u}}_m(X_m^{\hat{\mathbf{u}}})\right)\right],\\ k_n(x) &= E_{n,x}\left[K\left(X_T^{\hat{\mathbf{u}}}\right)\right],\end{aligned}$$

and satisfy the recursions

$$\begin{aligned}\mathbf{A}^{\hat{\mathbf{u}}} h_n^m(x) &= 0, \quad 0 \le n \le m,\\ h_m^m(x) &= H(x, \hat{\mathbf{u}}_m(x)),\\ \mathbf{A}^{\hat{\mathbf{u}}} k_n(x) &= 0, \quad 0 \le n \le T,\\ k_T(x) &= K(x).\end{aligned}$$

7.2 Infinite horizon

We now consider special the case when the time horizon T is infinite, so the value functional is given by

$$J_n(x,\mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{\infty}\varphi(k-n)H(X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}}))\right] \tag{38}$$

In this case we have no F term so the extended equation takes the form

$$\sup_u\left\{(\mathbf{A}^u V)_n(x) + H(x,u) + \sum_{m=n+1}^{\infty}\Delta\varphi(m-n)\mathbf{P}^u h_n^m(x)\right\} = 0$$

We may now use the fact that the setup is time homogeneous. The value function, will then be time invariant and we have

$$J_n(x, \mathbf{u}) = J(x, \mathbf{u}),$$

where

$$J(x, \mathbf{u}) = E_x \left[\sum_{k=0}^{\infty} \varphi(k) H(X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}}) \right]$$

The equilibrium control, as well as the equilibrium value function, will also be time invariant so we have

$$\begin{aligned} \hat{\mathbf{u}}_n(x) &= \hat{\mathbf{u}}(x), \\ V_n(x) &= V(x). \end{aligned}$$

We also see that h will be time invariant in the sense that $h_n^m(x) = h_0^{m-n}(x)$ so we make the definition

$$h_m(x) = h_n^{n+m}(x) = h_0^m(x),$$

and we note that

$$h_0(x) = H(x, \hat{\mathbf{u}}(x)).$$

Using the definition of $h_n^k(x)$ we also see that in fact

$$V(x) = \sum_{k=0}^{\infty} \varphi(k) h_k(x).$$

The Bellman equation now takes the form

$$\sup_u \left\{ \mathbf{A}^u V_n(x) + H(x, u) + \sum_{k=1}^{\infty} \Delta\varphi(k) \mathbf{P}^u h_k(x) \right\} = 0$$

where, by definition,

$$\begin{aligned} \mathbf{P}^u h_k(x) &= \mathbf{P}^u h_n^{n+k}(x) = E_{nx} \left[h_{n+1}^{n+k}(X_{n+1}^u) \right] \\ &= E_{0x} \left[h_0^{k-1}(X_1^u) \right] = E_x \left[h_{k-1}(X_1^u) \right] \end{aligned}$$

We can thus state the final result.

Proposition 7.2 *For the infinite horizon problem (38) the extended Bellman system has the form*

$$\sup_u \left\{ \mathbf{A}^u V_n(x) + H(x, u) + \sum_{k=1}^{\infty} \Delta\varphi(k) \mathbf{P}^u h_k(x) \right\} = 0$$

or, alternatively,

$$V(x) = \sup_u \left\{ H(x, u) + E_x \left[V(X_1^u) \right] + E_x \left[\sum_{k=1}^{\infty} \Delta\varphi(k) h_{k-1}(X_1^u) \right] \right\}.$$

In both cases, the function sequence $h^k(x)$ is determined by the recursion

$$\begin{aligned} h_{k+1} &= E_x\left[h_k(X_1^{\hat{\mathbf{u}}})\right], \\ h_0(x) &= H(x, \hat{\mathbf{u}}(x)). \end{aligned}$$

8 Quasi-hyperbolic discounting

In this section we study the special case when the discount function φ is a "quasi-hyperbolic discounting function". More precisely we assume that φ has the form

$$\varphi(0) = 1, \tag{39}$$

$$\varphi(k) = \beta\delta^k, \quad k = 1, 2, 3, \ldots \tag{40}$$

where $\beta \geq 0$ and $0 < \delta < 1$. We restrict ourselves to infinite horizon problems of the type studied in Section 7.2.

8.1 The extended Bellman equation

In the case of quasi-hyperbolic discounting, Proposition 7.2 simplifies considerably. We have

$$\begin{aligned} \Delta\varphi(1) &= \beta\delta - 1, \\ \Delta\varphi(k) &= \beta\delta^{k-1}(\delta - 1), \quad k = 1, 2, \ldots \end{aligned}$$

Using the relation

$$V(x) = \sum_{k=0}^{\infty} \varphi(k) h^k(x)$$

we obtain

$$\begin{aligned} &E_x\left[\sum_{k=1}^{\infty} \Delta\varphi(k) h^{k-1}(X_1^u)\right] \\ &= (\beta\delta - 1)E_x\left[h^0(X_1^u)\right] + \beta(\delta - 1)\sum_{k=2}^{\infty} \delta^{k-1} E_x\left[h^{k-1}(X_1^u)\right] \\ &= (\beta\delta - 1)E_x\left[h^0(X_1^u)\right] + (\delta - 1)\sum_{k=1}^{\infty} \beta\delta^k E_x\left[h^k(X_1^u)\right] \\ &= (\delta - 1)E_x\left[h^0(X_1^u)\right] + (\delta - 1)\sum_{k=1}^{\infty} \beta\delta^k E_x\left[h^k(X_1^u)\right] + \delta(\beta - 1)E_x\left[h^0(X_1^u)\right] \\ &= (\delta - 1)E_x\left[V(X_1^u)\right] + \delta(\beta - 1)E_x\left[h^0(X_1^u)\right]. \end{aligned}$$

Plugging this into the recursive equation in Proposition 7.2 we thus have our main result for the hyperbolic discounting case.

Proposition 8.1 *For a reward functional of the form (38) and quasi-hyperbolic discounting, the extended Bellman system takes the form*

$$V(x) = \sup_u \left\{ \delta E_x \left[V(X_1^u) \right] + H(x,u) + \delta(\beta - 1) E_x \left[H(X_1^u, \hat{\mathbf{u}}(X_1^u)) \right] \right\}. \quad (41)$$

Remark 8.1 *We note that if $\beta = 1$ we have the standard exponential discounting $\varphi(k) = \delta^k$ and the extended Bellman equation in Proposition 8.1 trivializes to*

$$V(x) = \sup_u \left\{ \delta E_x \left[V(X_1^u) \right] + H(x,u) \right\}.$$

which is the classical Bellman equation for the exponential discounting case.

8.2 An example with logarithmic utility

We now exemplify the theory developed above by studying a simple case of portfolio optimization with log utility and quasi-hyperbolic discounting.

More precisely we consider a financial market consisting of a risky asset with price process S and a risk free asset with price process B. Recalling that for any process ξ the process $\Delta\xi$ is defined by $\Delta\xi_n = \xi_n - \xi_{n-1}$, we assume that the dynamics of S and B are given by

$$\begin{aligned} \Delta S_{n+1} &= S_n Y_{n+1}, \\ \Delta B_{n+1} &= B_n r, \end{aligned}$$

where the random returns $\{Y_n\}_{n=1}^{\infty}$ is a sequence of i.i.d. random variables with $Y_n \geq -1$, and the short rate r is assumed to be constant.

We denote by c_n the consumption at time n and by w_n the portfolio weight on the risky asset at time n. Furthermore, we let X_n denote the market value of the portfolio at time n, just before consumption. In other words: X_n is the value of the portfolio going into time n. At time n we consume c_n, and rebalance the portfolio with remaining value $X_n - c_n$. With these definitions it is easy to see that the dynamics of the value process are given by

$$X_{n+1} = (X_n - c_n)(R + wZ_{n+1})$$

where $R = 1 + r$ and $Z_n = Y_n - r$. We now consider a portfolio problem with log utility and non exponential discounting, so the reward functional is given by

$$J_n(x, \mathbf{c}, \mathbf{w}) = E_{n,x} \left[\sum_{k=n}^{\infty} \varphi(k-n) \ln c_k \right]$$

We specialize to quasi-hyperbolic discounting so from Proposition 8.1 we obtain

$$V(x) = \sup_{w \in R, c \leq x} \left\{ \delta E_x \left[V(X_1^u) \right] + \ln c + \delta(\beta - 1) E_x \left[\ln \left(\hat{c}(X_1^u) \right) \right] \right\} \quad (42)$$

where we use u as shorthand for (c, w). We now make an *Ansatz*, i.e. a trial solution of the form

$$\begin{aligned} V(x) &= A \ln x + B, \\ \hat{c}(x) &= Dx, \end{aligned}$$

where A, B and D are constants. With this Ansatz we easily obtain

$$E_x \left[V(X_1^u) \right] = A \ln(x - c) + Ag(w) + B.$$

where we have used the notation

$$g(w) = E_x \left[\ln \left(R + wZ \right) \right].$$

Given a specification for the distribution of Z, the function g is thus a known concave function of the real variable w. We assume that the distribution of Z is such that g has a maximum point w^*, and we denote $g(w^*)$ by $\hat{g}$.

In the same way we obtain

$$E_x \left[\ln \left(\hat{c}_1(X_1^u) \right) \right] = E_x \left[\ln(DX_1^u) \right] = \ln D + \ln(x - c) + g(w) \tag{43}$$

Using these results we see that the extended Bellman equation takes the form

$$\begin{aligned} A \ln x + B &= \sup_{w,c} \left\{ \delta \left[A \ln(x - c) + Ag(w) + B \right] \right. \\ &+ \left. \ln c + \delta(\beta - 1) \left[\ln D + \ln(x - c) + g(w) \right] \right\} \end{aligned}$$

We thus have to optimize

$$\delta(A + \beta - 1) g(w)$$

over w, and optimize

$$\delta(A + \beta - 1) \ln(x - c) + \ln c$$

over $c \leq x$. Assuming that $A + \beta - 1 > 0$ (see Remark 8.2) we obtain

$$\begin{aligned} \hat{w} &= w^*, \\ \hat{c} &= \frac{x}{1 + \delta(A + \beta - 1)}. \end{aligned}$$

This allows us to identify D in the Ansatz as

$$D = \frac{1}{1 + \delta(A + \beta - 1)}.$$

Plugging all these expressions into the extended Bellman equation gives us

$$\begin{aligned} A \ln x + B &= \delta A \ln x + \delta A \ln \left[\frac{\delta(A + \beta - 1)}{1 + \delta(A + \beta - 1)} \right] + \delta A \hat{g} + \delta B \\ & + \ln x - \ln \left[1 + \delta(A + \beta - 1) \right] \\ & - \delta(\beta - 1) \ln \left[1 + \delta(A + \beta - 1) \right] + \delta(\beta - 1) \hat{g} \\ & + \delta(\beta - 1) \ln x + \delta(\beta - 1) \ln \left[\frac{\delta(A + \beta - 1)}{1 + \delta(A + \beta - 1)} \right] \end{aligned}$$

We can now identify the terms containing $\ln x$ and the constant terms to obtain

$$
\begin{aligned}
A &= \delta A + 1 + \delta(\beta - 1), \\
B &= \delta A \ln \left[\frac{\delta(A + \beta - 1)}{1 + \delta(A + \beta - 1)} \right] + \delta A \hat{g} + \delta B \\
&\quad - \ln \left[1 + \delta(A + \beta - 1) \right] \\
&\quad - \delta(\beta - 1) \ln \left[1 + \delta(A + \beta - 1) \right] + \delta(\beta - 1)\hat{g} \\
&\quad + \delta(\beta - 1) \ln \left[\frac{\delta(A + \beta - 1)}{1 + \delta(A + \beta - 1)} \right]
\end{aligned}
$$

This gives us A as

$$
A = \frac{1 + \delta(\beta - 1)}{1 - \delta}
$$

and, by using the second equation, we can easily obtain an explicit (but extremely messy) expression for B. We can now summarize the results.

Proposition 8.2 *With assumptions as above, the equilibrium consumption strategy is given by*

$$
\hat{c}(x) = \frac{1 - \delta}{1 - \delta + \delta\beta} x.
$$

The equilibrium weight on the risky asset is constant and given by

$$
\hat{w} = w^*
$$

where w^* is the maximum point of the function g defined above. Furthermore, the equilibrium value function has the form

$$
V(x) = A \ln x + B,
$$

where

$$
A = \frac{1 + \delta(\beta - 1)}{1 - \delta}
$$

and B can be computed from the equation above.

Remark 8.2 *We see that $\frac{1-\delta}{1-\delta+\delta\beta} < 1$ implying that $\hat{c}(x) < x$ so we do indeed have an interior optimum for c.*

8.3 Two equivalent standard problems

In this section we comment on the relation of the problem in Section 8.2 to standard problems with exponential discounting.

From Proposition 8.2 we note that it we set $\beta = 1$ then we have a standard problem with exponential discounting, and the optimal consumption is given by

$$
\hat{c}(x) = (1 - \delta)x.
$$

For the general case when $\beta \neq 1$, we see that if we define γ by

$$\gamma = \frac{\delta\beta}{1 - \delta + \delta\beta} \tag{44}$$

then we can write the equilibrium consumption in Proposition 8.2 as

$$\hat{c}(x) = (1 - \gamma)x,$$

which is the optimal rule for a model with exponential discounting. We have thus proved the following result which has the same structure as a corresponding result in [1].

Proposition 8.3 *The equilibrium consumption rule in Proposition 8.2 coincides with the optimal consumption rule for a standard problem with exponential discounting of the form*

$$E_{n,x}\left[\sum_{k=n}^{\infty} \gamma^{k-n} \ln c_k\right]$$

where γ is defined by (44).

Note that the equivalent standard problem defined in Proposition 8.3 is **not** the equivalent standard problem defined by Proposition 4.1. To identify the equivalent problem of Proposition 4.1 it is easiest to recall (42)

$$V(x) = \sup_{w \in R, c \leq x} \left\{ \delta E_x \left[V(X_1^u) \right] + \ln c + \delta(\beta - 1) E_x \left[\ln \left(\hat{c}(X_1^u) \right) \right] \right\}$$

From this we see that if we define the function $U(x, c, w)$ by

$$U(x, c, w) = \ln c + \delta(\beta - 1) E_x \left[\ln \left(\hat{c}(X_1^u) \right) \right]$$

then the extended bellman equation takes the form

$$V(x) = \sup_{w \in R, c \leq x} \left\{ \delta E_x \left[V(X_1^u) \right] + U(x, c, w) \right\}$$

which is the Bellman equation for a standard problem. Using Proposition 8.2 and (43) we thus have the following result.

Proposition 8.4 *Define the function $U(x, c, w)$ by*

$$U(x, c, w) = \ln c + \ln D + \ln(x - c) + g(w)$$

with D and g defined as in Section 8.2. Then the equilibrium value function and the equilibrium control coincides with the optimal value function and the optimal control for the standard problem to maximize

$$E_x\left[\sum_{k=0}^{\infty} \delta^k U(X_k, c_k, w_k)\right]$$

We see that in Proposition 8.3 we keep the local utility function $\ln c$ but change the discount factor from δ to γ, whereas in Proposition 8.4 we keep the discount factor δ but change the local utility function. As was noted in Section 4, the utility function U in Proposition 8.4 is quite strange and does not seem to allow for an easy economic interpretation.

9 Further examples

In this section we give some further applications of the abstract theory. Some of these examples are known from the literature and some are new. The main point of the examples is to illustrate how to handle the abstract theory.

9.1 Mean variance portfolios

In this example we study a multi-period version of the classical mean variance problem. A continuous time version of this problem is studied in some detail in [2], where the authors use a different methodology than in the present paper, and in [6] the mean variance results have recently been massively extended to a general semi martingale framework. We thus make no claim of originality, apart from the fact that we use a general methodology.

We consider a financial market consisting of a risky asset with price process S and a standard risk free bank account with price process B. Recalling that for any process X the process ΔX is defined by $\Delta X_n = X_n - X_{n-1}$, we assume that the dynamics of S and B are given by

$$\begin{aligned} \Delta S_{n+1} &= S_n Y_{n+1}, \\ \Delta B_{n+1} &= B_n r, \end{aligned}$$

where the random returns $\{Y_n\}_{n=1}^T$ is a sequence of i.i.d. random variables, and the short rate r is assumed to be constant.

We denote by u_n the dollar amount invested in the risky asset at time n, and by X_n we denote the dollar value of the portfolio at time n. For a self financing portfolio with no consumption the dynamics of X are easily seen to be given by the expression

$$\Delta X_{n+1} = rX_n + u_n (Y_{n+1} - r),$$

so, using the notation $Z_n = Y_n - r$, and $R = 1 + r$, we have

$$X_{n+1} = RX_n + u_n Z_{n+1}.$$

We have no constraints on the control u and the reward functional is given by the standard mean variance criterion

$$J_n(x, \mathbf{u}) = E_{n,x}[X_T^{\mathbf{u}}] - \frac{\gamma}{2} Var_{n,x}(X_T^{\mathbf{u}})$$

Recalling that $Var(X) = E[X^2] - (E[X])^2$, we see that, in the notation of Theorem 3.1, we have

$$F(y, x) = x - \frac{\gamma}{2}x^2, \quad G(y, x) = \frac{\gamma}{2}x^2.$$

We can now apply Theorem 3.1. Since $F(y, x) = F(x)$ we have no need of the f_n sequence (the f_n terms in the Bellman equation cancel), so the extended

Bellman equation becomes

$$\begin{aligned}\sup_{u\in R}\left\{(\mathbf{A}^u V)_n(x) - \mathbf{A}^u(G\diamond g)_n(x) + \mathbf{H}_g^u G_n(x)\right\} &= 0,\\ V_T(x) &= x,\end{aligned}$$

Since $G(y,x) = G(x) = \frac{\gamma}{2}x^2$ we have $(G\diamond g)_n(x) = G(g_n(x)) = \frac{\gamma}{2}g_n^2(x)$, and after some trivial calculations the extended Bellman equation reduces to the equation

$$\begin{aligned}\sup_{u\in R}\left\{(\mathbf{A}^u V)_n(x) + \frac{\gamma}{2}(\mathbf{P}^u g_n(x))^2 - \frac{\gamma}{2}\mathbf{P}^u g_n^2(x)\right\} &= 0,\\ V_T(x) &= x,\end{aligned}$$

or, in more concrete terms,

$$\begin{aligned}V_n(x) &= \sup_{u\in R}\left\{E_{n,x}\left[V_{n+1}(X_{n+1}^u)\right] + \frac{\gamma}{2}\left(E_{n,x}\left[g_{n+1}(X_{n+1}^u)\right]\right)^2\right.\\ &- \left.\frac{\gamma}{2}E_{n,x}\left[g_{n+1}^2(X_{n+1}^u)\right]\right\},\\ V_T(x) &= x,\end{aligned}$$

where the recursion for g is given by

$$\begin{aligned}g_n(x) &= E_{n,x}\left[g_{n+1}(X_{n+1}^{\hat{\mathbf{u}}})\right]\\ g_T(x) &= x\end{aligned}$$

It is now natural to make the Ansatz (trial solution)

$$\begin{aligned}V_n(x) &= A_n x + B_n,\\ g_n(x) &= a_n x + b_n,\end{aligned}$$

and try to derive recursive equations for A, B, a, and b. Using this Ansatz, as well as the dynamics for X, the extended Bellman equation reduces considerably, and we obtain

$$A_n x + B_n x = A_{n+1}Rx + B_{n+1} + \sup_u\left\{-\frac{\gamma}{2}a_{n+1}^2\sigma^2u^2 + A_{n+1}\mu u\right\}$$

where we have used the notation

$$\begin{aligned}\mu &= E[Z],\\ \sigma^2 &= Var[Z].\end{aligned}$$

We can now easily determine the equilibrium control as

$$\hat{\mathbf{u}}_n(x) = \frac{A_{n+1}\mu}{\gamma a_{n+1}^2\sigma^2}$$

and inserting this into the equation above we obtain

$$A_n x + B_n = A_{n+1} R x + B_{n+1} + \frac{1}{2} \frac{A_{n+1}^2 \mu^2}{\gamma a_{n+1}^2 \sigma^2},$$

which, after identifying coefficients, gives us the recursions

$$\begin{aligned} A_n &= R A_{n+1}, \\ B_n &= B_{n+1} + \frac{1}{2} \frac{A_{n+1}^2 \mu^2}{\gamma a_{n+1}^2 \sigma^2}, \\ A_T &= 1, \\ B_T &= 0. \end{aligned}$$

We still need to determine the a_n sequence and to this end we plug the Ansatz $g_n(x) = a_n x + b_n$, and the previously derived expression for $\hat{\mathbf{u}}$, into the recursion

$$\begin{aligned} g_n(x) &= E_{n,x}\left[g_{n+1}(X_{n+1}^{\hat{\mathbf{u}}})\right] \\ g_T(x) &= x. \end{aligned}$$

We thus obtain

$$a_n x + b_n = a_{n+1} R x + b_{n+1} + \frac{A_{n+1} \mu^2}{\gamma a_{n+1}^2 \sigma^2}$$

This gives us

$$\begin{aligned} a_n &= a_{n+1} R, \\ b_n &= b_{n+1} + \frac{A_{n+1} \mu^2}{\gamma a_{n+1}^2 \sigma^2}, \\ a_T &= 1, \\ b_T &= 0. \end{aligned}$$

These equations are easy to solve as

$$\begin{aligned} A_n &= R^{T-n}, \\ B_n &= (T-n) \frac{\mu^2}{2\gamma \sigma^2}, \\ a_n &= R^{T-n}, \\ b_n &= \frac{\mu^2}{\sigma^2} \sum_{k=n+1}^{T} R^{-(T-k)} \end{aligned}$$

and we have the final result.

Proposition 9.1 *For the mean variance problem above, we have*

$$V_n(x) = R^{T-n} x + (T-n) \frac{\mu^2}{2\gamma \sigma^2},$$

and the equilibrium control is given by

$$\hat{\mathbf{u}}_n(x) = \frac{\mu}{\gamma \sigma^2} R^{-(T-n-1)}.$$

9.2 Mean variance portfolios with state dependent risk aversion

Going back to the mean variance portfolio of the previous section we see that the equilibrium control is of the form

$$\hat{\mathbf{u}}_n(x) = c_n,$$

where c is a deterministic function of time. We also recall that the economic interpretation of u_n is that u_n is the **dollar amount** invested at time n. This implies in particular that for the equilibrium control of the previous section, dollar amount invested in the risky asset does **not** depend on the level of current wealth x.

The problem with this result is that it is economically unreasonable, since it implies that you will invest the same number of **dollars** in the stock if your wealth is 100 dollars as you would if your wealth is 100,000,000 dollars. This is of course well known fact concerning the mean-variance and the exponential utility functions.

The deeper reason for this anomaly is the fact that the risk aversion parameter γ is assumed to be a constant, which is clearly unreasonable. A person's risk preference certainly depends on how wealthy she is, and hence the obvious implication is that we should explicitly **allow γ to depend on current wealth**.

The goal of the present section is to investigate if we can obtain a more economically reasonable result if we introduce a state dependent risk aversion in the mean variance problem. A simple dimension analysis of the utility function tells us that the risk parameter γ should have the dimension $(dollar)^{-1}$, so the obvious choice is to study a state dependent γ of the form $\gamma(x) = \gamma/x$.

We thus consider a reward functional of the form

$$E_{t,x}\left[X_T\right] - \frac{\gamma}{x} Var_{t,x}\left[X_T\right],$$

where the risk parameter $\frac{\gamma}{x}$ now depends also on current wealth x. We note in passing that this specification has the pleasing property that the risk aversion is decreasing in wealth. In [5] we have studied this problem in a continuous time setting. We thus have the reward functional above and the wealth dynamics

$$X_{n+1} = RX_n + u_n Z_{n+1}$$

from the previous section. In terms of Theorem 3.1 we have

$$F(x,y) = y - \frac{\gamma}{x} y^2, \quad G(x,y) = \frac{\gamma}{x} y^2,$$

and the extended Bellman equation has the form

$$\begin{aligned} \sup_{u \in \mathcal{U}} \{ \left(\mathbf{A}^u V\right)_n(x) - \left(\mathbf{A}^u f\right)_n(x,x) + \left(\mathbf{A}^u f^x\right)_n(x) \\ - \mathbf{A}^u \left(G \diamond g\right)_n(x) + \mathbf{H}_g^u G_n(x) \} &= 0, \\ V_T(x) = x, \end{aligned}$$

The function sequences $f_n(x,y)$ and $g_n(x)$ are defined by

$$\begin{aligned} f_n(x,y) &= E_{n,x}\left[X_T^{\hat{\mathbf{u}}}\right] - \frac{\gamma}{y}E_{n,x}\left[\left(X_T^{\hat{\mathbf{u}}}\right)^2\right], \\ g_n(x) &= E_{n,x}\left[X_T^{\hat{\mathbf{u}}}\right]. \end{aligned}$$

We conjecture an equilibrium control of the form $u_n(x) = c_{n+1}(x)$, and under this conjecture the following Ansatz is natural.

$$\begin{aligned} V_n(x) &= A_n x, \\ f_n(x,y) &= \alpha_n x - \beta_n \frac{\gamma}{y} x^2, \\ g_n(x) &= \alpha_n x. \end{aligned}$$

Since we obviously have

$$V_n(x) = f_n(x,x) - \frac{\gamma}{x} g_n^2(x),$$

we must have the relation

$$A_n = \alpha_n - \gamma\beta_n + \gamma\alpha_n^2.$$

For future use we note that

$$\begin{aligned} E_{n,x}\left[X_{n+1}^u\right] &= Rx + \mu u, \\ E_{n,x}\left[\left(X_{n+1}^u\right)^2\right] &= R^2x^2 + 2Rx\mu u + u^2\sigma^2. \end{aligned}$$

Using this, the Ansatz, and the relation $A_n = \alpha_n - \gamma\beta_n + \gamma\alpha_n^2$ we obtain

$$\begin{aligned} \left(\mathbf{A}^u f\right)_n(x,x) &= E_{n,x}\left[f_{n+1}\left(X_{n+1}^u, X_{n+1}^u\right)\right] - f_n(x,x) \\ &= \left(\alpha_{n+1} - \gamma\beta_{n+1}\right)(Rx + \mu u) - f_n(x,x), \\ \left(\mathbf{A}^u f^x\right)_n(x) &= E_{n,x}\left[f_{n+1}\left(X_{n+1}^u, x\right)\right] - f_n(x,x) \\ &= \alpha_{n+1}(Rx + \mu u) - \gamma\beta_{n+1}x^{-1}\left(R^2x^2 + 2Rx\mu u + u^2\sigma^2\right) - f_n(x,x) \end{aligned}$$

We also have

$$\begin{aligned} \left(G \diamond g_n\right)(x) &= \frac{\gamma}{x} g_n^2(x), \\ \mathbf{A}^u\left(G \diamond g\right)_n(x) &= E_{n,x}\left[\left(G \diamond g\right)_{n+1}\left(X_{n+1}^u\right)\right] - \left(G \diamond g_n\right)(x) \\ &= \alpha_{n+1}^2\gamma(Rx + \mu u) - \left(G \diamond g_n\right)(x), \\ \mathbf{H}_g^u G_n(x) &= G\left(x, E_{n,x}\left[g_{n+1}(X_{n+1}^u)\right]\right) - G(x, g_n(x)) \\ &= \gamma x^{-1}\alpha_{n+1}^2(Rx + \mu u)^2 - G(x, g_n(x)). \end{aligned}$$

After some simplifications we can then write the extended Bellman equation as

$$\begin{aligned} A_n x &= \sup_{u \in \mathcal{U}}\left\{\alpha_{n+1}(Rx + \mu u) - \beta_{n+1}\gamma x^{-1}\left(R^2x^2 + 2Rx\mu u + u^2\sigma^2\right)\right. \\ &+ \left.\alpha_{n+1}^2\gamma x^{-1}\left(R^2x^2 + 2Rx\mu u + u^2\mu^2\right)\right\}. \end{aligned}$$

The optimal u is now easily calculated as

$$\widehat{\mathbf{u}}_n(x) = c_{n+1}x$$

where the c_n sequence is given by

$$c_n = \frac{1}{2} \cdot \frac{\alpha_n\mu - 2\beta_n\gamma R\mu + 2\alpha_n\gamma R\mu}{\beta_n\gamma\sigma^2 - \alpha_n^2\gamma\mu^2}.$$

To get recursions for the α and β sequences we use the recursions

$$\begin{aligned} f_n(x,y) &= E_{n,x}\left[f_{n+1}\left(X_T^{\hat{\mathbf{u}}}, y\right)\right], \quad f_T(x,y) = x - \frac{\gamma}{y}x^2, \\ g_n(x) &= E_{n,x}\left[g_{n+1}\left(X_T^{\hat{\mathbf{u}}}\right)\right], \quad g_T(x) = x. \end{aligned}$$

Inserting $\hat{u}_n(x) = c_{n+1}x$ into these recursions and identifying coefficients for the x and x^2 terms gives us

$$\begin{aligned} \alpha_n &= \alpha_{n+1}\left(R + c_{n+1}\mu\right), \\ \alpha_T &= 1, \\ \beta_n &= \beta_{n+1}\left(R^2 + 2R\mu c_{n+1} + \sigma^2 c_{n+1}^2\right), \\ \beta_T &= 1. \end{aligned}$$

Since c_{n+1} is determined by α_{n+1} and β_{n+1}, the recursions above will determine the α and β sequences.

It only remains to check that the extended Bellman equation is satisfied. Plugging $\hat{u}_n(x) = c_{n+1}x$ into the Bellman equation and simplifying, will finally leave us with the formula

$$\alpha_n^2 = \alpha_{n+1}^2\left(R^2 + 2R\mu c_{n+1} + \mu^2 c_{n+1}^2\right)$$

which is trivially satisfied because of the recursion for α above. We ahve thus proved the following.

Proposition 9.2 *For the mean variance problem with state dependent risk aversion of the form $\gamma(x) = \gamma/x$, we have*

$$\begin{aligned} V_n(x) &= \left(\alpha_n - \gamma\beta_n + \gamma\alpha_n^2\right)x, \\ \widehat{\mathbf{u}}_n(x) &= c_{n+1}x \end{aligned}$$

where the α, β, and c sequences are recursively defined above.

9.3 A time inconsistent linear quadratic regulator

In this engineering type problem the reward functional, which in this case is to be minimized, is given by

$$J_n(x,\mathbf{u}) = E_{n,x}\left[\sum_n^{T-1} u_k^2\right] + \left(E_{n,x}\left[X_T\right]\right)^2,$$

with scalar dynamics

$$X_{n+1} = aX_n + bu_n + \sigma W_{n+1}.$$

The scalars a, b and σ are assumed to be known, and the sequence of random variables $\{W_n;\ n = 1, \dots, T\}$ is i.i.d. with distribution $N[0,1]$. There is no restriction on the control.

For this example we can use Theorem 3.2 with

$$C_{n,k}(x, y, u) = u^2, \quad F_n(x, y) \equiv 0, \quad G_n(x, y) = y^2.$$

and the extended Bellman equation has the form

$$\begin{aligned} \inf_{u \in R} \left\{ \mathbf{A}^u V_n(x) + u^2 - \mathbf{A}^u (G \circ g)_n (x) + \mathbf{H}_g^u G_n(x) \right\} &= 0, \\ V_T(x) &= x^2. \end{aligned}$$

where g is determined by the recursion

$$\begin{aligned} \mathbf{A}^{\hat{\mathbf{u}}} g_n(x) &= 0, \quad n = 0, \dots, T-1, \\ g_T(x) &= x, \end{aligned}$$

and we have

$$\begin{aligned} \mathbf{A}^u (G \circ g)_n (x) &= E_{n,x}\left[g^2 \left(X_{n+1}^u \right) \right] - g^2(x), \\ \mathbf{H}_g^u G_n(x) &= \left\{ E_{n,x}\left[g\left(X_{n+1}^u \right) \right] \right\}^2 - g^2(x). \end{aligned}$$

The recursion for V is thus

$$V_n(x) = \inf_{u \in R} \left\{ E_{n,x}\left[V_{n+1}\left(X_{n+1}^u \right) \right] + u^2 + E_{n,x}^2\left[g\left(X_{n+1}^u \right) \right] - E_{n,x}\left[g^2 \left(X_{n+1}^u \right) \right] \right\}.$$

We make the natural Ansatz

$$\begin{aligned} V_n(x) &= A_n x^2 + C_n, \\ g_n(x) &= \alpha_n x. \end{aligned}$$

To simplify the calculations we now note that

$$E_{n,x}^2\left[g\left(X_{n+1}^u \right) \right] - E_{n,x}\left[g^2 \left(X_{n+1}^u \right) \right] = -Var_{n,x}\left[g\left(X_{n+1}^u \right) \right].$$

Using the Ansatz and the X-dynamics we then have

$$E_{n,x}^2\left[g\left(X_{n+1}^u \right) \right] - E_{n,x}\left[g^2 \left(X_{n+1}^u \right) \right] = -\alpha_{n+1}^2 \sigma^2,$$

so the extended Bellman equation takes the form

$$A_n x^2 + C_n = \sup_u \left\{ A_{n+1} \left(ax + bu + \sigma W_{n+1} \right)^2 + u^2 - \alpha_{n+1}^2 \sigma^2 \right\}$$

The equilibrium control is thus given by

$$\hat{u}_n = -ab\frac{A_{n+1}}{A_{n+1}^2 + 1} \cdot x$$

Plugging this into the recursion we obtain the following recursion for A.

$$\begin{aligned} A_n &= {A_{n+1}}^2 + \frac{A_{n+1}^2 a^2 b^4}{\left(A_{n+1}^2 + 1\right)^2} - 2\frac{A_{n+1} a^2 b^2}{A_{n+1}^2 + 1}, \\ A_T &= 1. \end{aligned}$$

The C sequence is determined by

$$C_n = \left\{A_{n+1} - \alpha_{n+1}^2\right\}\sigma^2.$$

The α sequence is determined by the recursion for g, and we obtain

$$\begin{aligned} \alpha_n &= \alpha_{n+1} a \left\{1 - \frac{A_{n+1} b^2}{A_{n+1}^2 b + 1}\right\}, \\ \alpha_T &= 1. \end{aligned}$$

We have thus proved the following

Proposition 9.3 *The equilibrium value function and the equilibrium control are given by*

$$\begin{aligned} V_n(x) &= A_n x^2 + C_n, \\ \hat{u}_n(x) &= -ab\frac{A_{n+1}}{A_{n+1}^2 + 1} \cdot x \end{aligned}$$

where the A, C and α sequences are given by the recursions above.

Going on to the equivalent standard problem, the extended Bellman equation was seen to have the form

$$\begin{aligned} V_n(x) &= \inf_{u \in R}\left\{E_{n,x}\left[V_{n+1}\left(X_{n+1}^u\right)\right] + u^2 - \alpha_{n+1}\sigma^2\right\}, \\ V_T(x) &= x^2, \end{aligned}$$

so the corresponding equivalent time consistent problem of Section 4 is to minimize the functional

$$E_{n,x}\left[\sum_n^{T-1}\left(u_k^2 - \alpha_{n+1}\sigma^2\right)\right] + E_{n,x}\left[X_T^2\right],$$

The optimal control for this problem is of course the same as for the problem to minimize

$$E_{n,x}\left[\sum_n^{T-1} u_k^2\right] + E_{n,x}\left[X_T^2\right],$$

which is a standard linear quadratic regulator problem.

9.4 Another time inconsistent linear quadratic regulator

We now study another small variation of a time inconsistent linear quadratic regulator, where the reward functional (again to be minimized) is given by

$$J_n(x,\mathbf{u}) = E_{n,x}\left[\sum_{n}^{T-1} u_k^2\right] + E_{n,x}\left[(X_T - x)^2\right],$$

with scalar dynamics

$$X_{n+1} = aX_n + bu_n + \sigma W_{n+1}.$$

In this example we thus want to stay close to the current state x while penalizing the control energy. Since the current state $x = X_t$ is dynamic, this leads again to time inconsistency.

For this problem we have the extended Bellman equation

$$\begin{aligned}
\inf_{u \in R}\left\{\mathbf{A}^u V_n(x) + u^2 - \mathbf{A}^u f_n(x,x) + \mathbf{A}^u f_n^x(x)\right\} &= 0, \\
V_T(x) &= 0,
\end{aligned}$$

where

$$\begin{aligned}
f_n^y(x) &= f_n(x,y), \\
f_n(x,y) &= E_{n,x}\left[\left(X_T^{\hat{\mathbf{u}}} - y\right)^2\right], \\
f_T(x,y) &= (x-y)^2, \\
\mathbf{A}^u f_n(x,x) &= E_{n,x}\left[f_{n+1}(X_{n+1}^u, X_{n+1}^u)\right] - f_{n+1}(x,x), \\
\mathbf{A}^u f_n^x(x) &= E_{n,x}\left[f_{n+1}(X_{n+1}^u, x\right] - f_{n+1}(x,x).
\end{aligned}$$

We conjecture that the equilibrium law $\hat{\mathbf{u}}$ is linear in x. Given this conjecture, and the expressions above, the following Ansatz is more or less obvious.

$$\begin{aligned}
V_n(x) &= A_n x^2 + C_n, \\
f_n(x,y) &= \alpha_n x^2 + y^2 + 2\beta_n xy + \delta_n.
\end{aligned}$$

An easy calculation gives us

$$\begin{aligned}
E_{n,x}\left[X_{n+1}^u\right] &= ax + bu, \\
E_{n,x}\left[\left(X_{n+1}^u\right)^2\right] &= a^2x^2 + b^2u^2 + 2abxu + \sigma^2,
\end{aligned}$$

so, using the Ansatz, the extended Bellman equation takes the form

$$\begin{aligned}
A_n x^2 + C_n &= \inf_{u \in R}\left\{u^2 + D_{n+1}(a^2x^2 + b^2u^2 + 2abxu + \sigma^2) + \right. \\
&+ \left. x^2 + 2\beta_{n+1}(ax^2 + bxu) + C_{n+1}\right\}
\end{aligned}$$

where

$$D_n = A_n - 1 - 2\beta_n.$$

The candidate equilibrium control is thus given by

$$\hat{\mathbf{u}}_n(x) = -H_{n+1}x,$$

where

$$H_n = \frac{abD_n + b\beta_n}{bD_n + 1}$$

Plugging the expression for V into the Bellman equation, and identifying coefficients gives us the following recursions for A and C.

$$\begin{aligned}
A_n &= H_{n+1}^2\left(1+b^2D_{n+1}\right)+\left(a^2-2abH_{n+1}\right)D_{n+1}+1+2\beta_{n+1}-b\beta_{n+1}H_{n+1},\\
A_T &= 0,\\
C_n &= C_{n+1}+\sigma^2D_{n+1},\\
C_T &= 0.
\end{aligned}$$

Finally, the recursion for $f_n(x,y)$ gives us the following recursions for α and β

$$\begin{aligned}
\alpha_n &= \alpha_{n+1}\left(a-bH_{n+1}\right)^2, \quad \alpha_T = 1,\\
\beta_n &= \beta_{n+1}\left(a-bH_{n+1}\right), \quad \beta_T = 1.
\end{aligned}$$

10 Conclusion and future research

In this paper we have presented a fairly general class of time inconsistent stochastic control problems. Using a game theoretic perspective we have derived a system of equations for the determination of the subgame perfect Nash equilibrium control, as well as for the corresponding equilibrium value function. The system is an extension of the standard dynamic programming equation for time consistent problems, and we have studied several concrete problems in some detail. We have also shown that for every time inconsistent problem there is an corresponding time consistent problem, which is equivalent in the sense that the optimal control and the optimal value function for the standard problem coincide with the equilibrium control and the equilibrium value function for the time inconsistent problem.

The corresponding continuous time theory is developed in the second half of the working paper [3] and will shortly appear separately as [4].

Some obvious open questions are the following.

- Prove existence adn uniqueness for a reasonably general class of problems.
- Study (lack of) uniqueness for problems with infinite horizon.
- Does there exist an efficient martingale formulation of the theory?
- Is it possible to develop a convex duality theory for time inconsistent problems?
- Can we extend the theory from time inconsistent control problems, to time inconsistent n-person games?

References

[1] Barro, R. Ramsay meets Laibson in the neoclassical growth model. *The Quarterly Journal of Economics 114* (1999), 1125–1152.

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

[3] Björk, T., and Murgoci. A general theory of Markovian time inconsistent stochastic control problems. Working paper $http : //www.hhs.se/se/Search/Person/Pages/Person.aspx?PersonID = 37$, (2009).

[4] Björk, T., and Murgoci. A theory of Markovian time inconsistent stochastic control in continuous time. Working paper, Stockholm School of Economics, (2010).

[5] Björk, T., and Murgoci., and Zhou, X.. Mean–Variance Portfolio Optimization with State Dependent Risk Aversion. Forthcoming in *Mathematical Finance*, (2011).

[6] Czichowsky, C. Time-consistent mean-variance portfolio selection in discrete and continuous time, 2012. forthcoming *Finance and Stochastics.*

[7] Ekeland, I., and Lazrak, A. Being serious about non-commitment: subgame perfect equilibrium in continuous time. *arXiv:math/0604264v1 [math.OC]* (2008), 533–537.

[8] Ekeland, I., and Privu, T. Investment and consumption without commitment. *Mathematics and Financial Economics 2* (2008), 57 – 86

[9] Goldman, S. Consistent plans. *Review of Economic Studies 47* (1980), 533–537.

[10] Harris, C., and Laibson, D. Dynamic choices of hyperbolic consumers. *Econometrica 69* (2001), 935–957.

[11] Krusell, P., and Smith, A. Consumption and savings decisions with quasi-geometric discounting. *Econometrica 71* (2003), 366–375.

[12] Kryger, E., and Steffensen, M. Some solvable portfolio problems with quadratic and collective objectives. Working paper $http : //papers.ssrn.com/sol3/papers.cfm?abstract\ id = 1577265$, (2010).

[13] Luttmer, E., and Mariotti, T. Subjective discounting in an exchange economy. *Journal of Political Economy 111* (2003), 959–989.

[14] Peleg, B., and Menahem, E. On the existence of a consistent course of action when tastes are changing. *Review of Economic Studies 40* (1973), 391–401.

[15] Pollak, R. Consistent planning. *Review of Economic Studies 35* (1968), 185–199.

[16] Strotz, R. Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies 23* (1955), 165–180.

[17] Vieille, N., and Weibull, J. Multiple solutions under quasi-exponential discounting. *Economic Theory 39* (2009), 513–526.