A General Theory of Markovian Time Inconsistent Stochastic Control Problems *

Tomas Björk,
Department of Finance,
Stockholm School of Economics,
e-mail: tomas.bjork@hhs.se

Agatha Murgoci,
Department of Finance,
Copenhagen Business School,
e-mail: avm.fi@cbs.dk

November 15, 2010

First version July 13, 2008

Abstract

We develop a theory for 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 Nash subgame perfect equilibrium points. For a general controlled Markov process and a fairly general objective functional we derive an extension of the standard Hamilton-Jacobi-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. All known examples of time inconsistency 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. We also study some concrete examples.

Key words: Time consistency, time inconsistent control, dynamic programming, time inconsistency, stochastic control, hyperbolic discounting, mean-variance, Bellman equation, Hamilton-Jacobi-Bellman.

*The authors are greatly indebted to Ivar Ekeland, Ali Lazrak, Traian Pirvu, Suleyman Basak, Mogens Steffensen, and Eric Böse-Wolf for very helpful comments.

Contents

1 Introduction

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

$$E\left[\int_0^T C\left(s, X_s, u_s\right) ds + F(X_T)\right].$$

where X is some controlled Markov process, u_s is the control applied at time s, and F, C are given functions. A typical example is when X is a controlled scalar SDE of the form

$$dX_t = \mu(X_t, u_t)dt + \sigma(X_t, u_t)dW_t,$$

with some initial condition $X_0 = x_0$. Later on in the paper we will allow for more general dynamics than those of an SDE, but in this informal section we restrict ourselves for simplicity to the SDE case.

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). We restrict ourselves to **control laws**, i.e., the control at time s, given that $X_s = y$, is of the form $\mathbf{u}(s, y)$ where the control law $\mathbf{u}$ is a deterministic function. We then embed the problem above in a family of problems indexed by the initial point. More precisely we consider, for every (t, x), the problem $\mathcal{P}_{t,x}$ of maximizing

$$E_{t,x}\left[\int_t^T C\left(s, X_s, u_s\right) ds + F(X_T)\right].$$

given the initial condition $X_t = x$. Denoting the optimal control law for $\mathcal{P}_{t,x}$ by $\mathbf{u}_{tx}(s, X_s)$ and the corresponding optimal value function by $V(t, x)$ we see that the original problem corresponds to the problem $\mathcal{P}_{0,x_0}$.

We note that *ex ante* the optimal control law $\mathbf{u}_{tx}(s, X_s)$ for the problem $\mathcal{P}_{t,x}$ must be indexed by the initial point (t, x). However, 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 full time interval $[0, T]$, then it is also optimal for any subinterval $[t, T]$. Given the Bellman principle, it is easy to informally derive the Hamilton-Jacobi-Bellman (henceforth HJB) equation

$$\begin{aligned}
\frac{\partial V}{\partial t}(t,x) + \sup_u \left\{ C(t,x,u) + \mu(t,x,u)\frac{\partial V}{\partial x}(t,x)\frac{1}{2}\sigma^2(t,x,u)\frac{\partial^2 V}{\partial x^2}(t,x) \right\} &= 0, \\
V(T,x) &= F(x),
\end{aligned}$$

for the determination of V. One can (with considerable effort) show rigorously that, given enough regularity, the optimal value function will indeed satisfy the

HJB equation. On can also (quite easily) prove a verification theorem which says that if V is a solution of the HJB equation, then V is indeed the optimal value function for the control problem, and the optimal control law is given by the maximizing u in the HJB equation.

We end this section by listing some important points concerning time consistency.

Remark 1.1 *The main reasons for the time consistency of the problem above are as follows.*

- *The integral term $C(s, X_s, u_s)$ in the problem $\mathcal{P}_{t,x}$ is allowed to depend on s, X_s and u_s. It is* **not** *allowed to depend on the initial point (t, x).*
- *The terminal evaluation term is allowed to be of the form $E_{t,x}[F(X_T)]$, i.e the expected value of a non-linear function of the terminal value X_T. Time consistency is then a relatively simple consequence of the law of iterated expectations. We are* **not** *allowed to have a term of the form $G(E_{t,x}[X_T])$, 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 (t, 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 standard Black-Scholes model for an underlying stock price S, as well as a bank account B with short rate r.

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dB_t &= r B_t dt. \end{aligned}$$

We consider a self financing portfolio based on S and B where u_t is the number of dollars invested in the risky asset S, and c is the consumption rate at time t. Denoting the market value process of this portfolio by X, we have

$$dX_t = [rX_t + (\alpha - r)u_t - c_t]dt + \sigma u_t dW_t,$$

and we now consider three indexed families of optimization problems. In all cases the (naive) objective is to maximize the objective functional $J(t, x, \mathbf{u})$, where (t, x) is the initial point and $\mathbf{u}$ is the control law.

1. Hyperbolic discounting

$$J(t, x, \mathbf{u}) = E_{t,x}\left[\int_t^T \varphi(s-t)h(c_s)dt + \varphi(T-t)F(X_T)\right]$$

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 t enters in the integral (see Remark 1.1). Obviously; if φ is exponential so $\varphi(s-t) = e^{-a(s-t}$ then we can factor out e^{at} and convert the problem into a standard problem with objective functional

$$J(t, x, \mathbf{u}) = E_{t,x}\left[\int_t^T e^{-as}h(c_s)dt + e^{-aT}F(X_T)\right]$$

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

2. Mean variance utility

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

This case is a continuous time 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 $E_{t,x}\left[X_T\right]$, but we penalize the risk by the conditional variance $\frac{\gamma}{2}Var_{t,x}\left(X_T\right)$. This looks innocent enough, but we recall the elementary formula

$$Var[X] = E[X^2] - E^2[X].$$

Now, in a standard time consistent problem we are allowed to have terms like $E_{t,x}\left[F(X_T)\right]$ 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 $\left(E_{t,x}\left[X\right]\right)^2$. 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

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

In this particular example we basically want to maximize log utility of terminal wealth. In a standard problem we would have the objective $E_{t,x}\left[\ln\left(X_T - d\right)\right]$ 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 $X_t = x$ at time t. 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

(t, x) we determine the control law $\hat{\mathbf{u}}$ which maximizes $J(t, x, \mathbf{u})$, then at some later point (s, X_s) the control law $\hat{\mathbf{u}}$ will no longer be optimal for the functional $J(s, X_s, \mathbf{u})$. 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 (s, X_s) the control law $\hat{\mathbf{u}}$ will not be optimal for the functional $J(s, X_s, \mathbf{u})$. 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 t, where player t can be viewed as the future incarnation of ourselves (or rather of our preferences) at time t. Given this point of view, it is natural to look for Nash equilibria for the game, and this is exactly our approach.

In continuous time it is far from trivial to formulate this intuitive idea in precise terms. We will do this in the main text below but a rough picture of the game is as follows.

- We consider a game with one player at each point t in time. This player is referred to as "player t". You may think of player t as a future incarnation of yourself, but conceptually it may be easier to think of the game as being played by a continuum of completely different individuals.
- Depending on t and on the position x in space, player t will choose a control action. This action is denoted by $\mathbf{u}(t, x)$, so the strategy chosen by player t is given by the mapping $x \longmapsto \mathbf{u}(t, x)$.
- A control law $\mathbf{u}$ can thus be viewed as a complete description of the chosen strategies of all players in the game.
- The reward to player t is given by the functional $J(t, x, \mathbf{u})$. We note that in the three examples of the previous section it is clear that $J(t, x, \mathbf{u})$ does not depend on the actions taken by any player s for $s < t$, so in fact J does only depend on the restriction of the control law $\mathbf{u}$ to the time interval

$[t, T]$. It is also clear that this is really a game, since the reward to player t does not only depend on the strategy chosen by himself, but also by the strategies chosen by all players coming after him in time.

We can now loosely define the concept of a "Nash subgame perfect equilibrium point" of the game. This is a control law $\hat{\mathbf{u}}$ satisfying the following condition:

- Choose an arbitrary point t in time.
- Suppose that every player s, for all $s > t$, will use the strategy $\hat{\mathbf{u}}(s, \cdot)$.
- Then the optimal choice for player t is that he/she also uses the strategy $\hat{\mathbf{u}}(t, \cdot)$.

The problem with this "definition" in continuous time is that, for example in a diffusion framework without impulse controls, the individual player t does not really influence the outcome of the game at all. He/she only chooses the control at the single point t, and since this is a time set of Lebesgue measure zero, the control dynamics will not be influenced. For a proper definition we need some sort of limiting argument, which is given in the main text below.

1.4 Previous literature

The game theoretic approach to time inconsistency using Nash equilibrium points as above has a long history starting with [13] where a deterministic Ramsay problem is studied. Further work along this line in continuous and discrete time is provided in [6], [8], [10], [11], and [14].

Recently there has been renewed interest in these problems. In the interesting, and mathematically very advanced, papers [4] and [5], 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. To our knowledge, these papers were the first to provide a precise definition of the equilibrium concept in continuous time. The authors derive, among other things, an extension of the HJB equation to a system of PDEs including an integral term, and they also provide a rigorous verification theorem. They also, in a tour de force, derive an explicit solution for the case when the discounting function is a weighted sum of two exponential discount functions. Our present paper is much inspired by these papers, in particular for the definition of the equilibrium law.

In [1] 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. The authors derive an extension of the Hamilton Jacobi Bellman equation and manages, by a number of very clever ideas, to solve this equation explicitly for the basic problem, and also for the above mentioned extensions. The paper has two limitations: Firstly, from a mathematical perspective it is somewhat heuristic, the equilibrium concept is never given a precise definition, and no verification theorem is provided. Secondly, and more importantly, the methodology depends heavily on the use of a "total variance formula", which in some sense (partially) replaces the iterated expectations formula in a standard problem. This implies that the basic methodology cannot be extended beyond the mean variance case. The paper is extremely thought provoking, and we have benefited greatly from reading it.

The recent working paper [9] uses the theory of the present paper and contains several interesting new applications. In [3] the author undertakes a deep study of the mean variance problem within in a general semi martingale framework.

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 SDEs driven by a Wiener and a point process.

- We consider a functional of the form

$$J(t,x,\mathbf{u}) = E_{t,x}\left[\int_t^T C\left(x, X_s^{\mathbf{u}}, \mathbf{u}(X_s^{\mathbf{u}})\right) ds + F(x, X_T^{\mathbf{u}})\right] + G\left(x, E_{t,x}\left[X_T^{\mathbf{u}}\right]\right).$$

We see that with the choice of functional above, time inconsistency enters at several different points. Firstly we have the appearance of the present state x in the local utility function C, as well as in the functions F and G. As a consequence of this, the utility function changes as the state changes. At time t we have, for example, the utility function $F(X_t, X_T)$ which we want to maximize as a function of X_T, but at a later time s we have the utility function $F(X_s, X_T)$. This obviously leads to time inconsistency. Secondly, in the term $G\left(x, E_{t,x}\left[X_T^{\mathbf{u}}\right]\right)$ we have, even forgetting about the appearance of x, a non linear function G acting on the conditional expectation, again leading to time inconsistency.

Note that, for notational simplicity we have not explicitly included dependence on running time t. This can always be done by letting running time be one component of the state process X, so the setup above also allows for expressions like $F(t, x, X_T^{\mathbf{u}})$ etc, thus allowing (among many other things) for hyperbolic discounting.

This setup is studied in some detail in continuous as well as in discrete time. The discrete time results are parallel to those in continuous time, and our main results in continuous time are as follows.

- We provide a precise definition of the Nash equlibrium concept. (This is done along the lines of [4] and [5]).
- We derive an extension of the standard Hamilton-Jacobi-Bellman equation to a non standard system of equations for the determination of the equilibrium value function V.
- We prove a verification theorem, showing that the solution of the extended HJB system is indeed the equilibrium value function, and that the equilibrium strategy is given by the optimizer in the equation system.
- 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 startegy for the time inconsistent problem.
- We solve some specific test examples.

Our framework and results extends the existing theory considerably. As we noted above, hyperbolic discounting is included as a special case of the theory. The mean variance example from above is of course also included. More precisely it is easy to see that it corresponds to the case when

$$C = 0, \quad F(x, y) = y - \frac{\gamma}{2} y^2, \quad G(x, y) = \frac{\gamma}{2} y^2$$

We thus extend the existing literature by allowing for a considerably more general utility functional, and a completely general Markovian structure. The existence of the associated equivalent standard control problem is to our knowledge a completely new result.

1.6 Structure of the paper

Since the equilibrium concept in continuous time is a very delicate one, we start by studying a discrete time version of our problem in Section 2. In discrete time there are no conceptual problems with the equilibrium concept, but the arguments are sometimes quite delicate, the expressions are rather complicated, and great care has to be taken. It is in fact in this section that the main work is done. In Section 5 we exemplify the theory by studying the special case of non exponential discounting.

In Section 7 we study the continuous time problem by taking formal limits for a discretized problem, and using the results of the Section 2. This leads to an extension of the standard HJB equation to a system of equations with an embedded static optimization problem. The limiting procedure described above is done in an informal manner and it is largely heuristic, so in order to prove that the derived extension of the HJB equation is indeed the correct one we also provide a rigorous proof of a verification theorem. In Section 8 we prove the existence of the associated standard control problem, and in Sections 9-11 we study several examples.

2 Discrete time

Since the theory is conceptually much easier in discrete time than in continuous time, we start by presenting the discrete time version.

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. 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 $\mathbf{u}$, whereas a possible value of the mapping will be denoted without boldface, like, $u \in \mathcal{U}$.

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.

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

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 $f_n(X_n^{\mathbf{u}})$ is then a martingale under the measure induced by $\mathbf{u}$ if and only if the sequence $\{f_n\}$ satisfies the equation*

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

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 $E_{n,x}^{\mathbf{u}}[\cdot]$, 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 $E_{n,x}\left[F(X_k^{\mathbf{u}})\right]$ 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 $\mathbf{u}$, and a fixed time horizon T, we consider the functional

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

Obviously, the functional J depends only on the restriction of the control law $\mathbf{u}$ to the time set $k = n, n+1, \ldots, T-1$.

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!". The complicating factor here 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 $k, k+1, \ldots, T$ (for $k > n$) is not necessarily optimal for the problem $\mathcal{P}_{k,X_k^{\mathbf{u}}}$. There are two reasons for this time inconsistency:

- The shape of the utility functional depends explicitly on the initial position x in space, as can be seen in the appearance of x in the expression $F(x, X_T)$ and similarly for the other terms. In other words, as the X process moves around, our utility function changes, so at time t this part of the utility function will have the form $F(X_t, X_T)$.

- For a standard time consistent control problem we are allowed to have expressions like $E_{n,x}\left[G(X_T)\right]$ 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 $G\left(E_{n,x}\left[X_T^{\mathbf{u}}\right]\right)$, which 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 the present state x in the expression $G\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right)$.

The moral of all this is that we have a family of time inconsistent problems or, alternatively, we have time inconsistent preferences. 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, as outlined above, take a game theoretic approach and we now go on the describe this in some detail.

2.3 The game theoretic formulation

The idea, which appears already in [13], 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. 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 $\mathbf{u}$, 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* $\bar{\mathbf{u}}$ *on the time set* $n, n+1, \ldots, T-1$ *by setting, for any* $y \in \mathcal{X}$*,*

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

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

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

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 all players coming after him will use the control $\hat{\mathbf{u}}$, then it is optimal for player number n also to use $\hat{\mathbf{u}}_n$.

Remark 2.1 *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 induction.*

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.

2.4 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.

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

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 the derivation only for the case when the objective functional does not contain the sum $\sum_{k=n}^{T-1} C\left(x, X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}})\right)$ in (1), and thus has 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 provide the result for the general case in Section 2.4.4. The derivation of this is completely parallel to that of the simplified case.

We start by making the observation that 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 control law $\mathbf{u}$ (restricted to the interval $[n, k-1]$) so for $k > n+1$ we use the notation $X_k^{\mathbf{u}}$.

We now go on to the recursion arguments. From the definition of J we have

$$J_{n+1}(X_{n+1}^u, \mathbf{u}) = E_{n+1}\left[F(X_{n+1}^u, X_T^{\mathbf{u}})\right] + G\left(X_{n+1}^u, E_{n+1}\left[X_T^{\mathbf{u}}\right]\right), \tag{3}$$

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 definitions which will play a central role in the sequel.

Definition 2.3 *For any control law* $\mathbf{u}$*, we define the function sequences* $\{f_n^{\mathbf{u}}\}$ *and* $\{g_n^{\mathbf{u}}\}$*, where* $f_n^{\mathbf{u}}, g_n^{\mathbf{u}} : \mathcal{X} \times \mathcal{X} \to \mathbf{R}$ *by.*

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

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 processes $f_n^{\mathbf{u},y}(X_n^{\mathbf{u}})$ and $g_n^{\mathbf{u}}(X_n^{\mathbf{u}})$ are martingales under the measure generated by $\mathbf{u}$. We thus have the following result.

Lemma 2.1 *For every fixed control law* $\mathbf{u}$ *and every fixed choice of* $y \in \mathcal{X}$*, the function sequence* $\{f_n^{\mathbf{u},y}\}$ *satisifes the recursion*

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

The sequence $\{g_n^{\mathbf{u}}\}$ *satisifes the recursion*

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

Going back to (3) we note that, from the Markovian structure and the definitions above, we have

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

We can now write (3) as

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

Taking expectations gives us

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

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

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

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

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

and 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 these identities into the recursion above, we can now collect the findings so far.

Lemma 2.2 *The value function J satisfies the following recursion.*

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

2.4.2 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 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. 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

$$\sup_{u \in \mathcal{U}} J_n(x, \mathbf{u}) = V_n(x). \tag{4}$$

We now make a small variation of Definition 3.

Definition 2.4 *We define the function sequences $\{f_n\}_{n=0}^T$, and $\{g_n\}_{n=0}^T$, where $f_n : \mathcal{X} \times \mathcal{X} \to \mathbf{R}$, and $g_n : \mathcal{X} \to \mathbf{R}$ by.*

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

We 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 2.2, the basic relation (4) now reads

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

W 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

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

We can thus write the recursion as

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

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

$$\left. - \left(E_{n,x} \left[G \left(X_{n+1}^u, g_{n+1}(X_{n+1}^u) \right) \right] - G \left(x, E_{n,x} \left[g_{n+1}(X_{n+1}^u) \right] \right) \right) \right\} = 0.$$

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 rewrite the third line of the recursion as

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

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

Definition 2.5 *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

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

We now introduce the last piece of new notation.

Definition 2.6 *With notation as above we define the function sequence* $\left\{ \mathbf{H}_g^u G \right\}_k$ *by*

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

Finally, we may state the main result for discrete time models.

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

$$\begin{aligned} \sup_{u \in \mathcal{U}} \{ (\mathbf{A}^u V)_n (x) - (\mathbf{A}^u f)_n (x,x) + (\mathbf{A}^u f^x)_n (x) & \\ - \mathbf{A}^u (G \diamond g)_n (x) + \mathbf{H}_g^u G_n(x) \} &= 0, \quad (5) \\ V_T(x) = F(x,x) + G(x,x), & \quad (6) \end{aligned}$$

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*

$$\begin{aligned} \mathbf{A}^{\hat{\mathbf{u}}} f_n^y(x) &= 0, \quad n = 0, \dots, T-1, \quad (7) \\ f_T^y(x) &= F(y,x), \quad (8) \end{aligned}$$

   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.*

$$\begin{aligned} \mathbf{A}^{\hat{\mathbf{u}}} g_n(x) &= 0, \quad n = 0, \dots, T-1, \quad (9) \\ g_T(x) &= x, \quad (10) \end{aligned}$$

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

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

4. *In the recursions above, the* $\hat{\mathbf{u}}$ *occurring in the expression* $\mathbf{A}^{\hat{\mathbf{u}}}$ *is the equilibrium control law.*

We now have some comments on this result.

- The first point to notice is that we have a **system** of recursion equation (5)-(10) for the simultaneous determination of V, f and g.
- 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 $(\mathbf{A}^u f)_n (x,x) + (\mathbf{A}^u f^x)_n (x)$ in the V-equation (5) cancel, and the system reduces to the standard Bellman equation

$$\begin{aligned} (\mathbf{A}^u V)_n (x) &= 0, \\ V_T(x) &= F(x). \end{aligned}$$

- In order to solve the V-equation (5) 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 (5).
- 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

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

    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 (5) 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.
- In the present **discrete time** setting, the situation is, however, simpler than the fixed point argument above may lead us to believe. In fact; looking closer at the recursions, it turns out that the system for V, f, and g is a formalization of the recursive strategy outlined in Remark 2.1.

2.4.3 Explicit time dependence

In many examples, the objective functional exhibits explicit time dependence, in the sense that the functions F and G depend explicitly on running time n, so J has the form

$$J_n(x, \mathbf{u}) = E_{n,x}\left[F_n(x, X_T^{\mathbf{u}})\right] + G_n\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right). \tag{11}$$

This case is in fact included in the theory developed above by considering the extended state process (n, X_n), where n has trivial dynamics. Since problems of this form are quite common we nevertheless provide the result for easy reference. The novelty for this case is that the function sequence

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

will be replaced by the sequence

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

Proposition 2.2 *For a functional of the form (11), the following hold.*

1. *The equilibrium value function V satisfies the equation.*

$$\begin{aligned} \sup_{u\in\mathcal{U}} \{ (\mathbf{A}^u V)_n(x) - (\mathbf{A}^u f)_{nn}(x,x) + (\mathbf{A}^u f^{nx})_n(x) \\ - \mathbf{A}^u (G \diamond g)_n(x) + \mathbf{H}_g^u G_n(x) \} &= 0, \quad (12) \\ V_T(x) = F_T(x,x) + G_T(x,x), \quad (13) \end{aligned}$$

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

2. *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*

$$\begin{aligned} \mathbf{A}^{\hat{\mathbf{u}}} f_n^{ky}(x) &= 0, \quad n = 0, \ldots, T-1, \quad (14) \\ f_T^{ky}(x) &= F_k(y,x), \quad (15) \end{aligned}$$

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

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

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

$$\begin{aligned} \mathbf{A}^{\hat{\mathbf{u}}} g_n(x) &= 0, \quad n = 0, \ldots, T-1, \quad (16) \\ g_T(x) &= x, \quad (17) \end{aligned}$$

4. *The function sequence $(G \diamond g)$ is defined by*

$$(G \diamond g)_n(x) = G_n(x, g_n(x))$$

5. *The term $\mathbf{H}_g^u G_n(x)$ is defined by*

$$\mathbf{H}_g^u G_n(x) = G_n(x, P^u g_n(x)) - G_n(x, g_n(x))$$

6. *The probabilistic interpretations of f and g are given by*

$$\begin{aligned} f_{nk}(x,y) &= E_{n,x}\left[F_k\left(y, X_T^{\hat{\mathbf{u}}}\right) \right], \quad (18) \\ g_n(x) &= E_{n,x}\left[X_T^{\hat{\mathbf{u}}} \right]. \quad (19) \end{aligned}$$

7. *In the expressions above, the $\hat{\mathbf{u}}$ occurring in the expression $\mathbf{A}^{\hat{\mathbf{u}}}$ is the equilibrium control law.*

Remark 2.2 *Note that for every fixed $k = 0, 1, 2, \ldots, T$ the function $f_{nk}(x,y)$ is defined for* **all** *$n = 0, 1, 2, \ldots, T$, not only for $n \leq k$.*

2.4.4 The general case

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

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{T-1} C_{n,k}\left(x, X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}})\right) + F_n(x, X_T^{\mathbf{u}})\right] + G_n\left(x, E_{n,x}\left[X_T^{\mathbf{u}}\right]\right). \tag{20}$$

The arguments for the C terms in the sum above are very similar to the previous arguments for the F term. It is thus natural to introduce an indexed function sequence defined by

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

where, as usual, $\hat{\mathbf{u}}$ denotes the equilibrium law. We then have the following result.

Theorem 2.2 *Consider a functional of the form (20), and assume that an equilibrium control law $\hat{\mathbf{u}}$ exists. Then the the equilibrium value function V satisfies the equation.*

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

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

$$V_T(x) = F_T(x,x) + G_T(x,x),$$

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*

$$\mathbf{A}^{\hat{\mathbf{u}}} f_n^{ky}(x) = 0, \quad n = 0, \ldots, T-1, \tag{21}$$

$$f_T^{ky}(x) = F_k(y, x), \tag{22}$$

   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.*

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

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

$$\begin{aligned}(\mathbf{A}^{\hat{\mathbf{u}}} c^{k,m,y})_n(x) &= 0, \quad 0 \le n \le m-1,\\ c_m^{k,m,y}(x) &= C_{k,m}(y, x, \hat{\mathbf{u}}_m(x)).\end{aligned}$$

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*

$$f_n^{ky}(x) = E_{n,x}\left[F_k\left(y, X_T^{\hat{\mathbf{u}}}\right)\right], \tag{23}$$

$$g_n(x) = E_{n,x}\left[X_T^{\hat{\mathbf{u}}}\right]. \tag{24}$$

$$c_n^{k,m,y}(x) = E_{n,x}\left[C_{k,m}\left(y, X_m^{\hat{\mathbf{u}}}, \hat{\mathbf{u}}_m\left(X_m^{\hat{\mathbf{u}}}\right)\right)\right] \tag{25}$$

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

2.4.5 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. Let us thus consider a value functional of the form

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{\infty} C_{n,k}\left(x, X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}})\right)\right]. \tag{26}$$

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

Proposition 2.3 *Consider a functional of the form (26), and assume that an equilibrium control law $\hat{\mathbf{u}}$ exists. Then the the corresponding equilibrium value function V satisfies the equation.*

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

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 (25).

For the infinite horizon case, existence and uniqueness of an equilibrium control is highly nontrivial. See Section 4 for more details.

2.4.6 A scaling result

In this section we derive a small scaling result, which is sometimes quite useful. Consider the objective functional (20) above and denote, as usual, the equilibrium control and value function by $\hat{\mathbf{u}}$ and V respectively. Let $\varphi : R \to R$ be a fixed real valued function and consider a new objective functional J^{φ}, defined by,

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

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 2.4 *With notation as above we 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. ∎

3 An equivalent time consistent problem

The object of the present section is to provide a surprising link 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 (23)-(25). We now define the function h by

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

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

$$E_{n,x}\left[\sum_{k=n}^{T} h_k(X_k, u_k) + H(X_T)\right]. \tag{27}$$

We have thus proved the following result.

Proposition 3.1 *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 (27).*

We immediately remark that Proposition 3.1 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.

Furthermore, Proposition 3.1 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 3.1 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".

4 Existence and uniqueness

So far we have not discussed existence and uniqueness of an equilibrium control, and for these issues there is a marked difference between the case $T < \infty$ and the case $T = \infty$.

For the case $T < \infty$, existence and uniqueness is not complicated. Theorem 2.2 does in fact give us a concrete backward recursion for the equilibrium value function, starting at $n = T$. For a well posed problem, i.e. a problem where the supremum is attained in the extended Bellman equation for $n = T-1, T-2, \ldots$ the equilibrium control and the corresponding value function are thus determined recursively. In principle it may of course happen that for some n there is more than one global optimum in the extended Bellman equation, but this is a non-generic situation in the sense that it is not structurally stable. We thus conclude that for the case $T < \infty$ we have generic existence and uniqueness of the equilibrium control and value function.

The case $T = \infty$ is much 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, the argument above for finite T will no longer hold, and we conjecture that we may generically have multiple equilibria. See [14] for a detailed study of the existence of multiple equilibria in a concrete case.

5 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

$$J_n(x, \mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{T-1} \varphi(k-n) H(X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}})) + \varphi(T-n) K(X_T)\right] \quad (28)$$

- 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.

5.1 A general discount function

We see that, in the notation of Theorem 2.2 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

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

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 5.1 *The extended Bellman system for the non exponential discounting problem (28) is given by*

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

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

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

5.2 Infinite horizon

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

$$J_n(x,\mathbf{u}) = E_{n,x}\left[\sum_{k=n}^{\infty} \varphi(k-n)H(X_k^{\mathbf{u}}, \mathbf{u}_k(X_k^{\mathbf{u}}))\right] \tag{29}$$

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 5.2 *For the infinite horizon problem (29) 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

$$\begin{aligned} h^{k+1} &= E_x\left[h^k(X_1^{\hat{\mathbf{u}}})\right], \\ h^0(x) &= H(x, \hat{\mathbf{u}}(x)). \end{aligned}$$

5.3 Quasi-hyperbolic discounting

We now turn to the special case when φ is a "quasi-hyperbolic discounting function". More precisely we assume that φ has the form

$$\varphi(0) = 1, \tag{30}$$
$$\varphi(k) = \beta\delta^k, \quad k = 1, 2, 3, \dots \tag{31}$$

where $\beta \geq 0$ and $0 < \delta < 1$. In this case we have

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

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 5.2 we thus have our main result for the hyperbolic discounting case.

Proposition 5.3 *For the case (30)-(31), of quasi-hyperbolic discounting, the extended Bellman system takes the form*

$$V(x) = \sup_u \left\{\delta E_x\left[V(X_1^u)\right] + H(x,u) + \delta(\beta - 1)E_x\left[H(X_1^u, \hat{\mathbf{u}}(X_1^u))\right]\right\}. \tag{32}$$

Remark 5.1 *We note that if $\beta = 1$ we have the standard exponential discounting $\varphi(k) = \delta^k$ and the extended Bellman equation in Proposition 5.3 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.

5.4 A variational equation

We now specialize further to the case when the X process is governed by a stochastic difference equation.

Assumption 5.1 *We assume that X and u are real valued, and that X has dynamics according to*

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

We furthermore assume that μ and σ are deterministic functions, and that $\{Y_n\}_{n=1}^{\infty}$ is a sequence of i.i.d. random variables.

Our objective is to derive a variational equation, and to this end we note that in the presnt setting we can write the extended Bellman equation (32) as follows where, for brevity, we use the notation $Y = Y_1$.

$$\begin{aligned} V(x) \quad = \quad & \sup_u \left\{ \delta E_x \left[V \left(\mu(x,u) + \sigma(x,u)Y \right) \right] + H(x,u) \right. \\ & \left. + \, \delta(\beta - 1) E_x \left[H(\mu(x,u) + \sigma(x,u)Y, \hat{\mathbf{u}}(\mu(x,u) + \sigma(x,u)Y)) \right] \right\} \end{aligned}$$

If we assume an interior optimum we can easily compute the first order condition for optimum by setting the u derivative equal to zero. The result is as follows.

$$\begin{aligned} & \delta E_x \left[V_x(\widehat{X}_1) \right] \mu_u \left(x, \hat{\mathbf{u}}(x) \right) + \delta E_x \left[V_x(\widehat{X}_1) \cdot Y \right] \sigma_u \left(x, \hat{\mathbf{u}}(x) \right) \\ + \quad & H_u \left(x, \hat{u}(x) \right) \\ + \quad & \delta(\beta - 1) E_x \left[H_x(\widehat{X}_1, \hat{\mathbf{u}}(\widehat{X}_1)) \right] \mu_u \left(x, \hat{\mathbf{u}}(x) \right) \\ + \quad & \delta(\beta - 1) E_x \left[H_x(\widehat{X}_1, \hat{\mathbf{u}}(\widehat{X}_1)) \cdot Y \right] \sigma_u \left(x, \hat{\mathbf{u}}(x) \right) \\ + \quad & \delta(\beta - 1) E_x \left[H_u(\widehat{X}_1, \hat{\mathbf{u}}(\widehat{X}_1)) \cdot \hat{\mathbf{u}}_x(\widehat{X}_1) \right] \mu_u \left(x, \hat{\mathbf{u}}(x) \right) \\ + \quad & \delta(\beta - 1) E_x \left[H_u(\widehat{X}_1, \hat{\mathbf{u}}(\widehat{X}_1)) \cdot \hat{\mathbf{u}}_x(\widehat{X}_1) Y \right] \sigma_u \left(x, \hat{\mathbf{u}}(x) \right) = 0. \end{aligned} \tag{33}$$

Here we have used the notation

$$\widehat{X}_1 = X_1^{\hat{\mathbf{u}}(x)} = \mu(x, \hat{\mathbf{u}}(x)) + \sigma(x, \hat{\mathbf{u}}(x)) Y,$$

and lower case indices, like μ_u and $\hat{\mathbf{u}}_x$, denote partial derivatives.

From the envelope theorem we furthermore obtain

$$\begin{aligned} V_x(x) &= \delta E_x\left[V_x\left(\widehat{X}_1\right)\right]\mu_x(x,\hat{\mathbf{u}}(x)) + \delta E_x\left[V_x\left(\widehat{X}_1\right)\cdot Y\right]\sigma_x(x,\hat{\mathbf{u}}(x)) \\ &+ H_x(x,\hat{\mathbf{u}}(x)) \\ &+ \delta(\beta-1)E_x\left[H_x(\widehat{X}_1,\hat{\mathbf{u}}(\widehat{X}_1))\right]\mu_x\left(x,\hat{\mathbf{u}}(x)\right) \\ &+ \delta(\beta-1)E_x\left[H_x(\widehat{X}_1,\hat{\mathbf{u}}(\widehat{X}_1))\cdot Y\right]\sigma_x\left(x,\hat{\mathbf{u}}(x)\right) \\ &+ \delta(\beta-1)E_x\left[H_u(\widehat{X}_1,\hat{\mathbf{u}}(\widehat{X}_1))\cdot \hat{\mathbf{u}}_x(\widehat{X}_1)\right]\mu_x\left(x,\hat{\mathbf{u}}(x)\right) \\ &+ \delta(\beta-1)E_x\left[H_u(\widehat{X}_1,\hat{\mathbf{u}}(\widehat{X}_1))\cdot \hat{\mathbf{u}}_x(\widehat{X}_1)Y\right]\sigma_x\left(x,\hat{\mathbf{u}}(x)\right) \end{aligned} \tag{34}$$

We now introduce a another crucial assumption.

Assumption 5.2 *We assume that μ and σ are of the form*

$$\mu(x,u)=\mu(x-u), \quad \sigma(x,u)=\sigma(x-u)$$

This is equivalent to saying that

$$\mu_x=-\mu_u, \quad \sigma_x=-\sigma_u,$$

and we may now insert (33) above into (34) to obtain

$$V_x(x)=H_x(x,\hat{\mathbf{u}}(x))+H_u(x,\hat{\mathbf{u}}(x)).$$

Substituting this expression into (34) gives us the following result.

Proposition 5.4 *Given Assumptions 5.1 and 5.2, and assuming smoothness of all entities as well as an interior optimum in the extended Bellman equation, we have the following variational equation.*

$$\begin{aligned} H_u(x,\hat{\mathbf{u}}(x)) &= \delta E_x\left[H_u\left(\widehat{X}_1,\hat{\mathbf{u}}(\widehat{X}_1)\right)\right]\mu'(x-\hat{\mathbf{u}}(x)) \\ &+ \delta E_x\left[H_u\left(\widehat{X}_1,\hat{\mathbf{u}}(\widehat{X}_1)\right)\cdot Y\right]\sigma'(x-\hat{\mathbf{u}}(x)) \\ &+ \delta\beta E_x\left[H_x(\widehat{X}_1,\hat{\mathbf{u}}(\widehat{X}_1))\right]\mu'\left(x-\hat{\mathbf{u}}(x)\right) \\ &+ \delta\beta E_x\left[H_x(\widehat{X}_1,\hat{\mathbf{u}}(\widehat{X}_1))\cdot Y\right]\sigma'\left(x-\hat{\mathbf{u}}(x)\right) \\ &+ \delta(\beta-1)E_x\left[H_u(\widehat{X}_1,\hat{\mathbf{u}}(\widehat{X}_1))\cdot \hat{\mathbf{u}}'(\widehat{X}_1)\right]\mu'\left(x-\hat{\mathbf{u}}(x)\right) \\ &+ \delta(\beta-1)E_x\left[H_u(\widehat{X}_1,\hat{\mathbf{u}}(\widehat{X}_1))\cdot \hat{\mathbf{u}}'(\widehat{X}_1)Y\right]\sigma'\left(x-\hat{\mathbf{u}}(x)\right) \end{aligned} \tag{35}$$

5.5 The Harris and Laibson model

A special case of the setup above is the model studied in [7]. In their model, the X dynamics have the form

$$X_{n+1}=R\left(X_n-c_n\right)+Y_{n+1}.$$

The economic interpretation is that the state variable X denotes cash-on-hand, the control c denotes consumption, and R is the capitalization factor. The reward functional has the form

$$E_x \left[U(c_1) + \beta \sum_{n=1}^{\infty} \delta^n U(c_n) \right]$$

Comparing to the more general setup in Section 5.4 we thus have

$$\begin{aligned} u &= c, \\ \mu(x,c) &= R(x-c), \\ \sigma(x,c) &= 1, \\ H(x,c) &= U(c) \end{aligned}$$

We thus have

$$\mu' = R, \quad \sigma' = 0, \quad H_x = 0, \quad H_u = U'$$

and, inserting this into the variational equation (35), we obtain

$$U'(\hat{c}(x)) = \delta R E_x \left[U'\left(\widehat{c}(\widehat{X}_1) \right) \right] + \delta(\beta - 1) R E_x \left[U'\left(\widehat{c}(\widehat{X}_1) \right) \cdot \widehat{c}'(\widehat{X}_1) \right]$$

which, for the case of an interior optimum, is precisely the variational equation (8) in [7].

6 Application 2: 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 [1], where the authors use a different methodology than in the present paper. We thus make no claim of originality, apart from the fact that we do not restrict ourselves to normally distributed returns. More precisely 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

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

Recalling that $Var(X) = E\left[X^2\right] - \left(E\left[X\right]\right)^2$, we see that, in the notation of Theorem 2.1, 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 2.1. 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

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

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

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

or, in more concrete terms,

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

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

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

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

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

which, after identifying coefficients, gives us the recursions

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

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

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

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 6.1 *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

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

7 Continuous time

We now turn to the more delicate case of continuous time models. We start by presenting the basic setup in term of a fairly general controlled Markov process. We then formulate the problem and formally define the continuous time equilibrium concept. In order to derive the relevant extension of the Hamilton-Jacobi-Bellman equation we discretize, use our previously derived results in discrete time, and go to the limit. Since the limiting procedure is somewhat informal we need to prove a formal verification theorem, showing the connection between the extended HJB equation and the previously defined equilibrium concept.

7.1 Setup

We consider, on the time interval $[0, T]$ a **controlled Markov process** in continous time. The process X lives 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 way that controls are influencing the dynamics of the process is formalized by specifying the controlled infinitesimal generator of X.

Definition 7.1 *For any fixed $u \in \mathcal{U}$ we denote the corresponding infinitesimal generator by $\mathbf{A}^u$. For a control law $\mathbf{u}$, the corresponding generator is denoted by $\mathbf{A}^{\mathbf{u}}$.*

As an example: of X is a controlled SDE of the form

$$dX_t = \mu(X_t, u_t)dt + \sigma(X_t, u_t)dW_t,$$

then we have, for any real valued function $f(t,x)$, and for any **fixed** $u \in \mathcal{U}$

$$\mathbf{A}^u f(t,x) = \frac{\partial f}{\partial t}(t,x) + \mu(x,u)\frac{\partial f}{\partial x}(t,x) + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 f}{\partial x^2}(t,x)$$

For a **control law** $\mathbf{u}(t,x)$ we have

$$\mathbf{A}^{\mathbf{u}} f(t,x) = \frac{\partial f}{\partial t}(t,x) + \mu(x,\mathbf{u}(t,x))\frac{\partial f}{\partial x}(t,x) + \frac{1}{2}\sigma^2(x,\mathbf{u}(t,x))\frac{\partial^2 f}{\partial x^2}(t,x).$$

By the Kolmogorov backward equation, the infinitesimal generator will, for any control law $\mathbf{u}$, determine the distribution of the process X, and to stress this fact we will use the notation $X_t^{\mathbf{u}}$. In particular we will have, for each $h \in \mathbf{R}$ an operator $\mathbf{P}_h^{\mathbf{u}}$, operating on real valued functions of the form $f(t,x)$, and defined as

$$\mathbf{P}_h^{\mathbf{u}} f(t,x) = E\left[f(t+h, X_{t+h}^{\mathbf{u}}) \middle| X_t = x \right]. \tag{36}$$

We also recall that

$$\mathbf{A}^{\mathbf{u}} = \frac{d\mathbf{P}_h^{\mathbf{u}}}{dh}|_{h=0}. \tag{37}$$

7.2 Basic problem formulation

For a fixed $(t,x) \in [0,T] \times \mathcal{X}$, a fixed control law $\mathbf{u}$, we consider, as in discrete time, the simplified functional

$$J(t,x,\mathbf{u}) = E_{t,x}\left[F(x, X_T^{\mathbf{u}})\right] + G\left(x, E_{t,x}\left[X_T^{\mathbf{u}}\right]\right). \tag{38}$$

The most general case will be treated in Section 7.3.2. As in discrete time we have the game theoretic interpretation that, for each point t in time we have a player ("player t") choosing u_t who wants to maximize the functional above. Player t can, however, only affect the dynamics of the process X by choosing the control u_t exactly at time t. At another time, say s, the control u_s will be chosen by player s. We again attack this problem by looking for a Nash subgame perfect equilibrium point. The intuitive picture is exactly like in continuous time: An equilibrium strategy $\hat{\mathbf{u}}$ is characterized by the property that if all players on the half open interval $(t,T]$ uses $\hat{\mathbf{u}}$, then it is optimal for player t to use $\hat{\mathbf{u}}$.

However, in continuous time this is not a bona fide definition. Since player t can only choose the control u_t exactly at time t, he only influences the control on a time set of Lebesgue measure zero, and for most models this will have no effect whatsoever on the dynamics of the process. We thus need another definition of the equilibrium concept, and we follow [4] and [5], who were the first to use the definition below.

Definition 7.2 *Consider a control law* $\hat{\mathbf{u}}$ *(informally viewed as a candidate equlibrium law). Choose a fixed* $u \in \mathcal{U}$*, a fixed real number* $h > 0$*. Also fix an arbitrarily chosen initial point* (t, x)*. Define the control law* $\mathbf{u}_h$ *by*

$$\mathbf{u}_h(s, y) = \begin{cases} u, & \text{for } t \le s < t+h, \quad y \in \mathcal{X} \\ \hat{\mathbf{u}}(s, y), & \text{for } t+h \le s \le T, \quad y \in \mathcal{X} \end{cases}$$

If

$$\liminf_{h \to 0} \frac{J(t, x, \hat{\mathbf{u}}) - J(t, x, \mathbf{u}_h)}{h} \ge 0,$$

for all $u \in \mathcal{U}$*, we say that* $\hat{\mathbf{u}}$ *is an* **equilibrium** *control law. The equilibrium* **value function** V *is defined by*

$$V(t, x) = J(t, x, \hat{\mathbf{u}}).$$

Remark 7.1 *This is our continuous time formalization of the corresponding discrete time equilibrium concept. Note the necessity of dividing by* h*, since for most models we trivially would have*

$$\lim_{h \to 0} \{J(t, x, \hat{\mathbf{u}}) - J(t, x, \mathbf{u}_h)\} = 0.$$

We also note that we do not get a perfect correspondence with the discrete time equilibrium concept, since if the limit above equals zero for all $u \in \mathcal{U}$*, it is not clear that this corresponds to a maximum or just to a stationary point.*

7.3 The extended HJB equation

We now assume that there exists an equilibrium control law $\hat{\mathbf{u}}$ (not necessarily unique) and we go on to derive and extension of the standard Hamilton-Jacobi-Bellman (henceforth HJB) equation for the determination of the corresponding value function V. To clarify the logical structure of the derivation we outline our strategy as follows.

- We discretize (to some extent) the continuous time problem. We then use our results from discrete time theory to obtain a discretized recursion for $\hat{\mathbf{u}}$ and we then let the time step tend to zero.
- In the limit we obtain our continuous time extension of the HJB equation. Not surprisingly it will in fact be a system of equations.
- In the discretizing and limiting procedure we mainly rely on informal heuristic reasoning. In particular we have do **not** claim that the derivation is a rigorous one. The derivation is, from a logical point of view, only of motivational value.
- We then go on to show that our (informally derived) extended HJB equation is in fact the "correct" one, by proving a rigorous verification theorem.

7.3.1 Deriving the equation

In this section we will, in an informal and heuristic way, derive a continuous time extension of the HJB equation. Note again that we have no claims to rigor in the derivation, which is only motivational. To this end we assume that there exists an equilibrium law $\hat{\mathbf{u}}$ and we argue as follows.

- Choose an arbitrary initial point (t, x). Also choose a "small" time increment $h > 0$.
- Define the control law $\mathbf{u}_h$ on the time interval $[t, T]$ by

$$\mathbf{u}_h(s, y) = \begin{cases} u, & \text{for } t \le s < t + h, \quad y \in \mathcal{X} \\ \hat{\mathbf{u}}(s, y), & \text{for } t + h \le s \le T, \quad y \in \mathcal{X} \end{cases}$$

- If now h is "small enough" we expect to have

$$J(t, x, \mathbf{u}_h) \le J(t, x, \hat{\mathbf{u}}),$$

  and in the limit as $h \to 0$ we should have equality if $u = \hat{\mathbf{u}}(t, x)$.

If we now use our discrete time results, with n and $n+1$ replaced by t and $t+h$, we obtain the inequality

$$\left(\mathbf{A}_h^u V\right)(t, x) - \left(\mathbf{A}_h^u f\right)(t, x, x) + \left(\mathbf{A}_h^u f^x\right)(t, x) - \mathbf{A}_h^u \left(G \diamond g\right)(t, x) + \left(\mathbf{H}_h^u g\right)(t, x) \le 0$$

where

$$\left(\mathbf{A}_h^u V\right)(t, x) = E_{t,x}\left[V(t + h, X_{t+h}^u)\right] - V(t, x)$$

and similarly for the other terms. We now divide the inequality by h and let h tend to zero. The the operator $\mathbf{A}_h^u$ will converge to the infinitesimal operator $\mathbf{A}^u$, but the limit of $h^{-1}\left(\mathbf{H}_h^u g\right)(t, x)$ requires closer investigation.

We have in fact

$$\left(\mathbf{H}_h^u g\right)(t, x) = G^x(E_{t,x}\left[g(t + h, X_{t+h}^u)\right]) - G^x(g(t, x)).$$

Furthermore we have the approximation

$$E_{t,x}\left[g(t + h, X_{t+h}^u)\right] = g(t, x) + \mathbf{A}^u g(t, x) + o(h),$$

and using a standard Taylor approximation for G^x we obtain

$$G^x(E_{t,x}\left[g(t + h, X_{t+h}^u\right]) = G^x(g(t, x)) + G_y^x(g(t, x)) \cdot \mathbf{A}^u g(t, x) + o(h),$$

where

$$G_y^x(y) = \frac{\partial G^x}{\partial y}(y).$$

We thus obtain

$$\lim_{h \to 0} \frac{1}{h}\left(\mathbf{H}_h^u g\right)(t, x) = G_y^x(g(t, x)) \cdot \mathbf{A}^u g(t, x).$$

Collecting all results we arrive at our proposed extension of the HJB equation. To stress the fact that the arguments above are largely informal we state the equation as a definition rather than as proposition.

Definition 7.3 *The extended HJB system of equations for the Nash equilibrium problem is defined as follows.*

$$\begin{aligned}
\sup_{u\in\mathcal{U}} \{ (\mathbf{A}^u V)(t,x) - (\mathbf{A}^u f)(t,x,x) + (\mathbf{A}^u f^x)(t,x) & \\
- \mathbf{A}^u (G \diamond g)(t,x) + (\mathbf{H}^u g)(t,x) \} &= 0, \quad 0 \le t \le T, \\
\mathbf{A}^{\hat{\mathbf{u}}} f^y(t,x) &= 0, \quad 0 \le t \le T, \\
\mathbf{A}^{\hat{\mathbf{u}}} g(t,x) &= 0, \quad 0 \le t \le T, \\
V(T,x) &= F(x,x) + G(x,x), \\
f^y(T,x) &= F(y,x), \\
g(T,x) &= x.
\end{aligned}$$

Here $\hat{\mathbf{u}}$ *is the control law which realizes the supremum in the first equation, and* f^y*,* $G \diamond g$*, and* $\mathbf{H}g$ *are defined by*

$$\begin{aligned}
f^y(t,x) &= f(t,x,y) \\
(G \diamond g)(t,x) &= G(x, g(t,x)), \\
\mathbf{H}^u g(t,x) &= G_y(x, g(t,x)) \cdot \mathbf{A}^u g(t,x), \\
G_y(x,y) &= \frac{\partial G}{\partial y}(x,y).
\end{aligned}$$

We now have some comments on the extended HJB system.

- The first point to notice is that we have a **system** of recursion equation (5)-(10) for the simultaneous determination of V, f and g.
- 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 $(\mathbf{A}^u f)(t,x,x) + (\mathbf{A}^u f^x)(t,x)$ in the V-equation cancel, and the system reduces to the standard Bellman equation

$$\begin{aligned}
(\mathbf{A}^u V)(t,x) &= 0, \\
V(T,x) &= F(x).
\end{aligned}$$

- In order to solve the V-equation we need to know f and g but these are determined by the optimal control law $\hat{\mathbf{u}}$, which in turn is determined by the *sup*-part of the V-equation.
- We can view the system as a fixed point problem, where the optimal control law $\mathbf{u}$ solves an equation of the form $M(\mathbf{u}) = \mathbf{u}$. The mapping M is defined by the following procedure.
  - Start with a control $\mathbf{u}$.
  - Generate the functions f and g by the ODEs

$$\begin{aligned}
\mathbf{A}^{\mathbf{u}} f^y(t,x) &= 0, \\
\mathbf{A}^{\mathbf{u}} g(t,x) &= 0,
\end{aligned}$$

  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 the V-equation 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.

- The equations for g and f^y state that the processes $g(t, X_t^{\hat{\mathbf{u}}})$ and $f^y(t, X_t^{\hat{\mathbf{u}}}))$ are martingales. From the boundary conditions we then have the interpretation

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

- We note that the g function above appears, in a more restricted framework, already in [1], [4], and [5].

7.3.2 The general case

We now turn to the most general case of the present paper, where the functional J is given by

$$J(t,x,\mathbf{u}) = E_{t,x}\left[\int_t^T C\left(t,x,s,X_s^{\mathbf{u}}, \mathbf{u}(X_s^{\mathbf{u}})\right) ds + F(t,x,X_T^{\mathbf{u}})\right] + G\left(t,x,E_{t,x}\left[X_T^{\mathbf{u}}\right]\right). \tag{39}$$

Definition 7.4 *Given the objective functional (39) the extended HJB equation for V is given by*

$$\begin{aligned} \sup_{u\in\mathcal{U}} \Big\{ \left(\mathbf{A}^u V\right)(t,x) + C(t,x,t,x,u) - \int_t^T (\mathbf{A}^u c^s)(t,x,t,x) ds + \int_t^T (\mathbf{A}^u c^{txs})(t,x) ds \\ - \left(\mathbf{A}^u f\right)(t,x,t,x) + \left(\mathbf{A}^u f^{tx}\right)(t,x) - \mathbf{A}^u \left(G \diamond g\right)(t,x) + \left(\mathbf{H}^u g\right)(t,x) \Big\} &= 0, \end{aligned}$$

with boundary condition

$$V(T,x) = F(T,x,x) + G(T,x,x).$$

Furthermore, the following hold.

1. *For each fixed s and y, the function $f^{sy}(t,x)$ is defined by*

$$\begin{aligned} \mathbf{A}^{\hat{\mathbf{u}}} f^{sy}(t,x) &= 0, \quad 0 \le t \le T \\ f^{sy}(T,x) &= F(s,y,x) \end{aligned}$$

2. The function $g(t,x)$ is defined by

$$\begin{aligned}\mathbf{A}^{\hat{\mathbf{u}}} g(t,x) &= 0, \quad 0 \le t \le T \\ g(T,x) &= x.\end{aligned}$$

3. For each fixed s, r and y, the function $c^{syr}(t,x)$ is defined by

$$\begin{aligned}(\mathbf{A}^{\hat{\mathbf{u}}} c^{syr})(t,x) &= 0, \quad 0 \le t \le r \\ c^{syr}(r,x) &= C(s,y,r,x,\hat{\mathbf{u}}_r(x))\end{aligned}$$

4. We have used the notation

$$\begin{aligned}f(t,x,s,y) &= f^{sy}(t,x) \\ c^{s}(t,x,t,x) &= c^{txs}(t,x), \\ (G \diamond g)(t,x) &= G(t,x,g(t,x)), \\ \mathbf{H}^{u} g(t,x) &= G_y(t,x,g(t,x)) \cdot \mathbf{A}^{u} g(t,x), \\ G_y(x,y) &= \frac{\partial G}{\partial y}(t,x,y).\end{aligned}$$

5. The probabilistic interpretations of f, g, and c are given by

$$\begin{aligned}f^{sy}(t,x) &= E_{t,x}\left[F(s,y,X_T^{\hat{\mathbf{u}}})\right], \quad 0 \le t \le T \\ g(t,x) &= E_{t,x}\left[X_T^{\hat{\mathbf{u}}}\right], \quad 0 \le t \le T \\ c^{syr}(t,x) &= E_{t,x}\left[C\left(s,y,r,X_r^{\hat{\mathbf{u}}},\hat{\mathbf{u}}(X_r^{\hat{\mathbf{u}}})\right)\right], \quad 0 \le t \le r.\end{aligned}$$

6. In these expressions, $\hat{\mathbf{u}}$ always denotes the equilibrium control law.

Remark 7.2 *We can easily extend the result above to the case when the term $G(t,x,E_{t,x}[X_T^{\mathbf{u}}])$ is replaced by*

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

for some function h. In this case we simply define g by

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

The extended HJB system then looks exactly as in Definition 7.4 above, apart from the fact that the boundary condition for g is changed to

$$g(T,x) = h(x).$$

See [9] for an interesting application.

Remark 7.3 *It is of course also possible to include terms of the type*

$$K(t,s,E_{tx}[b(X_s^{\mathbf{u}})])$$

in the integral term of the value functional. The structure of the resulting HJB system is fairly obvious but we have omitted it since the present HJB system is, in our opinion, complicated enough as it is.

7.3.3 A simple special case

We see that the general extended HJB equation is quite complicated. In many concrete cases there are, however, cancellations between different terms in the equation. The simplest case occurs when the objective functional has the form

$$J(t,x,\mathbf{u}) = E_{t,x}\left[F(X_T^{\mathbf{u}})\right] + G\left(E_{t,x}\left[X_T^{\mathbf{u}}\right]\right),$$

and X is a scalar diffusion of the form

$$dX_t = \mu(X_t,u_t)dt + \sigma(X_t,u_t)dW_t.$$

In this case the extended HJB equation has the form

$$\sup_{u\in\mathcal{U}}\left\{\mathbf{A}^u V(t,x) - \mathbf{A}^u\left[G\left(g(t,x)\right)\right] + G'(g(t,x))\mathbf{A}^u g(t,x)\right\} = 0,$$

and a simple calculation shows that

$$-\mathbf{A}^u\left[G\left(g(t,x)\right)\right] + G'(g(t,x))\mathbf{A}^u g(t,x) = -\frac{1}{2}\sigma^2(x,u)G''(g(t,x))g_x^2(t,x),$$

where $g_x = \frac{\partial g}{\partial x}$. Thus the extended HJB equation becomes

$$\sup_{u\in\mathcal{U}}\left\{\mathbf{A}^u V(t,x) - \frac{1}{2}\sigma^2(x,u)G''(g(t,x))g_x^2(t,x)\right\} = 0, \tag{40}$$

We will use this result in Section 10 below.

7.3.4 A scaling result

We now extend the results from Section 2.4.6 to continuous time. Let us thus consider the objective functional (39) above and denote, as usual, the equilibrium control and value function by $\hat{\mathbf{u}}$ and V respectively. Let $\varphi : R \to R$ be a fixed real valued function and consider a new objective functional J_φ, defined by,

$$J_\varphi(t,x,\mathbf{u}) = \varphi(x)J(t,x,\mathbf{u}), \quad n = 0,1,\ldots,T$$

and denote the corresponding equilibrium control and value function by $\hat{\mathbf{u}}_\varphi$ and V_φ respectively. Since player No t is (loosely speaking) trying to maximize $J_\varphi(t,x,\mathbf{u})$ over u_t, and $\varphi(x)$ is just a scaling factor which is not affected by u_t the following result is intuitively obvious. The formal proof is, however, not quite trivial.

Proposition 7.1 *With notation as above we have*

$$\begin{aligned}
V_\varphi(t,x) &= \varphi(x)V(t,x), \\
\hat{\mathbf{u}}_\varphi(t,x) &= \hat{\mathbf{u}}(t,x).
\end{aligned}$$

Proof. For notational simplicity we consider the case when J is of the form

$$J(t,x,\mathbf{u}) = E_{t,x}\left[F(x, X_T^{\mathbf{u}})\right] + G\left(x, E_{t,x}\left[X_T^{\mathbf{u}}\right]\right). \tag{41}$$

The proof for the general case has exactly the same structure.

For J as above we have the extended HJB system

$$\begin{aligned}
\sup_{u\in\mathcal{U}} \{ \left(\mathbf{A}^u V\right)(t,x) - \left(\mathbf{A}^u f\right)(t,x,x) + \left(\mathbf{A}^u f^x\right)(t,x) & \\
- \mathbf{A}^u \left(G \diamond g\right)(t,x) + G_y(x, g(t,x)) \cdot \mathbf{A}^u g(t,x) \} &= 0, \quad 0 \le t \le T, \\
\mathbf{A}^{\hat{\mathbf{u}}} f^y(t,x) &= 0, \quad 0 \le t \le T, \\
\mathbf{A}^{\hat{\mathbf{u}}} g(t,x) &= 0, \quad 0 \le t \le T, \\
V(T,x) &= F(x,x) + G(x,x), \\
f^y(T,x) &= F(y,x), \\
g(T,x) &= x.
\end{aligned}$$

We now recall the probabilistic interpretations

$$\begin{aligned}
V(t,x) &= E_{t,x}\left[F(x, X_T^{\hat{\mathbf{u}}})\right] + G\left(x, E_{t,x}\left[X_T^{\hat{\mathbf{u}}}\right]\right) \\
f(t,x,y) &= E_{t,x}\left[F(y, X_T^{\hat{\mathbf{u}}})\right], \\
g(t,x) &= E_{t,x}\left[X_T^{\hat{\mathbf{u}}}\right].
\end{aligned}$$

and the definition

$$\left(G \diamond g\right)(t,x) = G(x, g(t,x)).$$

From this it follows that

$$V(t,x) = f(t,x,x) + \left(G \diamond g\right)(t,x),$$

so the first HJB equation above can be written

$$\sup_{u\in\mathcal{U}} \left\{ \left(\mathbf{A}^u f^x\right)(t,x) + G_y(x, g(t,x)) \cdot \mathbf{A}^u g(t,x) \right\} = 0.$$

We now turn to J_φ, which can be written

$$J(t,x,\mathbf{u}) = E_{t,x}\left[F_\varphi(x, X_T^{\mathbf{u}})\right] + G_\varphi\left(x, E_{t,x}\left[X_T^{\mathbf{u}}\right]\right),$$

where

$$\begin{aligned}
F_\varphi(x,y) &= \varphi(x) F(x,y), \\
G_\varphi(x,y) &= \varphi(x) G(x,y),
\end{aligned}$$

and we note that

$$\frac{\partial G_\varphi}{\partial y}(x,y) = \varphi(x) G_y(x,y)$$

We thus obtain the HJB equation

$$\sup_{u\in\mathcal{U}} \left\{ \left(\mathbf{A}^u f_\varphi^x\right)(t,x) + \varphi(x) G_y(x, g_\varphi(t,x)) \cdot \mathbf{A}^u g_\varphi(t,x) \right\} = 0,$$

with f_φ and g_φ defined by

$$\begin{aligned}
\mathbf{A}^{\hat{\mathbf{u}}_\varphi} f_\varphi^y(t,x) &= 0, \\
\mathbf{A}^{\hat{\mathbf{u}}_\varphi} g_\varphi(t,x) &= 0, \\
f_\varphi^y(T,x) &= \varphi(y)F(y,x), \\
g_\varphi(T,x) &= x.
\end{aligned}$$

From this it follows that we can write

$$\begin{aligned}
f_\varphi(t,x,y) &= \varphi(y)f_0(t,x,y), \\
g_\varphi(t,x) &= g_0(t,x)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{A}^{\hat{\mathbf{u}}_\varphi} f_0^y(t,x) &= 0, \\
\mathbf{A}^{\hat{\mathbf{u}}_\varphi} g_0(t,x) &= 0, \\
f_0^y(T,x) &= F(y,x), \\
g_0(T,x) &= x.
\end{aligned}$$

and the HJB equation has the form

$$\sup_{u\in\mathcal{U}} \left\{\varphi(x)\left(\mathbf{A}^u f_0^x\right)(t,x) + \varphi(x)G_y(x,g_0(t,x)) \cdot \mathbf{A}^u g_0(t,x)\right\} = 0,$$

or, equivalently,

$$\sup_{u\in\mathcal{U}} \left\{\left(\mathbf{A}^u f_0^x\right)(t,x) + G_y(x,g_0(t,x)) \cdot \mathbf{A}^u g_0(t,x)\right\} = 0.$$

We thus see that the system for f, g, and $\hat{\mathbf{u}}$ is exactly the same as that for f_0, g_0, and $\hat{\mathbf{u}}_\varphi$. We thus have

$$\begin{aligned}
f_\varphi(t,x,y) &= \varphi(y)f(t,x,y), \\
g_\varphi(t,x) &= g(t,x), \\
\hat{\mathbf{u}}_\varphi &= \hat{\mathbf{u}}.
\end{aligned}$$

Moreover, since

$$V_\varphi(t,x) = f_\varphi(t,x,x) + \left(G_\varphi \diamond g_\varphi\right)(t,x),$$

we obtain

$$V_\varphi(t,x) = \varphi(x)V(t,x).$$

∎

7.3.5 A Verification Theorem

As we have noted above, the derivation of the continuous time extension of the HJB equation is rather informal. It seems reasonable to expect that the system in Definition 7.3 will indeed determine the equilibrium value function V, but so far nothing has been formally proved. However, the following two conjectures are natural.

- Assume that there exists an equilibrium law $\hat{\mathbf{u}}$ and that V is the corresponding value function. Assume furthermore that V is regular enough to allow allow $\mathbf{A}^u$ to operate on it (in the diffusion case this would imply $V \in C^{1,2}$). Define f and g by

$$
\begin{aligned}
f(t,x,y) &= E_{t,x}\left[F(y, X_T^{\hat{\mathbf{u}}})\right], && (42)\\
g(t,x) &= E_{t,x}\left[X_T^{\hat{\mathbf{u}}}\right]. && (43)
\end{aligned}
$$

  Then V satisfies the extended HJB system and $\hat{\mathbf{u}}$ realizes the supremum in the equation.

- Assume that V, f, and g solves the extended HJB system and that the supremum i the V-equation is attained for every (t,x). Then there exists an equlibrium law $\hat{\mathbf{u}}$, and it is given by the optimal u in the in the V-equation. Furthermore, V is the corresponding equilibrium value function, and f and g allow for the interpretations (42)-(43).

In this paper we do not attempt to prove the first conjecture. Even for a standard time consistent control problem, it is well known that this is technically quite complicated, and it typically requires the theory of viscosity solutions. We will, however, prove the second conjecture. This obviously has the form of a verification result, and from standard theory we would expect that it can be proved with a minimum of technical complexity.

Theorem 7.1 (Verification Theorem) *Assume that V, f, g is a solution of the extended system in Definition 7.3, and that the control law $\hat{\mathbf{u}}$ realizes the supremum in the equation. Then $\hat{\mathbf{u}}$ is an equilibrium law, and V is the corresponding value function. Furthermore, f and g can be interpreted according to (42)-(43).*

Proof. The proof consists of two steps:

- We start by showing that V is the value function corresponding to $\hat{\mathbf{u}}$, i.e. that $V(t,x) = J(t,x,\hat{\mathbf{u}})$, and that f and g have the interpretations (42)-(43).

- In the second step we then prove that $\hat{\mathbf{u}}$ is indeed an equilibrium control law.

To show that $V(t,x) = J(t,x,\hat{\mathbf{u}})$, we use the V equation to obtain:

$$\begin{aligned}\left(\mathbf{A}^{\hat{\mathbf{u}}}V\right)(t,x) - \left(\mathbf{A}^{\hat{\mathbf{u}}}f\right)(t,x,x) + \left(\mathbf{A}^{\hat{\mathbf{u}}}f^{x}\right)(t,x) & \\ -\mathbf{A}^{\hat{\mathbf{u}}}\left(G\diamond g\right)(t,x) + \left(\mathbf{H}^{\hat{\mathbf{u}}}g\right)(t,x) &= 0,\end{aligned}$$

where

$$\mathbf{H}^{\hat{\mathbf{u}}}g(t,x) = G_y(x,g(t,x))\cdot\mathbf{A}^{\hat{\mathbf{u}}}g(t,x).$$

Since V, f, and g satsifies the extended HJB, we also have

$$\left(\mathbf{A}^{\hat{\mathbf{u}}}f^{x}\right)(t,x) = 0, \tag{44}$$

$$\mathbf{A}^{\hat{\mathbf{u}}}g(t,x) = 0, \tag{45}$$

and we thus have the equation

$$\left(\mathbf{A}^{\hat{\mathbf{u}}}V\right)(t,x) - \left(\mathbf{A}^{\hat{\mathbf{u}}}f\right)(t,x,x) - \mathbf{A}^{\hat{\mathbf{u}}}\left(G\diamond g\right)(t,x) = 0. \tag{46}$$

We now use Dynkin's Theorem which says that if X is a process with infinitesimal operator $\mathbf{A}$, and if $h(t,x)$ is a sufficiently integrable real valued function, then the process

$$h(t,X_t) - \int_0^t \mathbf{A}h(s,X_s)ds$$

is a martingale. Using Dynkin's Theorem we thus have

$$E_{t,x}\left[V(T,X_T)\right] = V(t,x) + E_{t,x}\left[\int_t^T \mathbf{A}^{\hat{\mathbf{u}}}V(s,X_s^{\hat{\mathbf{u}}})ds\right],$$

and from (46) we obtain

$$E_{t,x}\left[V(T,X_T^{\hat{\mathbf{u}}})\right] = V(t,x)+E_{t,x}\left[\int_t^T \mathbf{A}^{\hat{\mathbf{u}}}f(s,X_s^{\hat{\mathbf{u}}},X_s^{\hat{\mathbf{u}}}ds)\right]+E_{t,x}\left[\int_t^T \mathbf{A}^{\hat{\mathbf{u}}}\left(G\diamond g\right)(s,X_s^{\hat{\mathbf{u}}})ds\right]$$

We again refer to Dynkin and obtain

$$\begin{aligned}E_{t,x}\left[\int_t^T \mathbf{A}^{\hat{\mathbf{u}}}f(s,X_s^{\hat{\mathbf{u}}},X_s^{\hat{\mathbf{u}}})ds\right] &= E_{t,x}\left[f(T,X_T^{\hat{\mathbf{u}}},X_T^{\hat{\mathbf{u}}})\right] - f(t,x,x), \\ E_{t,x}\left[\int_t^T \mathbf{A}^{\hat{\mathbf{u}}}\left(G\diamond g\right)(s,X_s^{\hat{\mathbf{u}}})ds\right] &= E_{t,x}\left[G(X_T,g(T,X_T^{\hat{\mathbf{u}}}))\right] - G(x,g(t,x)).\end{aligned}$$

Using this and the boundary conditions for f and g we get

$$\begin{aligned}E_{t,x}\left[F(X_T^{\hat{\mathbf{u}}},X_T^{\hat{\mathbf{u}}}) + G(X_T^{\hat{\mathbf{u}}},X_T^{\hat{\mathbf{u}}})\right] &= V(t,x) + E_{t,x}\left[F(X_T^{\hat{\mathbf{u}}},X_T^{\hat{\mathbf{u}}})\right] - f(t,x,x) \\ &+ E_{t,x}\left[G(X_T^{\hat{\mathbf{u}}},X_T^{\hat{\mathbf{u}}})\right] - G(x,g(t,x)),\end{aligned}$$

i.e.

$$V(t,x) = f(t,x,x) + G(x,g(t,x)). \tag{47}$$

Now, from (44)-(45) it follows that the processes $f^y(s, X_s^{\hat{\mathbf{u}}})$ and $g(s, X_s^{\hat{\mathbf{u}}})$ are martingales, so from the boundary conditions for f and g we obtain

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

This shows that f and g have the correct interpretation and, plugging it into (47) we obtain

$$
V(t,x) = E_{t,x}\left[F(x, X_T^{\hat{\mathbf{u}}})\right] + G(x, E_{t,x}\left[X_T^{\hat{\mathbf{u}}}\right]) = J(t,x,\hat{\mathbf{u}}).
$$

We now go on to show that $\hat{\mathbf{u}}$ is indeed an equilibrium law. To that end we construct, for any $h > 0$ and an arbitrary $u \in \mathcal{U}$, the control law $\mathbf{u}_h$ defined in Definition 7.2. From Lemma 2.2, applied to the points t and $t+h$, we have

$$
\begin{aligned}
J_n(x,\mathbf{u}) &= E_{t,x}\left[J(t+h, X_{t+h}^u, \mathbf{u})\right] \\
&- \left\{E_{t,x}\left[f^{\mathbf{u}}(t+h, X_{t+h}^u, X_{t+h}^u)\right] - E_{t,x}\left[f^{\mathbf{u}}(t+h, X_{t+h}^u, x)\right]\right\} \\
&- \left\{E_{t,x}\left[G\left(X_{t+h}^u, g^{\mathbf{u}}(t+h, X_{t+h}^u)\right)\right] - G\left(x, E_{t,x}\left[g^{\mathbf{u}}(t+h, X_{t+h}^u)\right]\right)\right\}.
\end{aligned}
$$

where, for ease of notation, we have suppressed the lower index h of $\mathbf{u}_h$. By the construction of $\mathbf{u}$ we have

$$
\begin{aligned}
J(t+h, X_{t+h}^u, \mathbf{u}) &= V(t+h, X_{t+h}^u), \\
f^{\mathbf{u}}(t+h, X_{t+h}^u, x) &= f(t+h, X_{t+h}^u, x), \\
g^{\mathbf{u}}(t+h, X_{t+h}^u) &= g(t+h, X_{t+h}^u),
\end{aligned}
$$

so we obtain

$$
\begin{aligned}
J_n(x,\mathbf{u}) &= E_{t,x}\left[V(t+h, X_{t+h}^u)\right] \\
&- \left\{E_{t,x}\left[f(t+h, X_{t+h}^u, X_{t+h}^u)\right] - E_{t,x}\left[f(t+h, X_{t+h}^u, x)\right]\right\} \\
&- \left\{E_{t,x}\left[G\left(X_{t+h}^u, g(t+h, X_{t+h}^u)\right)\right] - G\left(x, E_{t,x}\left[g(t+h, X_{t+h}^u)\right]\right)\right\}.
\end{aligned}
$$

Furthermore, from the V-equation we have

$$
\begin{aligned}
&\left(\mathbf{A}^u V\right)(t,x) - \left(\mathbf{A}^u f\right)(t,x,x) + \left(\mathbf{A}^u f^x\right)(t,x) \\
&\qquad -\mathbf{A}^u\left(G \diamond g\right)(t,x) + \left(\mathbf{H}^u g\right)(t,x) \leq 0,
\end{aligned}
$$

for all $u \in \mathcal{U}$. Discretizing this gives us

$$
\begin{aligned}
E_{t,x}\left[V(t+h, X_{t+h}^u)\right] - V(t,x) - \left\{E_{t,x}\left[f(t, X_{t+h}^u, X_{t+h}^u)\right] - f(t,x,x)\right\} \\
+E_{t,x}\left[f(t, X_{t+h}^u, x)\right] - f(t,x,x) \\
-E_{t,x}\left[G(t+h, g(t+h, X_{t+h}^u)\right] + G(x, g(t,x)) \\
+G(x, E_{t,x}\left[g(t+h, X_{t+h}^u)\right] - G(x, g(t,x)) \quad \leq \quad o(h),
\end{aligned}
$$

or, after simplification,

$$\begin{aligned} V(t,x) &\geq E_{t,x}\left[V(t+h, X^u_{t+h})\right] - E_{t,x}\left[f(t, X^u_{t+h}, X^u_{t+h})\right] + E_{t,x}\left[f(t, X^u_{t+h}, x)\right] \\ &- E_{t,x}\left[G(t+h, g(t+h, X^u_{t+h})\right] + G(x, E_{t,x}\left[g(t+h, X^u_{t+h})\right] + o(h). \end{aligned}$$

Using the expression for $J_n(x, \mathbf{u})$ above, and the fact that $V(t,x) = J(t,x,\hat{\mathbf{u}})$, we obtain

$$J(t,x,\hat{\mathbf{u}}) - J(t,x,\mathbf{u}) \geq o(h),$$

so

$$\liminf_{h \to 0} \frac{J(t,x,\hat{\mathbf{u}}) - J(t,x,\mathbf{u})}{h} \geq 0,$$

and we are done. ∎

8 An equivalent time consistent problem

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

$$\begin{aligned} \sup_{u \in \mathcal{U}} \{ (\mathbf{A}^u V)(t,x) + C(x,x,u) - \int_t^T (\mathbf{A}^u c^s)_t(x,x) ds + \int_t^T (\mathbf{A}^u c^{s,x})_t(x) ds \\ - (\mathbf{A}^u f)(t,x,x) + (\mathbf{A}^u f^x)(t,x) - \mathbf{A}^u (G \diamond g)(t,x) + (\mathbf{H}^u g)(t,x) \} &= 0. \end{aligned}$$

Let us now assume that there exists an equilibrium control law $\hat{\mathbf{u}}$. Using $\hat{\mathbf{u}}$ we can then construct c, f and g by solving the associated equations in Definition 7.3. We now define the function h by

$$\begin{aligned} h(t,x,u) &= C(x,x,u) - \int_t^T (\mathbf{A}^u c^s)_t(x,x) ds + \int_t^T (\mathbf{A}^u c^{s,x})_t(x) ds \\ &- (\mathbf{A}^u f)(t,x,x) + (\mathbf{A}^u f^x)(t,x) - \mathbf{A}^u (G \diamond g)(t,x) + (\mathbf{H}^u g)(t,x). \end{aligned}$$

With this definition of h, the equation for V above and its boundary condition become

$$\begin{aligned} \sup_{u \in \mathcal{U}} \{ (\mathbf{A}^u V)(t,x) + h(t,x,u) \} &= 0, \\ V(T,x) &= F(x,x) + G(x,x). \end{aligned}$$

We now observe, by inspection, that this is a standard HJB equation for the standard time consistent optimal control problem to maximize

$$E_{t,x}\left[\int_t^T h(s, X_s, u_s) ds + F(X_T, X_T) + G(X_T, X_T)\right]. \tag{48}$$

We have thus proved the following result.

Proposition 8.1 *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 (48).*

We immediately remark that the Proposition 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.

Furthermore, Proposition 8.1 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 8.1 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".

9 Example: Non exponential discounting

We now illustrate the theory developed above, and the first example we consider is a fairly general case of a control problem with non exponential discounting. The model is specified as follows.

- We consider a controlled, not necessarily time homogeneous, Markov process in continuous time with controlled infinitesimal generator $\mathbf{A}^u$.
- The value functional for player number t is given by

  $$J(t,c,\mathbf{u}) = E_{tx}\left[\int_t^T \varphi(t-s)H(X_s^{\mathbf{u}}, u\left(X_s^{\mathbf{u}}\right))ds + \varphi(T-t)\Gamma\left(X_T^{\mathbf{u}}\right)\right],$$

  where the discounting function $\varphi(t)$, the local utility function $H(x,u)$ and the final state utility function $\Gamma(x)$ are deterministic functions.
- Without loss of generality we assume that

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

In the notation of Definition 7.4 we see that we have no G-term so the extended HJB equation has the form

$$\begin{aligned}\sup_{u\in\mathcal{U}}\{\left(\mathbf{A}^u V\right)(t,x) + C(t,x,t,x,u) - \int_t^T (\mathbf{A}^u c^s)(t,x,t,x)ds + \int_t^T (\mathbf{A}^u c^{txs})(t,x)ds \\ - \left(\mathbf{A}^u f\right)(t,x,t,x) + \left(\mathbf{A}^u f^{tx}\right)(t,x)\} &= 0,\end{aligned}$$

We recall the relations

$$\begin{aligned}f^{sy}(t,x) &= E_{t,x}\left[F(s,y,X_T^{\hat{\mathbf{u}}})\right], \quad 0 \le t \le T \\ f(t,x,t,x) &= f^{tx}(t,x), \\ c^{rys}(t,x) &= E_{t,x}\left[C\left(r,y,s,X_s^{\hat{\mathbf{u}}},\hat{\mathbf{u}}(X_s^{\hat{\mathbf{u}}})\right)\right], \quad 0 \le t \le r. \\ c^s(t,x,t,x) &= c^{sxt}(t,x)\end{aligned}$$

as well as the notational convention that the operator $\mathbf{A}^u$ operates only on the variables within the parenthesis, whereas upper case indices are treated as constant parameters. In our case the y variable is not present, so we can write

$$\begin{aligned}f^{sy}(t,x) &= f^s(t,x) \\ c^{rys}(t,x) &= c^{rs}(t,x)\end{aligned}$$

and the extended HJB equation takes the form

$$\begin{aligned}\sup_{u\in\mathcal{U}}\{\left(\mathbf{A}^u V\right)(t,x) + C(t,t,x,u) - \int_t^T (\mathbf{A}^u c^s)(t,t,x)ds + \int_t^T (\mathbf{A}^u c^{ts})(t,x)ds \\ - \left(\mathbf{A}^u f\right)(t,t,x) + \left(\mathbf{A}^u f^{t}\right)(t,x)\} &= 0,\end{aligned}$$

where

$$\begin{aligned}f^s(t,x) &= E_{t,x}\left[F(s,X_T^{\hat{\mathbf{u}}})\right], \quad 0 \le t \le T, \\ f(t,t,x) &= f^t(t,x), \\ c^{rs}(t,x) &= E_{t,x}\left[C\left(r,s,X_s^{\hat{\mathbf{u}}},\hat{\mathbf{u}}(X_s^{\hat{\mathbf{u}}})\right)\right], \\ c^s(t,t,x) &= c^{ts}(t,x).\end{aligned}$$

In our case we furthermore have

$$\begin{aligned}C(t,s,x,u) &= \varphi(s-t)H(x,u), \\ F(t,x) &= \varphi(T-t)\Gamma(x).\end{aligned}$$

so we can write

$$\begin{aligned}f^s(t,x) &= \varphi(T-s)\gamma(t,x), \\ c^{rs}(t,x) &= \varphi(s-r)h^s(t,x), \\ c^s(t,t,x) &= \varphi(s-t)h^s(t,x)\end{aligned}$$

where γ and h are defined by

$$
\begin{aligned}
\gamma(t,x) &= E_{t,x}\left[\Gamma(X_T^{\hat{\mathbf{u}}})\right], \quad 0 \leq t \leq T, \\
h^s(t,x) &= E_{t,x}\left[H\left(X_s^{\hat{\mathbf{u}}}, \hat{\mathbf{u}}(X_s^{\hat{\mathbf{u}}})\right)\right], \quad 0 \leq t \leq s.
\end{aligned}
$$

We now easily obtain

$$
\begin{aligned}
C(t,t,x,u) &= H(x,u), \\
\mathbf{A}^u c^{rs}(t,x) &= \varphi(s-r)\mathbf{A}^u h^s(t,x), \\
\mathbf{A}^u c^s(t,t,x) &= \varphi(s-t)\mathbf{A}^u h^s(t,x) - \varphi'(t-s)h^s(t,x), \\
\mathbf{A}^u f^t(t,x) &= \varphi(T-t)\mathbf{A}^u \gamma^s(t,x), \\
\mathbf{A}^u f(t,t,x) &= \varphi(T-t)\mathbf{A}^u \gamma(t,x) - \varphi'(T-t)\gamma(t,x).
\end{aligned}
$$

We have thus proved the following result.

Proposition 9.1 *The extended HJB system has the form*

$$
\begin{aligned}
\sup_{u \in \mathcal{U}} \left\{ \mathbf{A}^u V(t,x) + H(x,u) + \int_t^T \varphi'(s-t)h^s(t,x)ds + \varphi'(T-t)\gamma(t,x) \right\} &= 0, \\
V(T,x) &= \Gamma(x),
\end{aligned}
$$

where h^s and γ are determined by

$$
\begin{aligned}
\mathbf{A}^{\hat{\mathbf{u}}} h^s(t,x) &= 0, \quad 0 \leq t \leq s, \\
\mathbf{A}^{\hat{\mathbf{u}}} \gamma(t,x) &= 0, \quad 0 \leq t \leq T, \\
h^s(s,x) &= H\left(x, \hat{\mathbf{u}}_s(x)\right), \\
\gamma(T,x) &= \Gamma(x).
\end{aligned}
$$

This generalizes the corresponding results in [4] and [5], where special cases are treated in great detail.

10 Example: Mean-variance control

In this example we will consider dynamic mean variance optimization. This is a continuous time version of a standard Markowitz investment problem, where we penalize the risk undertaken by the conditional variance. As noted in the introduction, in a Wiener driven framework this example is studied intensively in [1], where the authors also consider the case of multiple assets, as well as the case of a hidden Markov process (unobservable factors) driving the parameters of the asset price dynamics. For illustrative purposes we first consider the simplest possible case of a Wiener driven single risky asset and, without any claim of originality, re-derivethe corresponding results of [1]. We then extend the model in [1] and study the case when the risky asset is driven by a point process as well as by a Wiener process.

10.1 The simplest case

We consider a market formed by a risky asset with price process S and a risk free money account with price process B. The price dynamics are given by

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma S_t dW_t, \\ dB_t &= r B_t dt, \end{aligned}$$

where α and σ are known constants, and r is the constant short rate.

Let u_t be the amount of money invested in the risky asset at time t. The value X_t of a self-financing portfolio based on S and B will then evolve according to the SDE

$$dX_t = [rX_t + (\alpha - r)u_t]dt + \sigma u_t dW_t. \tag{49}$$

Our value functional is given by

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

so we want to maximize expected return with a penalty term for risk. Remembering the definition for the conditional variance

$$Var_{t,x}[X_T] = E_{t,x}[X_T^2] - E_{t,x}^2[X_T],$$

we can re-write our objective functional as

$$J(t, x, \mathbf{u}) = E_{t,x}\left[F(X_T^{\mathbf{u}})\right] - G(E_{t,x}\left[X_T^{\mathbf{u}}\right])$$

where $F(x) = x - \frac{\gamma}{2}x^2$ and $G(x) = \frac{\gamma}{2}x^2$. As seen in the previous sections, the term $G(E_{t,x}\left[X_T\right])$ leads to a time inconsistent game theoretic problem.

The extended HJB equation is then given by the following PDE system:

$$\begin{aligned} \sup_u \left\{ [rx + (\alpha - r)u]V_x + \frac{1}{2}\sigma^2 u^2 V_{xx} - \mathcal{A}^u (G \circ g) + \mathbf{H}^u g \right\} &= 0, \\ V(T, x) &= x, \\ \mathcal{A}^{\hat{u}} g &= 0, \\ g(T, x) &= x, \end{aligned}$$

where lower case index denotes the corresponding partial derivative. This case is covered in Section 7.3.3, and from (40) we can simplify to

$$\begin{aligned} V_t + \sup_u \left\{ [rx + (\alpha - r)u]V_x + \frac{1}{2}\sigma^2 u^2 V_{xx} - \frac{\gamma}{2}\sigma^2 u^2 g_x^2 \right\} &= 0 \\ V(T, x) &= x \\ \mathcal{A}^{\hat{u}} g &= 0 \\ g(T, x) &= x \end{aligned}$$

Given the linear structure of the dynamics, as well as of the boundary conditions, it is natural to make the *Ansatz*

$$
\begin{aligned}
V(t,x) &= A(t)x + B(t) \\
g(t,x) &= a(t)x + b(t).
\end{aligned}
$$

With this trial solution the HJB equation becomes

$$
\begin{aligned}
A_t x + B_t + \sup_u \left\{ [rx + (\alpha - r)u]A - \frac{\gamma}{2}\sigma^2 u^2 a^2 \right\} &= 0, & (50) \\
a_t x + b_t + [rx + (\alpha - r)\hat{u}]a &= 0, & (51) \\
A(T) &= 1, \\
B(T) &= 0, \\
a(T) &= 1, \\
b(T) &= 0.
\end{aligned}
$$

We first solve the static problem embedded in (50). From the first order condition, we obtain the optimal control as

$$
\hat{u}(t,x) = \frac{1}{\gamma}\frac{\alpha - r}{\sigma^2}\frac{A(t)}{a^2(t)},
$$

so the optimal control does not depend on x. Substituting this expression for $\hat{u}$ into (50) we obtain:

$$
A_t x + B_t + Arx + \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2}\frac{A^2}{a^2} = 0.
$$

By separation of variables we then get the following system of ODEs.

$$
\begin{aligned}
A_t + Ar &= 0, \\
A(T) &= 1, \\
B_t + \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2}\frac{A^2}{a^2} &= 0, \\
B(T) &= 0.
\end{aligned}
$$

We immediately obtain

$$
A(t) = e^{r(T-t)}.
$$

Inserting this expression for A into the second ODE yields

$$
\begin{aligned}
B_t + \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2}\frac{e^{2r(T-t)}}{a^2} &= 0, & (52) \\
B(T) &= 0.
\end{aligned}
$$

This equation contain the unknown function a, and to determine this we use equation (51). Inserting the expression for $\hat{u}$ into (51) gives us

$$\begin{aligned} a_t x + b_t + rxa + \frac{1}{\gamma}\frac{(\alpha - r)^2}{\sigma^2}\frac{e^{r(T-t)}}{a} &= 0, \\ a(T) &= 1, \\ b(T) &= 0. \end{aligned}$$

Again we have separation of variables and obtain the system

$$\begin{aligned} a_t + ar &= 0, \\ b_t + \frac{1}{\gamma}\frac{(\alpha - r)^2}{\sigma^2}\frac{e^{r(T-t)}}{a} &= 0. \end{aligned}$$

This yields

$$a(t) = e^{r(T-t)},$$

and the ODE for b then takes the form

$$\begin{aligned} b_t &= \frac{1}{\gamma}\frac{(\alpha - r)^2}{\sigma^2}, \\ b(T) &= 0. \end{aligned}$$

We thus have

$$b(t) = \frac{1}{\gamma}\frac{(\alpha - r)^2}{\sigma^2}(T - t).$$

Introducing the results in the optimal control formula, we get

$$\hat{u}(t, x) = \frac{1}{\gamma}\frac{\alpha - r}{\sigma^2}e^{-r(T-t)}.$$

Using the expression for a above, we can go back to equation (52) which now takes the form

$$B_t + \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2} = 0,$$

so

$$B(t) = \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2}(T - t).$$

Thus, the optimal value function is given by

$$V(t, x) = e^{r(T-t)}x + \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2}(T - t)$$

We summarize the results as follows:

Proposition 10.1 *For the model above we have the following results.*

- *The optimal amount of money invested in a stock is given by*

$$\hat{\mathbf{u}}(t,x) = \frac{1}{\gamma}\frac{\alpha - r}{\sigma^2}e^{-r(T-t)}.$$

- *The equilibrium value function is given by*

$$V(t,x) = e^{r(T-t)}x + \frac{1}{2\gamma}\frac{(\alpha - r)^2}{\sigma^2}(T-t).$$

- *The expected value of the optimal portfolio is given by*

$$E_{t,x}\left[X_T\right] = e^{r(T-t)}x + \frac{1}{\gamma}\frac{(\alpha - r)^2}{\sigma^2}(T-t).$$

Using Proposition 8.1 we can also construct the equivalent standard time consistent optimization problem. An easy calculation gives us the following result.

Proposition 10.2 *The equivalent (in the sense of Proposition 8.1) time consistent problem is to maximize the functional*

$$\max_u E_{t,x}\left[X_T - \frac{\gamma\sigma^2}{2}\int_t^T e^{2r(T-s)}u_s^2 ds\right]$$

given the dynamics (49).

We note in passing that

$$\sigma^2 u_t^2 dt = d\langle X\rangle_t.$$

10.2 A point process extension

We will now present an extension of the mean variance problem when the stock dynamics are driven by a jump diffusion. We consider a single risky asset with price S and a bank account with price process B. The results below can be easily extended to the case of multiple assets, but for ease of exposition, we restrict ourselves to the scalar case. The dynamics are given by

$$\begin{aligned} dS_t &= \alpha(t,S_t)S_t dt + \sigma(t,S_t)S_t dW_t + S_{t-}\int_{\mathcal{Z}}\beta(z)\mu(dz,dt), \\ dB_t &= r dt. \end{aligned}$$

Here W is a scalar Wiener process and μ is a marked point process on the mark space $\mathcal{Z}$ with deterministic intensity measure $\lambda(dz)$. Furthermore, $\alpha(t,s)$, $\sigma(t,s)$ and $\beta(z)$ are known deterministic functions and r is a known constant.

As before u_t denotes the amount of money invested in the stock at time t, and X is the value process for a self financing portfolio based on S and B. The dynamics of X_t are then given by

$$dX_t = [rX_t + (\alpha(t, S_t, Y_t) - r)u]dt + \sigma(t, S_t, Y_t)udW_t + u_{t-} \int_Z \beta(z)\mu(dz, dt).$$

Again we study the case of mean-variance utility, i.e.

$$J(t, x, \mathbf{u}) = E_{t,x}\left[X_T^{\mathbf{u}}\right] - \frac{\gamma}{2} Var_{t,x}\left(X_T^{\mathbf{u}}\right).$$

The extended HJB system now has the form

$$\begin{aligned}
\sup_u \left\{\mathcal{A}^u V(t, x, s) - \mathcal{A}^u (G \circ g)(t, x, s) + \left(\mathbf{H}^u g\right)(t, x, s)\right\} &= 0, && (53)\\
V(T, x, s) &= x, &&\\
\mathcal{A}^{\hat{u}} g &= 0, && (54)\\
g(T, x, s) &= x. &&
\end{aligned}$$

As before, we make the *Ansatz*

$$\begin{aligned}
V(t, x, s) &= A(t)x + B(t, s),\\
g(t, x, s) &= a(t)x + b(t, s),\\
A(T) &= 1,\\
B(T, s) &= 0\\
a(T) &= 1,\\
b(T, s) &= 0.
\end{aligned}$$

After some simple but tedious calculations, equation (53) can be re-written as

$$\begin{aligned}
&\sup_u \{A_t x + B_t + A\left[rx + (\alpha - r)u\right] + \alpha s B_s + \frac{1}{2}\sigma^2 s^2 B_{ss} + Au \int_Z \beta(z)\lambda(dz)\\
&+ \int_Z [\underbrace{B(t, s(1 + \beta(z))) - B(t, s)}_{\Delta_\beta B(t,s,z)}]\lambda(dz) - \frac{1}{2}\sigma^2\gamma[au + b_s s]^2\\
&-\frac{\gamma}{2}\int_Z [au\beta(z) + \underbrace{b(t, s(1 + \beta(z))) - b(t, s)}_{\Delta_\beta b(t,s,z)}]^2 \lambda(dz)\} = 0 \qquad (55)
\end{aligned}$$

First, we solve the embedded static problem in (55)

$$\max_u \left\{(\alpha - r)Au + Au \int_Z \beta(z)\lambda(dz) - \frac{1}{2}\sigma^2[au + b_s s]^2 - \frac{\gamma}{2}\int_Z [au\beta(z) + \Delta_\beta b]\lambda(dz)\right\}$$

and obtain the optimal control

$$\hat{\mathbf{u}}(t, x, s) = \frac{\left[\alpha(t, s) - r + \int_Z \beta(z)\lambda(dz)\right] A(t)}{\gamma a^2(t)[\sigma^2(t, s) + \int_Z \beta^2(z)\lambda(dz)]} - \frac{\sigma(t, s)b_s(t, s)s + \int_Z \beta(z)\Delta_\beta b(t, s, z)\lambda(dz)}{a(t)[\sigma^2(t, s) + \int_Z \beta^2(z)\lambda(dz)]}$$

Again we see that the optimal control does not depend on x. We can plug the optimal control into equation (55) and as before, we can separate variables to obtain an ODE for $A(t)$ and a PIDE for $B(t,s)$. The ODE for A is

$$\begin{aligned} A_t + rA &= 0 \\ A(T) &= 0 \end{aligned}$$

with solution $A(t) = e^{r(T-t)}$. The PIDE for $B(t,s)$ becomes

$$\begin{aligned} &B_t + (\alpha - r)\hat{u} + \alpha s B_s + \frac{1}{2}\sigma^2 s^2 B_{ss} + A\hat{u}\int_Z \beta(z)\lambda(dz) && (56) \\ &+ \int_Z [\Delta_\beta B]\lambda(dz) - \frac{1}{2}\sigma^2\gamma[a(t)\hat{u} + b_s s]^2 && (57) \\ &-\frac{\gamma}{2}\int_Z [a(t)\hat{u}\beta(z) + \Delta_\beta b]^2\lambda(dz) = 0 && (58) \\ &B(T,s) = 0 && (59) \end{aligned}$$

In order to solve this we need to determine the functions $a(t)$ and $b(t,s)$. To this end we use (54). This can be rewritten as

$$\begin{aligned} &a_t x + b_t + [rx + (\alpha - r)\hat{u}]a + \alpha s b_s + \frac{1}{2}\sigma^2 s^2 b_{ss} \\ &+ a\hat{u}\int_Z \beta(z)\lambda(dz) + \int_Z \Delta_\beta b\lambda(dz) = 0. && (60) \end{aligned}$$

with the appropriate boundary conditions for a and b. By separation of variables we obtain the ODE

$$\begin{aligned} a_t + ra &= 0, \\ a(T) &= 1, \end{aligned}$$

and the PIDE

$$\begin{aligned} b_t + (\alpha - r)\hat{u}a + \alpha s b_s + \frac{1}{2}\sigma^2 s^2 b_{ss} + a\hat{u}\int_Z \beta(z)\lambda(dz) + \int_Z \Delta_\beta b\lambda(dz) &= 0, \\ b(T,s) &= 0. \end{aligned}$$

From the ODE we have $a(t) = e^{r(T-t)}$ and, plugging this expression into the previous formula for $\hat{u}$, gives us

$$\hat{u} = \frac{\alpha - r + \int_Z \beta(z)\lambda(dz)}{\gamma[\sigma^2 + \int_Z \beta^2(z)\lambda(dz)]}e^{-r(T-t)} - \frac{\sigma b_s s + \int_Z \beta(z)\Delta_\beta b\lambda(dz)}{[\sigma^2 + \int_Z \beta^2(z)\lambda(dz)]}e^{-r(T-t)}$$

We can now insert this expression, as well as the formula for a, into the PIDE for b above to obtain the PIDE

$$
\begin{aligned}
& b_t + \left[\alpha - \frac{\left[\alpha - r + \int_Z \beta(y)\lambda(dy)\right]\sigma^2}{\left[\sigma^2 + \int_Z \beta^2(y)\lambda(dy)\right]}\right] s b_s + \frac{1}{2}\sigma^2 s^2 b_{ss} + \frac{\left[\alpha - r + \int_Z \beta(y)\lambda(dy)\right]^2}{\gamma\left[\sigma^2 + \int_Z \beta^2(y)\lambda(dy)\right]} \\
& \int_Z \Delta_\beta b(t,s,z)\left\{1 - \frac{\alpha - r + \int_Z \beta(y)\lambda(dy)}{\left[\sigma^2 + \int_Z \beta^2(y)\lambda(dy)\right]}\beta(z)\right\}\lambda(dz) = 0 \\
& b(T,s) = 0
\end{aligned}
$$

This rather forbidding looking equation cannot in general be solved explicitly, but by applying a Feynman-Kac representation theorem we can represent the solution as

$$
b(t,s) = E^Q_{t,s}\left[\int_t^T \frac{\left[\alpha(\tau,S_\tau) - r + \int_Z \beta(z)\lambda(dz)\right]^2}{\gamma\left[\sigma^2(\tau,S_\tau) + \int_Z \beta^2(z)\lambda(dz)\right]} d\tau\right]. \tag{61}
$$

Here the measure Q is absolutely continuous w.r.t. P, and the likelihood process

$$
L_t = \frac{dQ}{dP} \quad \text{on } \mathcal{F}_t,
$$

has dynamics given by

$$
dL_t = L_t \varphi dW_t + L_{t-}\int_Z \eta(z)\left[\mu(dz,dt) - \lambda(dz)dt\right],
$$

with φ and η given by

$$
\begin{aligned}
\varphi(t,s) &= -\frac{\left[\alpha(t,s) - r + \int_Z \beta(y)\lambda(dy)\right]\sigma(t,s)}{\left[\sigma^2(t,s) + \int_Z \beta^2(y)\lambda(dy)\right]}, \\
\eta(t,s,z) &= -\frac{\left[\alpha(t,s) - r + \int_Z \beta(y)\lambda(dy)\right]}{\left[\sigma^2(t,s) + \int_Z \beta^2(y)\lambda(dy)\right]}\beta(z).
\end{aligned}
$$

From the Girsanov Theorem it follows that the Q intensity λ^Q, of the point process $\mu(dt,dz)$ is given by

$$
\lambda^Q(t,s,dz) = \left\{1 - \frac{\left[\alpha(t,s) - r + \int_Z \beta(y)\lambda(dy)\right]}{\left[\sigma^2(t,s) + \int_Z \beta^2(y)\lambda(dy)\right]}\beta(z)\right\}\lambda(dz),
$$

and that

$$
dW_t = -\frac{\left[\alpha(t,S_t) - r + \int_Z \beta(y)\lambda(dy)\right]\sigma(t,S_t)}{\left[\sigma^2(t,S_t) + \int_Z \beta^2(y)\lambda(dy)\right]}dt + dW^Q_t
$$

where W^Q is Q a Wiener process. A simple calculation now shows that the Q dynamics of the stock prices S are given by

$$
dS_t = rS_t dt + S_t\sigma(t,S_t)dW^Q_t + S_{t-}\int_{\mathcal{Z}} \beta(z)\left[\mu(dt,dz) - \lambda^Q(t,S_t,dz)\right]
$$

so the measure Q is in fact a risk neutral martingale measure, and it is easy to check (see for example [2]) that Q is in fact the so called "minimal martingale measure" used in the context of local risk minimization and developed in [12] and related papers. This fact was, in a Wiener process framework, observed already in [1].

Performing similar calculations, one can show that the solution of the PIDE (58)-(59) can be represented as

$$\begin{aligned} B(t,s) &= E^Q_{t,s}\left[\int_t^T \frac{(\alpha - r + \int_Z \beta(z)\lambda(dz))^2}{\gamma[\sigma^2 + \int_Z \beta^2(z)\lambda(dz)]} d\tau\right] \\ &\quad - E^Q_{t,s}\left[\int_t^T \frac{1}{2}\sigma^2\gamma[e^{r(T-\tau)}\hat{u} + b_s s]^2 d\tau\right] \\ &\quad - E^Q_{t,s}\left[\int_t^T \frac{\gamma}{2}\int_Z [e^{r(T-\tau)}\hat{u}\beta(z) + \Delta_\beta b]^2\lambda(dz)d\tau\right], \end{aligned} \tag{62}$$

with Q as above.

We can finally summarize our results.

Proposition 10.3 *With notation as above, the following hold.*

- *The optimal amount of money invested in a stock is given by*

$$\mathbf{\hat{u}} = \frac{(\alpha - r + \int_Z \beta(z)\lambda(dz))}{\gamma[\sigma^2 + \int_Z \beta^2(z)\lambda(dz)]} e^{-r(T-t)} - \frac{(\sigma b_s s + \int_Z \beta(z)\Delta_\beta b(z)\lambda(dz))}{[\sigma^2 + \int_Z \beta^2(z)\lambda(dz)]} e^{-r(T-t)}$$

- *The mean-variance utility of the optimal portfolio is given by*

$$U(t,x) = e^{r(T-t)}x + b(t,s)$$

  where $b(t,s)$ is given by stochastic representation (61).

- *The expected terminal value of the optimal portfolio is*

$$E_{t,x}[X_T] = xe^{r(T-t)} + B(t,s)$$

  where x is the present portfolio value and $B(t,s)$ is given by the stochastic representation (62).

11 Example: The time inconsistent linear quadratic regulator

In this example we consider a problem where the utility function at T depends on the current state X_t. The model is specified as follows.

- The value functional for player t is given by

$$E_{tx}\left[\frac{1}{2}\int_t^T u_s^2 ds\right] + \frac{\gamma}{2}E_{tx}\left[(X_T - x)^2\right]$$

where γ is a positive constant.

- The state process X is scalar with dynamics

$$dX_t = [aX_t + bu_t]\,dt + \sigma dW_t,$$

where a, b and σ are given constants.

- The control u is scalar with no constraints.

This is a time inconsistent version of the classical linear quadratic regulator. Loosely speaking we want control the system so that the final sate X_T is close to x while at the same time keeping the control energy (formalized by the integral term) is small. The time inconsistency stems from the fact that the target point $x = X_t$ is changing as time goes by.

For this problem we have

$$\begin{aligned} F(x,y) &= y - x, \\ C(u) &= \frac{1}{2}u^2 \end{aligned}$$

and as usual we introduce the functions $f^y(t,x)$ and $f(t,x,x)$ by

$$\begin{aligned} f^y(t,x) &= E_{tx}\left[\left(X_T^{\hat{\mathbf{u}}} - y\right)^2\right], \\ f(t,x,y) &= f^y(t,x). \end{aligned}$$

The extended HJB system takes the form

$$\begin{aligned} \inf_u\left\{\mathbf{A}^u V(t,x) + \frac{1}{2}u^2 + \mathbf{A}^u f^x(t,x) - \mathbf{A}^u f(t,x,x)\right\} &= 0, \quad 0 \le t \le T \\ V(T,x) &= 0, \\ \mathbf{A}^{\hat{\mathbf{u}}} f^y(t,x) &= 0, \quad 0 \le t \le T \\ f^y(T,x) &= \frac{\gamma}{2}(x-y)^2. \end{aligned}$$

From the X dynamics we see that

$$\mathbf{A}^u = \frac{\partial}{\partial t} + (ax + bu)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2},$$

so we obtain

$$\begin{aligned} \mathbf{A}^u f^y(t,x) &= \frac{\partial f^y}{\partial t}(t,x) + (ax+bu)\frac{\partial f^y}{\partial x}(t,x) + \frac{1}{2}\sigma^2\frac{\partial^2 f^y}{\partial x^2}(t,x) \\ \mathbf{A}^u f(t,x,x) &= \frac{\partial f}{\partial t} + (ax+bu)\left[\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\right] + \frac{1}{2}\sigma^2\left[\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial x \partial y}\right], \end{aligned}$$

where the derivatives of f in the last equation are evaluated at (t,x,x). We thus obtain the following form of the HJB equation, where for shortness of notation we denote partial derivatives by lower case index so, for example, $f_x = \frac{\partial f}{\partial x}$. All V terms are evaluated at (t,x) and all f terms are evaluated at (t,x,x).

$$
\begin{aligned}
\inf_u \left\{ V_t + (ax+bu)V_x + \frac{1}{2}\sigma^2 V_{xx} + \frac{1}{2}u^2 - (ax+bu)f_y - \sigma^2 f_{xy} - \sigma^2 f_{yy} \right\} &= 0, \\
V(T,x) &= 0.
\end{aligned}
$$

The coupled system for f^y is given by

$$
\begin{aligned}
f_t^y(t,x) + [ax + b\hat{\mathbf{u}}(t,x)]\, f_x^y(t,x) + \frac{1}{2}\sigma^2 f_{xx}^y(t,x) &= 0, \quad 0 \le t \le T, \\
f^y(T,x) &= \frac{\gamma}{2}(x-y)^2.
\end{aligned}
$$

The first order condition in the HJB equation gives us

$$
\hat{\mathbf{u}}(t,x) = b\left\{ f_y(t,x,x) - V_x(t,x) \right\}
$$

and, inspired of the standard regulator problem, we now make the Ansatz (attempted solution)

$$
\begin{aligned}
f(t,x,y) &= A(t)x^2 + B(t)y^2 + C(t)xy + D(t)x + F(t)y + H(t), && (63) \\
V(t,x) &= \alpha(t)x^2 + \beta(t)x + \epsilon(t), && (64)
\end{aligned}
$$

where all coefficients are deterministic functions of time. We now insert the Ansatz into the HJB system, and perform an extremely large number of extremely boring calculations. As a result of these calculations, it turns out that the variables separate in the expected way and we have the following result.

Proposition 11.1 *For the time inconsistent regulator, we have the structure (63)-(64), where the coefficient functions solve the following system of ODEs.*

$$
\begin{aligned}
A_t + 2aA + 2b^2 A\left(2B + C - 2\alpha\right) &= 0, \\
B_t &= 0, \\
C_t + aC &= 0, \\
D_t + aD + b^2 D\left(2B + C - 2\alpha\right) + 2b^2 A(F-\beta) &= 0, \\
F_t + b^2 CF - \beta b^2 C &= 0, \\
H_t + b^2(F-\beta)D + \sigma^2 A &= 0, \\
\alpha_t + 2a\alpha + 2\alpha b^2\left(2B - C - 2\alpha\right) + \frac{1}{2}b^2\left(2B + C - 2\alpha\right)^2 - a\left(2B + C\right) & \\
-b^2\left(2B + C\right)\left(2B + C - 2\alpha\right) &= 0, \\
\beta_t + \alpha\beta + \beta b^2\left(2B + C - 2\alpha\right) + 2\alpha b^2\left(F - \beta\right) + b^2\left(2B + C - 2\alpha\right)\left(F - \beta\right) & \\
-aF - b^2 F\left(2B + C - 2\alpha\right) - b^2\left(F - \beta\right)\left(2B + C\right) &= 0, \\
\epsilon_t + b^2\beta\left(F - \beta\right) + \sigma^2 a + \frac{1}{2}b^2\left(F - \beta\right)^2 - b^2\left(F - \beta\right)F - \sigma^2 C - \sigma^2 B &= 0.
\end{aligned}
$$

With boundary conditions

$$\begin{aligned} A(T) = \frac{\gamma}{2}, \quad B(T) &= \frac{\gamma}{2}, \quad C(T) = -\gamma, \\ D(T) = 0, \quad F(T) &= 0, \quad H(T) = 0, \\ \alpha(T) = 0, \quad \beta(T) &= 0, \quad \epsilon(T) = 0. \end{aligned}$$

12 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, in discrete as well as in continuous time, 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. I particular we have 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.

Some obvious open questions are the following.

- Prove existence for a reasonably general class of problems with infinite horizon.
- 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?

References

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

[2] Björk, T., and Slinko, I. Towards a general theory of good deal bounds. *Review of Finance 10* (2006), 221–260.

[3] Czichowsky, C. Time-consistent mean-variance portfolio selection in discrete and continuous time, 2010. Working paper ETH.

[4] Ekeland, I., and Lazrak, A. Being serious about non-commitment: subgame perfect equilibrium in continuous time, 2006. Preprint. University of British Columbia.

[5] Ekeland, I., and Privu, T. Investment and consumption without commitment, 2007. Preprint. University of British Columbia.

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

[7] Harris, C., and Laibson, D. Dynamic choices of hyperbolic consumers. *Econometrica 69*, 4 (2001), 935–957.

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

[9] 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).

[10] 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.

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

[12] Schweizer, M. On the minimal martingale measure and the föllmer-schweizer decomposition. *Stochastic Analysis and it's Applications 13* (1995), 573–599.

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

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