Report n. 133

Backward-Forward Stochastic Differential Utility: Existence, Consumption and Equilibrium Analysis

Fabio Antonelli, Emilio Barucci, Maria Elvira Mancino

Pisa, Febbraio 1999

Backward-Forward Stochastic Differential Utility: Existence, Consumption and Equilibrium Analysis

Fabio Antonelli *

Dipartimento di Scienze Ambientali
Università della Tuscia (Viterbo)
Via S. Camillo De Lellis - 01100 VITERBO -ITALY

Emilio Barucci

Dipartimento di Statistica e Matematica applicata all'Economia
Università degli Studi di Pisa
Via C. Ridolfi, 10-56124, Pisa, ITALY

Maria Elvira Mancino

DIMADEFAS
Università degli Studi di Firenze
Via Lombroso, 6/17 - 50134 Florence, ITALY

Keywords: Habit Formation, Asset Pricing, Malliavin calculus, Backward - Forward Stochastic Differential Equations.
Classification: (JEL 1995) C61, D11, D81, G12

*The first author was partially supported by the NATO/CNR Advanced Fellowships Program. We thank participants at the ESEM 1998 Conference (Berlin), S. Polidoro and J. Ma for useful discussions. The usual disclaimers apply.

Abstract

In this paper we introduce the Backward-Forward Stochastic Differential Utility. The agent's utility associated with a consumption plan is defined backward in time, but it is also affected by the agent's habit, defined forward in time. Therefore the utility process becomes the solution of a backward-forward stochastic differential equation. We consider the case when the agent's habit is affected not only by past consumption but also by past experienced expected utility levels. By using and extending to our needs the existence results known in the theory of Backward-Forward Stochastic Differential Equations, we prove existence of the utility process and we are able to give an explicit representation of the utility gradient in two particular cases. The first one is when the utility function is linear in the habit and the coefficients are deterministic. The second considers a utility process that can be written as a function of the instantaneous consumption and of the habit. In both cases we develop the general equilibrium analysis deriving the interest rate of equilibrium, the market prices of risk, and a version of the CAPM.

1 Introduction

The analysis of the agents' decisions under risk plays a central role in modern economic theory and mathematical finance. In an intertemporal setting, the standard framework is the time *Additive Expected Utility* (AEU) which extends the classical Von Neumann-Morgenstern expected utility theory to a multiperiod setting. The AEU is based on an axiomatization of the agents' preferences: provided that the agents' preference relation satisfies certain assumptions, the induced consumption plans order can be represented by the AEU.

This framework has recently come under attack. Two main objections are put forward in the literature: the first concerns the axiomatization and the representation of the agent's preferences through the AEU, the second the implications that the AEU assumption has on the dynamics of the economy (consumption, asset prices, interest rate, etc.) when joined by the *Rational Expectations* hypothesis.

The first objection is based on a long list of experimental results contradicting the AEU axioms and on some unpleasant theoretical facts. Experimental evidence has shown that the agents do not behave according to the Von Neumann-Morgenstern Axioms (e.g. the Allais paradox) and that the standard setting is not rich enough to analyze the agents' behavior when their beliefs are affected by information ambiguity (e.g. the Ellsberg paradox). Some of the theoretical problems concern the temporal additivity, i.e. the assumption that the agent's benefit from the consumption at time t does not depend on what happens at time $t' \neq t$. This requirement rules out the possibility to include durable goods and habit formation in the analysis. Moreover in this setting the risk aversion and the willingness to substitute across time are mutually intertwined (high risk aversion/low substitution or low risk aversion/high substitution).

The second criticism comes from the intertemporal asset pricing theory developed since the seminal papers of [Lucas, 1978, Breeden, 1979, Cox et al., 1985]. Indeed, many difficulties arise in the attempt to reconcile the AEU asset pricing results with the empirical evidence. Among those, we recall the following: *the equity premium puzzle* (see e.g. [Mehra and Prescott, 1985]), the AEU dynamic optimization restrictions are rejected by the data (see e.g. [Hansen and Singleton, 1982]), consumption data show a dynamics smoother than the theoretical consumption process, excess volatility (see e.g. [West, 1988]).

The *Stochastic Differential Utility* (SDU) approach developed in [Epstein and Zin, 1989] and [Duffie and Epstein and Skiadas 1992] offers an appropriate tool to address some of these problems. We refer to [Constantinides, 1990] for a reexamination of the equity premium puzzle, to [Epstein and Wang, 1994] and [Epstein and Wang, 1995] for the analysis of the agent's decisions under uncertainty, to [Epstein, 1988] for the analysis of risk-aversion and intertemporal substitution, to [Constantinides, 1990, Detemple and Zapatero, 1991 and 1992, Detemple and Giannikos, 1996] for habit formation and finally to [Grossman and Laroque, 1990] and [Hindy and Huang, 1993] for durable capital goods.

The SDU is the continuous time version of the discrete time recursive utility function introduced in a stochastic environment by [Epstein and Zin, 1989], where the utility index associated with a discrete time consumption profile $\{c_t, c_{t+1}, \ldots\}$ is defined through an *aggregator* $\mathcal{W}$. At time t, this aggregator links the utility index for a consumption plan (V_t) to the current consumption c_t and to the certainty equivalent of the utility associated with that consumption profile from time $t+1$ onwards $\{c_{t+1}, \ldots\}$ (CV_{t+1}) : $V_t = \mathcal{W}(c_t, CV_{t+1})$. This way to model the agent's preferences allows to disentangle the risk aversion (through the certainty equivalence) from the willingness of intertemporal substitution (through the aggregator $\mathcal{W}$). The continuous time version of recursive utility was proposed in a deterministic setting in [Epstein, 1987] and in a stochastic one in [Duffie and Epstein and Skiadas 1992]. In the latter, the utility function is defined as the initial state of the solution of a backward stochastic differential equation identified by an aggregator, represented by a pair (f, A), where f is a differential version of $\mathcal{W}$ and A a local risk aversion measure.

One of the key features of the SDU with respect to the AEU is that the time separability is removed, allowing to introduce features such as local substitution (consumptions at nearby dates are at most perfect substitutes), durable goods (the satisfaction related to a good is not instantaneous but lasts for a period of time), habit formation (the instantaneous utility function depends on instantaneous consumption and on past consumption), endogeneous time preferences (the discount factor is not constant, it depends on consumption). This is done by assuming that the aggregator f is a function of the instantaneous consumption and of another process, defined forward in time, often representing a smoothed average of past consumption. In this general formulation, the SDU process becomes part of the solution of a backward-forward stochastic differential equation (BFSDE), where the utility process is defined backward in time, but is intertwined with the so called habit, which is the solution of a forward equation. In the current literature only decoupled systems have been considered, where the utility process depends on the habit process, but not viceversa. This requires that the agent's habit at time t depends on past consumption, but not on past expected utility. To remove this restriction, we suggest a more general utility process characterized by a full BFSDE. The agent's expected utility at time t is therefore affected by past consumption and by what the agent expected in the past about the future. We will call this utility process *Backward-Forward Stochastic Differential Utility* (BFSDU).

Existence of the utility process is proved in a general setting, whereas the optimal consumption problem and the equilibrium analysis are developed in two particular cases. The first one is a linear BFSDU with deterministic coefficients, the second one is a nonlinear BFSDU. In these two cases we characterize the Arrow-Debreu price process, we determine the equilibrium interest rate dynamics and the assets' risk premium. In the linear case we show that the equilibrium price process is smaller than the one obtained with the AEU and the risk premium is higher than the one obtained with the AEU. The general equilibrium analysis assuming a nonlinear BFSDU is more difficult. A conclusion similar to the one obtained with a linear utility function is obtained for the equilibrium price process, but no general result can be stated about the equilibrium interest rate and the risk premium. In general, a BFSDU does not help to solve the equity premium puzzle and does not induce a smoother consumption process.

The paper is organized as follows. In Section 2, we present our economy. In Section 3, we present the *backward-forward stochastic differential utility* and we recall the necessary conditions to ensure existence and uniqueness of the utility process. In Section 4, we analyze the optimal consumption problem via the martingale method. In Sections 5 and 6 we analyze two particular cases of BFSDU for which we are able to solve explicitly the optimal consumption problem. In the first case the BFSDU is linear with deterministic coefficients, while the second one considers nonlinearities but it assumes that the utility at time T can be written as a function of the current habit and of the instantaneous consumption.

2 The Economy and Related Literature

We consider a standard pure exchange one consumer economy with complete markets. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, on which a standard Brownian motion in $\mathbb{R}^d$ (W) is defined. The economy has a finite time horizon $[0,T]$ and W determines the flow of information through its natural filtration, augmented of the P-null sets and made right continuous, that we indicate by $\{\mathcal{F}_t : \ t \geq 0\}$. Let $\mathcal{F}_0$ be trivial.

We denote by

$$\mathcal{L}^2 = \{X : X \text{ is a predictable process such that } E(\int_0^T |X_s|^2 ds) < +\infty\},$$

and by $\mathcal{L}^2_+$, the space of $\mathcal{L}^2$ processes with values in $\mathbb{R}_+$

There are $d+1$ financial securities, which are continuously traded in frictionless markets and their equilibrium prices are denoted by S^i $(i = 0, \ldots, d)$. The 0-th security is the risk-free asset, its price is given by

$$S_t^0 = s_0^0 \exp\{\int_0^t r_u du\},$$

where r_t is a strictly positive, progressively measurable bounded process and $s_0^0 > 0$. The d-dimensional vector of the security prices $S^\top = (S^1, \ldots, S^d)$ (where $\top$ denotes transpose) instead satisfies

$$dS_t = \overline{S}_t[\mu_t^S dt + \sigma_t^S dW_t], \quad S_0 = s_0, \quad \overline{S} = \begin{pmatrix} S^1 & & 0 \\ & \ddots & \\ 0 & & S^d \end{pmatrix},$$

where the d-dimensional vector of mean returns μ^S and the $d \times d$ volatility matrix σ^S are bounded and progressively measurable and $s_0^i > 0$ for all $i = 1, \ldots, d$.

Each security pays dividends and the cumulative dividends process of security i is denoted by D_i. The vector of cumulative dividends satisfies

$$dD_t = \mu_t^D dt + \sigma_t^D dW_t,$$

where $\mu^D \in \mathbb{R}^{d\times 1}$ and $\sigma^D \in \mathbb{R}^{d\times d}$ are bounded and progressively measurable. Lastly, the gain process is defined as $G = S + D$, where the sum is done component by component and therefore it is an Itô process ($dG_t = [\mu_t^G dt + \sigma_t^G dW_t]$). The gain process can be written in returns rates as follows $\overline{S}_t[\mu_t dt + \sigma_t dW_t]$. Let σ_t be invertible.

In a complete markets economy there exists a unique equivalent martingale measure, called the *risk-neutral probability measure*, given by

$$Q(A) = E[\psi_T \mathbf{1}_A], \qquad A \in \mathcal{F}_T, \tag{1}$$

$$\psi_t = \exp\{-\int_0^t \lambda_s dW_s - \frac{1}{2}\int_0^t ||\lambda_s||^2 ds\} \tag{2}$$

where

$$\lambda_t = (\sigma_t)^{-1}[\mu_t - r_t \mathbf{1}], \quad \mathbf{1} = \underbrace{(1, \ldots, 1)}_{d}$$

denotes the market price of risk. Assuming no arbitrage, the discounted gain from trade is a martingale under $Q(\cdot)$.

The density ψ can be interpreted as the equilibrium price density of a *one consumer economy.* The agent is described by a pair (U, e), where $U : \mathcal{L}_+^2 \to \mathbb{R}$ is a utility functional and $e \in \mathcal{L}^2_+$ is an endowment process. Finally by $c \in \mathcal{L}^2_+$ we denote the consumption process and by $C_t = \gamma_0 + \int_0^t c_s ds$, $\gamma_0 > 0$, the cumulative consumption.

A portfolio process or trading strategy, $\pi \equiv (\pi^0, \overline{\pi}) = (\pi^0, \pi^1, \ldots, \pi^d)$, is a measurable, square integrable adapted process, where the i-th component represents the amount of money invested by the agent in the i-th asset.

Definition 2.1 : *A pair of consumption and portfolio policies, (c, π), for the representative agent is* **admissible**, *if it satisfies the budget constraint*

$$dX_t = (r_t X_t + e_t - c_t)dt + \overline{\pi}_t(\mu_t - r_t \mathbf{1})dt + \overline{\pi}_t \sigma_t dW_t, \quad X_0 = 0, \quad X_T \geq 0,$$

where X represents the agent's wealth and $X_T \geq 0$ is the no-bankruptcy condition. An admissible pair (c, π) is **optimal** *if there is no other admissible pair (c', π') such that $U(c') > U(c)$.*

Definition 2.2 : *A triple (S, c, π) is called an* **equilibrium** *if (c, π) is optimal, given the price processes S and the market clearing conditions $c_t = e_t$ (consumption good market) and $\overline{\pi}_t = 0$ (securities market) are satisfied for all $t \in [0, T]$.*

To keep the notation simple, we assume $d = 1$. The results can be easily extended to the multidimensional case.

The standard *Additive Expected Utility* U is defined as

$$U(C) = V_0(c), \quad V_t = E(\int_t^T e^{-\beta s} u(c_s) ds + v(X_T) | \mathcal{F}_t), \quad t \geq 0,$$

for a given consumption process c and a given utility function u identifying the agent. The factor β represents the discount factor ($\beta > 0$) and $v(X_T)$ the utility from wealth at time T.

The *Stochastic Differential Utility* (SDU) is instead defined as $U = V_0$, where V is the solution of the backward stochastic differential equation

$$dV_t = [-f(c_t, V_t) - \frac{1}{2} A(V_t) |\sigma_t^V|^2] dt + \sigma_t^V dW_t, \quad V_T = v(X_T).$$

The pair (f, A) is called an aggregator, it is given by two measurable functions $f : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$, $A : \mathbb{R} \rightarrow \mathbb{R}$. σ^V is a square integrable progressively measurable process. By taking conditional expectations, V_t may be equivalently rewritten as

$$V_t = E(\int_t^T [f(c_s, V_s) + \frac{1}{2} A(V_s) |\sigma_s^V|^2] ds + v(X_T) | \mathcal{F}_t).$$

As pointed out in [Duffie and Epstein and Skiadas 1992], any aggregator (f, A) admits an ordinally equivalent "normalized" aggregator $(\overline{f}, \overline{A})$, where $\overline{A} = 0$. In what follows we directly refer to f as a normalized aggregator. V becomes the solution of

$$V_t = E\left(\int_t^T f(Z_s(C), V_s)ds + V_T | \mathcal{F}_t\right), \tag{3}$$

where the general process $Z(C) \in \mathcal{L}_+^2$ may have different characterizations yielding different utility functions (see [Duffie and Skiadas, 1994]):

- *Additive Expected Utility*: $Z_t(C) = c_t$, $f(c_t, V_t) = u(c_t) - \beta V_t$,
- *Uzawa utility function*: (time varying discount factor)

$$Z_t(C) = c_t, \quad f(c_t, V_t) = u(c_t) - \beta(Z_t)V_t,$$

- *Habit formation*:

$$Z_t(C) = (c_t, y_t), \quad y_t = y_0 + \int_0^t h(c_s, y_s)ds, \quad f(c_t, y_t, V_t) = u(c_t, y_t) - \beta_t V_t,$$

- *Durable capital goods*:

$$Z_t(C) = \int_0^t k_{t-s} c_s ds, \quad f(Z_t, V_t) = u(Z_t) - \beta_t V_t,$$

  where k is a progressively measurable bounded process.

3 Backward-Forward Stochastic Differential Utility

In this section we present a generalization of the SDU by considering a utility process defined backward-forward in time. The backward-forward feature comes from the fact that the instantaneous utility from consumption depends on the agent's habit which is determined by past consumption/expected utility. This structure destroys the time additivity property and it is a step ahead from the models considered by [Constantinides, 1990, Detemple and Zapatero, 1991, Hindy and Huang, 1993, Detemple and Giannikos, 1996], in which the agent's habit is a function only of past consumption.

The habit is a weighted average of past consumption and of the conditional expected utility levels that the agent experienced in the past about the future consumption plan. This utility process captures the fact that the agent's prefereces are affected by past consumption and by what he expected in the past about the future. The utility from current consumption is negatively affected by the habit. If one's expectation in the past about his own future utility was high, then the agent is accustomed to that utility level and therefore he gets a low satisfaction from current consumption.

These ideas are formalized in the following system:

$$V_t = E(\int_t^T (u(c_s, Y_s) - \beta_s V_s)ds + \Gamma|\mathcal{F}_t) \tag{4}$$

$$Y_t = y_0 e^{-\int_0^t \alpha_u du} + \delta_t \int_0^t e^{-\int_s^t \alpha_u du}[\mu V_s + (1-\mu)c_s]ds, \tag{5}$$

where β, α, δ are bounded and positive adapted processes, $\mu \in [0,1]$, y_0 is a constant and Γ is a square integrable $\mathcal{F}_T$−measurable random variable. The properties of the function u will be specified later. We refer to $U(C) = V_0$ as the *Backward-Forward Stochastic Differential Utility* (BFSDU). The random variable Γ represents the utility at time T. We restrict our attention to an exponentially discounted utility process (β_t does not depend on c or y). Our analysis can be extended to an endogenous discount factor and to a time varying stochastic coefficient μ, under the appropriate conditions, with some computational costs.

The process Y describes the agent's habit, y_0 is the standard of living at time zero. The constant μ is the weight describing the forward/backward characterization of Y. If $\mu = 0$, then Y is independent of the utility process V and we obtain the classical backward habit formation process, as in the papers [Constantinides, 1990, Detemple and Zapatero, 1991, Detemple and Giannikos, 1996], that we mentioned before. If $\mu = 1$, we have the other extremal case, when the habit is affected only by the past expected utility. The processes α and δ measure the persistence of past habit and the effect of the instantaneous consumption on the habit. To simplify the analysis we will consider only the case of constant α, δ, β.

A complete backward-forward model incurs in some mathematical difficulties about the existence of the solution (V, Y). In general, this is not implied by the usual hypothesis of Lipschitz coefficients and some technical restrictions are needed, except when $\mu = 0$. In this case, the system is decoupled: for a given c, the habit process solves a forward equation independent of V, whose equation is instead solved backward in time, once Y is known. When $\mu > 0$, to ensure existence of the solution of the system (4)-(5) we use the results in [Antonelli, 1993] and those in [Ma et al., 1994]. In this field, it is important to quote also the results of [Hu and Peng, 1995], but we do not consider their approach here, since the required monotonicity condition is not verified by the coefficients of our equation.

To apply the results in [Antonelli, 1993] we make the following Assumption.

Assumption 3.1 :

A. y_0 is a positive constant, $\alpha,\ \delta > 0$.

B. $u : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}^+$ and there exists a constant $k > 0$ such that

$$|u(c,z) - u(c,w)| \leq k|z-w| \qquad \forall \ \ c \in \mathbb{R}_+,\ z, w \in \mathbb{R};$$

C. β is an adapted process, bounded by a constant M;

D. for any $c \in \mathcal{L}^2_+$, $\Gamma \in L^2(P)$ and $E(\int_0^T |u(c_s,0)|^2 ds) < +\infty$.

We recall that $\underline{\underline{S}}^2$ is the Banach space of semimartingales X so that $E(\sup_{0\leq t\leq T} |X_t|) < \infty$.

Proposition 3.2 : *Let Assumption 3.1 hold and set*

$$K = \max\{\max(k, \alpha), \max(\mu\delta, \sup_{0\leq s\leq T} |\beta_s|)\}.$$

If $\sqrt{8}KT < 1$, then there exists a unique pair (V,Y) in $\underline{\underline{S}}^2 \times \underline{\underline{S}}^2$ satisfying (4)-(5).

Proof: Let us rewrite (5) implicitly

$$Y_t = y_0 + \int_0^t (\delta[\mu V_s + (1-\mu)c_s] - \alpha Y_s) ds.$$

Under Assumption 3.1, the operator

$$L\begin{pmatrix} V_t \\ Y_t \end{pmatrix} = \begin{pmatrix} E(\int_t^T (u(c_s, Y_s) - \beta_s V_s) ds + \Gamma|\mathcal{F}_t) \\ y_0 + \int_0^t (\delta[\mu V_s + (1-\mu)c_s] - \alpha Y_s) ds \end{pmatrix}$$

goes from $\underline{\underline{S}}^2 \times \underline{\underline{S}}^2$ into itself. If $\sqrt{8}KT < 1$, L acts as a contraction on $\underline{\underline{S}}^2 \times \underline{\underline{S}}^2$, which is a Banach space and hence it identifies a unique fixed point. For more details we refer the reader to [Antonelli, 1993]. □

Under the same assumptions one can show, again by a contraction argument, that the system has a unique solution also when the initial condition for the habit is specified coherently with the dynamics of the process, that is $y_0 = \mu V_0 + (1-\mu)c_0$, provided that we choose T so that $\sqrt{10}KT < 1$.

In Section 5, we will show that our analysis can be fully developed when the equations are linear with deterministic coefficients. This simple case allows full generality for the terminal value Γ, but the linearity condition is quite stringent. To include some non-linearities in our model, we consider also the approach developed by [Ma et al., 1994] and

[Cvitanic and Ma, 1996]. This method associates a Partial Differential Equation (PDE) to the backward-forward system. To apply their technique, we have to impose that the final condition Γ depends on the consumption and on the habit at time T $(g(c_T, Y_T))$ and to use the process c as a state variable. Namely, for each consumption process $c \in \mathcal{L}^2_+$, with dynamics

$$dc_t = \mu(t, c_t)dt + \sigma(t, c_t)dW_t, \qquad c_0 = \gamma_0 > 0, \tag{6}$$

we want to find an adapted pair of processes (V, Y) that satisfies

$$V_t = E(g(c_T, Y_T) + \int_t^T [u(c_s, Y_s) - \beta_s V_s]ds|\mathcal{F}_t) \tag{7}$$

$$Y_t = y_0 + \int_0^t [\nu V_s + \eta c_s - \alpha Y_s]ds, \tag{8}$$

where $\nu = \delta\mu$ and $\eta = \delta(1-\mu)$ and g is a measurable $\mathbb{R}$-valued function on $\mathbb{R}^2$.

Equation (6) is clearly independent of the other two and by Picard's iterations we can prove

Proposition 3.3 : *Let $\mu(t,x)$ and $\sigma(t,x)$ be deterministic, continuous in t, globally Lipschitz in x with constant k_1. Then there exists a unique $c \in \underline{\underline{S}}^2$, solution of (6).*

Note that $c \in \underline{\underline{S}}^2$ implies $c \in \mathcal{L}^2$. Once the unique consumption process is determined, we proceed to solving equations (7)-(8). Having changed the nature of the final condition of V, we need to replace condition D of Assumption 3.1 with

D'. $g : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}^+ \cup \{0\}$ *is such that* $g(0,0) = 0$ *and for the same constant k, we have*

$$|g(x_1, y_1) - g(x_2, y_2)| \leq k(|x_1 - x_2| + |y_1 - y_2|).$$

for any $x_1, x_2, y_1, y_2 \in \mathbb{R}$, *besides* $E(\int_0^T |u(c_s, 0)|^2 ds) < +\infty$.

With the same notation as before for K, we can prove

Proposition 3.4 : *Let Assumption 3.1* A, B, C *and condition* D' *hold and assume that*

$$KT\sqrt{8(1+K^2)} < 1. \tag{9}$$

Then there exists a unique solution $(V, Y) \in \underline{\underline{S}}^2 \times \underline{\underline{S}}^2$ for (7)-(8).

Proof: We proceed exactly as for Proposition 3.2. For any given $c \in \underline{\underline{S}}^2$, we define the operator

$$L^c \begin{pmatrix} V_t \\ Y_t \end{pmatrix} = \begin{pmatrix} F^c(Y,V)_t \\ G^c(Y,V)_t \end{pmatrix} = \begin{pmatrix} y_0 + \int_0^t (\nu V_s + \eta c_s - \alpha Y_s) ds \\ E\left(\int_t^T [u(c_s, Y_s) - \beta_s V_s] ds + g(c_T, F(Y,V)_T) | \mathcal{F}_t \right) \end{pmatrix},$$

which, under our hypothesis, is again a contraction on $\underline{\underline{S}}^2 \times \underline{\underline{S}}^2$. □

Summarizing, we can affirm that for T small enough there exists a unique triple c, X, Y that satisfies equations (6),(7) and (8).

Remark 3.5 : If equation (6) is such that the solution c is positive, for instance in the linear case, then due to the linearity of (7) in V, its solution may be rewritten as

$$V_t = E\left(\mathrm{e}^{-\int_t^T \beta(r)dr} g(c_T, Y_T) + \int_t^T \mathrm{e}^{-\int_t^s \beta(r)dr} u(c_s, Y_s) ds | \mathcal{F}_t \right),$$

which implies that $V_t > 0$ for all ω and t, because of the properties of g and u. Consequently, being (8) linear in all variables, we have

$$Y_t = \mathrm{e}^{-\alpha t}\{y_0 + \int_0^t \mathrm{e}^{\alpha s}[\nu V_s + \eta c_s] ds\} \geq \mathrm{e}^{-\alpha T} y_0 > 0.$$

for all ω and t.

To address the optimal consumption problem and to carry out the equilibrium analysis we will need the standard concavity condition of $U(C)$ with respect to C. For the Backward SDU, this is readily obtained by assuming that the aggregator is concave in c (see [Duffie and Epstein and Skiadas 1992]). The situation is slightly more complex for a BFSDU. On this point we have the following Proposition

Proposition 3.6 *Let the same hypotheses as in Proposition 3.2 or 3.4 hold and let us assume that* $u, g : [0,\infty) \times (-\infty,\infty) \to (0,\infty)$ *are strictly concave in* c*, strictly decreasing in* y *and concave in the couple. Then for any* $c^1, c^2 \in \mathcal{L}^2$*, with respective cumulative consumptions* C^1, C^2*, and any constant* $\lambda \in [0,1]$*, we have*

$$U(\lambda C^1 + (1-\lambda)C^2) \geq \lambda U(C^1) + (1-\lambda)U(C^2).$$

Proof: Here we prove our statement only for the system (7)-(8), indeed when g does not depend on c and Y, the proof remains the same, since the final condition for the differences is simply 0.

For $c^i \in \mathcal{L}^2$ $(i = 1, 2)$ let (Y^i, V^i) be the corresponding solutions given by Proposition 3.4 and set

$$\begin{aligned} c_t^\lambda &= \lambda c_t^1 + (1-\lambda)c_t^2, \qquad Y_t^\lambda = \lambda Y_t^1 + (1-\lambda)Y_t^2, \qquad V_t^\lambda = \lambda V_t^1 + (1-\lambda)V_t^2 \\ u_t^\lambda &= \lambda u(c_t^1, Y_t^1) + (1-\lambda)u(c_t^2, Y_t^2), \qquad g_T^\lambda = \lambda g(c_T^1, Y_T^1) + (1-\lambda)g(c_T^2, Y_T^2). \end{aligned}$$

With this notation, the couple (Y^λ, V^λ) verifies the system

$$\begin{aligned} V_t^\lambda &= E(g_T^\lambda + \int_t^T [u_s^\lambda - \beta V_s^\lambda] ds | \mathcal{F}_t) \\ Y_t^\lambda &= y_0 + \int_0^t [\nu V_s^\lambda + \eta c_s^\lambda - \alpha Y_s^\lambda] ds. \end{aligned}$$

On the other hand, for T small enough, associated to the consumption c^λ, there exists a unique pair of processes (X^λ, U^λ) solution of

$$\begin{aligned} U_t^\lambda &= E(g(c_T^\lambda, X_T^\lambda) + \int_t^T [u(c_s^\lambda, X_s^\lambda) - \beta U_s^\lambda] ds | \mathcal{F}_t) \\ X_t^\lambda &= y_0 + \int_0^t [\nu U_s^\lambda + \eta c_s^\lambda - \alpha X_s^\lambda] ds. \end{aligned}$$

We want to show that $V_0^\lambda \leq U_0^\lambda$. When we take the differences between the above processes, we obtain the system

$$U_t^\lambda - V_t^\lambda = E(g(c_T^\lambda, Y_T^\lambda) + \int_t^T [(u(c_s^\lambda, Y_s^\lambda) - u_s^\lambda) - \beta(U_s^\lambda - V_s^\lambda)] ds | \mathcal{F}_t) \tag{10}$$

$$X_t^\lambda - Y_t^\lambda = \int_0^t [\nu(U_s^\lambda - V_s^\lambda) - \alpha(X_s^\lambda - Y_s^\lambda)] ds. \tag{11}$$

For this pair, we introduce the four stopping times

$$\begin{aligned} \tau_1 &= \inf\{t > 0 \,:\, X_t^\lambda - Y_t^\lambda < 0\} \wedge T \\ \tau_2 &= \inf\{t > 0 \,:\, X_t^\lambda - Y_t^\lambda > 0\} \wedge T \\ S_1 &= \inf\{t > 0 \,:\, U_t^\lambda - V_t^\lambda < 0\} \wedge T \\ S_2 &= \inf\{t > 0 \,:\, U_t^\lambda - V_t^\lambda > 0\} \wedge T. \end{aligned}$$

Being τ_i stopping times, the sets $\{\tau_i = 0\}$ are $\mathcal{F}_0-$ measurable and must have probability either 0 or 1, since the initial $\sigma-$algebra is trivial.

We want to show that $P(\tau_1 > 0) = 1$. Indeed, if this is true, then $P(\tau_2 = 0) = 1$ and, for a.e. ω, $X_t^\lambda - Y_t^\lambda \geq 0$ on $[0, \tau_1]$. On the other hand, by linearity, from (11) we may write

$$X_t^\lambda - Y_t^\lambda = \mathrm{e}^{-\alpha t} \int_0^t \mathrm{e}^{\alpha s} \nu (U_s^\lambda - V_s^\lambda) ds \tag{12}$$

and consequently, for a.e. ω, we have

$$0 \leq \lim_{t \downarrow 0} \frac{X_t^\lambda - Y_t^\lambda}{t} = \lim_{t \downarrow 0} \frac{\nu \mathrm{e}^{-\alpha t}}{t} \int_0^t \mathrm{e}^{\alpha s} (U_s^\lambda - V_s^\lambda) ds = \nu (U_0^\lambda - V_0^\lambda)$$

implying the concavity of the Utility function.

It remains to prove that $P(\tau_1 > 0) = 1$. By contradiction, we assume that $P(\tau_1 = 0) = 1$, then $P(\tau_2 > 0) = 1$ and with the same argument as before we can conclude that $U_0^\lambda - V_0^\lambda \leq 0$. If $U_0^\lambda - V_0^\lambda < 0$, the continuity of paths implies $P(S_1 > 0) = 1$; if $U_0^\lambda - V_0^\lambda = 0$, instead we have two possibilities, either $S_1 = 0$ or $S_1 > 0$ a.s. In both cases, $S_1 > 0$ a.s. leads to an immediate contradiction. In fact $U_s^\lambda - V_s^\lambda \geq 0$ on the random interval $[0, S_1]$ and, from (12), the same happens for $X_s^\lambda - Y_s^\lambda$, that is to say $S_1 \leq \tau_1$ a.s. and hence $P(\tau_1 > 0) = 1$.

Hence we can affirm that $S_1 = 0$ a.s. and $U_0^\lambda - V_0^\lambda = 0$. So $S_2 > 0$ a.s. and for this bounded stopping time we can apply the optional sampling theorem to rewrite (10) as

$$0 \geq (U_t^\lambda - V_t^\lambda) \mathbf{1}_{\{t \leq S_2\}} = E(\mathrm{e}^{-\beta S_2} (U_{S_2}^\lambda - V_{S_2}^\lambda) + \int_t^{S_2} \mathrm{e}^{-\beta(s-t)} [u(c_s^\lambda, X_s^\lambda) - u_s^\lambda] ds | \mathcal{F}_t) \mathbf{1}_{\{t \leq S_2\}},$$

which in particular holds for $t = 0$. On the other hand, by the continuity of paths

$$0 \geq U_{S_2}^\lambda - V_{S_2}^\lambda = 0 \mathbf{1}_{\{S_2 < T\}} + (U_T^\lambda - V_T^\lambda) \mathbf{1}_{\{S_2 = T\}}$$

and $U_T^\lambda - V_T^\lambda = g(c_T^\lambda, X_T^\lambda) - g_T^\lambda$. On $\{S_2 = T\}$, also $\tau_2 = T$, thus $X_T^\lambda - Y_T^\lambda \leq 0$ and by the assumptions on g (decreasing in y and strictly concave in c, concave in the couple) we have

$$g(c_T^\lambda, X_T^\lambda) - g_T^\lambda > 0$$

which contradicts the previous inequality, thus $\{S_2 = T\}$ has measure 0. It is easy to check that $S_2 \leq \tau_2$ a.s., so $X_s^\lambda \leq Y_s^\lambda$ on $[0, S_2]$ and the assumptions on u imply $u(c_s^\lambda, X_s^\lambda) - u_s^\lambda > 0$. Summarizing we have

$$\begin{aligned} 0 = U_0^\lambda - V_0^\lambda &= E(\mathrm{e}^{-\beta S_2} (U_{S_2}^\lambda - V_{S_2}^\lambda) + \int_t^{S_2} \mathrm{e}^{-\beta(s-t)} [u(c_s^\lambda, X_s^\lambda) - u_s^\lambda] ds) \\ &= E(\int_t^{S_2} \mathrm{e}^{-\beta(s-t)} [u(c_s^\lambda, X_s^\lambda) - u_s^\lambda] ds) > 0 \end{aligned}$$

which is a clear contradiction. □

To develop the equilibrium analysis assuming that the rapresentative agent is characterized by a BFSDU we will need an explicit representation of the solution of (7)-(8), at least at time 0. A linear BFSDU presents no problems, but problems arise when nonlinearities are introduced. In our analysis we follow [Cvitanic and Ma, 1996]. Being the terminal condition of V a smooth function of c_T and Y_T, a functional link carries through at all times, so there exists a function θ such that $V_t = \theta(t, c_t, Y_t)$. This function may be characterized as the solution a nonlinear degenerate parabolic partial differential equation associated to the BFSDE.

To prove this fact, we adapt the techniques of [Cvitanic and Ma, 1996], where a more regular case is considered. The link between Backward or Forward-Backward SDE's and quasi-linear parabolic PDE's has been extensively studied by many authors (e.g. see [Ma et al., 1994, Duffie and Lions, 1992]).

Some more regularity on the coefficients is needed and so we strenghten our hypotheses with the following

Assumption 3.7 :

(i) β *is uniformly bounded by* M *and* $y_0 > 0,\ \alpha,\ \delta,\ \mu > 0$;

(ii) The functions $\mu, \sigma : [0,T] \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ *are differentiable, with derivatives uniformly bounded by a constant* k_1;

(iii) the function $g : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}^+ \cup \{0\}$ *is differentiable with partials uniformly bounded by a constant* k;

(iv) $u : \mathbb{R}^+ \times \mathbb{R} \longrightarrow \mathbb{R}^+$ *is differentiable, with* $\left|\frac{\partial u}{\partial y}\right| \leq k$ *(k as in (iii)) and* $\left|\frac{\partial u}{\partial x}\right| \leq k_2$, *for some* $k_2 > 0$ *for all* x, y;

(v) u, g *are strictly increasing and concave in c, strictly decreasing in y and concave in the couple.*

We first prove that the solution shows continuous dependence on the parameters. First, let us extend u and g to all $\mathbb{R}\times\mathbb{R}$ by continuity and let us take t, x, y varying in $[0,T]\times\mathbb{R}\times\mathbb{R}$. We consider the following flows associated with our equations

$$c_s^{t,x} = x + \int_t^s \mu(r, c_r^{t,x})dr + \int_0^t \sigma(r, c_r^{t,x})dW_r, \qquad c_t^{t,x} = x \tag{13}$$

$$Y_s^{t,x,y} = y + \int_t^s [\nu V_r^{t,x,y} + \eta c_r^{t,x} - \alpha Y_r^{t,x,y}]dr, \qquad Y_y^{t,x,y} = y \tag{14}$$

$$V_s^{t,x,y} = E(g(c_T^{t,x}, Y_T^{t,x,y}) + \int_s^T [u(c_r^{t,x}, Y_s^{t,x,y}) - \beta(r)V_r^{t,x,y}]dr + |\mathcal{F}_s). \tag{15}$$

For any fixed $t_1, t_2 \in [0,T]$, $x_1, x_2 \in \mathbb{R}^+$ and $y_1, y_2 \in [y_0 e^{-\alpha T}, +\infty)$, we denote

$$c^i_\cdot = c^{t_i,x_i}_{\cdot \vee t_i}, \quad Y^i_\cdot = Y^{t_i,x_i,y_i}_{\cdot \vee t_i}, \quad V^i_\cdot = V^{t_i,x_i,y_i}_{\cdot \vee t_i}, \qquad i = 1,2$$

where $s \vee t$ stands for $\max(s,t)$.

Proposition 3.8 : *Under Assumption 3.7, the above flows are continuous in t, x, y. More specifically, for given t_1 and x_1, there exists a constant C_1 depending only on k_1, T, t_1, x_1, $\mu(r,0)$, $\sigma(r,0)$, such that*

$$E(\sup_{s\in[0,T]} |c^2_s - c^1_s|^2) \le C_1(|x_2 - x_1|^2 + |t_2 - t_1|). \tag{16}$$

Moreover, for given t_1, x_1 and y_1, provided that $\sqrt{8}K(K+1)T < 1$, there exists a constant C_2, depending only on $k, k_1, k_2, T, t_1, x_1, y_1$ such that

$$E(\sup_{s\in[0,T]} [|Y^2_s - Y^1_s| + |V^2_s - V^1_s|]^2) \le \frac{C_2(|x_2 - x_1|^2 + |y_2 - y_1|^2 + |t_2 - t_1|)}{1 - \sqrt{8}(K^2 + K)T}. \tag{17}$$

Proof: We rely again on the argument explained in the previous propositions, therefore here we just sketch the proof.

By the Lipschitz property, it is easy to verify that

$$\begin{aligned} |c^2_s - c^1_s|^2 &\le 5|x_2 - x_1|^2 + 5(s \vee t_2 - t_2)\int_{t_2}^{s\vee t_2} k_1^2 |c^2_r - c^1_r|^2 dr + [\int_{t_1 \wedge s}^{t_2 \wedge s} |\mu(r, c^1_r)| dr]^2 \\ &+ 5[\int_{t_2}^{s \vee t_2} |\sigma(r, c^2_r) - \sigma(r, c^1_r)| dW_r]^2 + [\int_{t_1\wedge s}^{t_2 \wedge s} |\sigma(r, c^1_r)| dW_r]^2, \end{aligned}$$

taking expectations and applying Doob's inequality we get

$$\begin{aligned} E(\sup_{0\le s\le t} |c^2_s - c^1_s|^2) \le\; & 5|x_2 - x_1|^2 + 5k_1^2(|t - t_2| + 1)\int_{t_2}^{t \vee t_2} E(\sup_{0 \le s \le r} |c^2_s - c^1_s|^2) dr \\ & + 5|t_2 - t_1|(1 + |t_2 - t_1|)\left[||c^1||^2_{\underline{S}^2_{[t_1,t_2]}} + \max_{0\le r\le T}(|\mu(r,0)|^2 + |\sigma(r,0)|^2)\right] \end{aligned}$$

and using Gronwall's inequality, we are able to derive (16).

Similarly we can show

$$\begin{aligned} |Y^2_s - Y^1_s| \le |y_2 - y_1| + K\int_{t_2}^{s\vee t_2} [|V^2_r - V^1_r| + |Y^2_r - Y^1_r| + |c^2_r - c^1_r|] dr \\ + \; K\int_{s\wedge t_1}^{s \wedge t_2} [|V^1_r| + |Y^1_r| + |c^1_r|] dr. \end{aligned}$$

On the other hand, because of the martingale representation theorem, there exist two predictable processes Z_r^1 and Z_r^2 such that

$$V_s^i = g(c_T^i, Y_T^i) + \int_{s\vee t_i}^T [u(c_r^i, Y_r^i) - \beta(r)V_r^i]dr - \int_{s\vee t_i}^T Z_r^i dW_r, \quad \text{with} \quad E\left(\int_0^T |Z_r^i|^2 dr\right) < +\infty$$

Taking the conditional expectation with respect to $\mathcal{F}_t$, by the Lipschitz property of the coefficients, we obtain

$$|V_s^2 - V_s^1| \le E\Big(K[|c_T^2 - c_T^1| + |Y_T^2 - Y_T^1|] + \int_{s\vee t_2}^T \{K[|V_r^2 - V_r^1| + |Y_r^2 - Y_r^1|] + k_2|c_r^2 - c_r^1|\}dr$$
$$+ \int_{s\vee t_1}^{s\vee t_2} \{K[|Y_r^1| + |V_r^1|] + k_2|c_r^1| + |u(0,0)|\}dr|\mathcal{F}_s\Big).$$

Summing the two differences, we get

$$|Y_s^2 - Y_s^1| + |V_s^2 - V_s^1| \le (K+1)|y_2 - y_1| + E\left((K^2+K)\int_{t_2}^T [|Y_r^2 - Y_r^1| + |V_r^2 - V_r^1|]dr|\mathcal{F}_s\right)$$
$$+ E\left(K|c_T^2 - c_T^1| + (K^2 \vee k_2)\int_{t_2}^T |c_r^2 - c_r^1|dr|\mathcal{F}_s\right)$$
$$+ (2K + K^2 \vee k_2)E\left(\int_{t_1}^{t_2} [|Y_r^1| + |V_r^1| + |c_r^1| + |u(0,0)|]dr|\mathcal{F}_s\right).$$

Squaring both sides, applying Cauchy-Schwarz inequality and Doob's inequality, we obtain, for some constant C depending only on T, K, k_2,

$$\|\, |Y^2 - Y^1| + |V^2 - V^1|\, \|_{\underline{\underline{S}}^2}^2 \le \frac{C^2}{1 - 8T(K+K^2)^2}\Big\{|y_2 - y_1|^2 + E(\int_{t_2}^T |c_r^2 - c_r^1|^2 dr)$$
$$+ |t_2 - t_1| E(\int_{t_1}^{t_2} [|Y_r^1|^2 + |V_r^1|^2 + |c_r^1|^2 + |u(0,0)|^2]dr)\Big\}$$
$$\le \frac{C^2}{1 - 8T(K+K^2)^2}\Big\{|y_2 - y_1|^2 + T\|c^2 - c^1\|_{\underline{\underline{S}}^2}^2$$
$$+ |t_2 - t_1|^2[\|V^1\|_{\underline{\underline{S}}^2_{[t_1,t_2]}}^2 + \|Y^1\|_{\underline{\underline{S}}^2_{[t_1,t_2]}}^2 + \|c^1\|_{\underline{\underline{S}}^2_{[t_1,t_2]}}^2 + |u(0,0)|^2]\Big\}$$

which gives our thesis, by virtue of (16). □

By hypothesis, all the coefficients occurring in the previous equations are deterministic and differentiable. By the standard technique of time shift and because of Blumenthal's 0-1 law, it is possible to show that the functions

$$\gamma(t,x) = c_t^{t,x}, \qquad \phi(t,x,y) = Y_t^{t,x,y}, \qquad \theta(t,x,y) = V_t^{t,x,y}$$

are all deterministic. Proposition 3.8 tells us that these functions are locally Lipschitz in x, y and Hölder of order $\frac{1}{2}$ in t, consequently their derivatives are defined a.s. and bounded on compacts. It is to be noted that basically the same proof gives also the continuity of $U(C)$ in c, with respect to $\underline{\underline{S}}^2$ norm.

Our next goal is to prove that $\theta(t,x,y)$ is a viscosity solution of a degenerate semilinear parabolic PDE. This will enable us to deduce some useful properties of V. First we would like to remind the notion of viscosity solution for second order operators.

Definition 3.9 : *Let $L = L(t, \theta, D\theta, D^2\theta)$ be an elliptic (possibly degenerate) operator and let us consider the PDE problem in a certain domain $\mathcal{O} \subseteq [0,T] \times \mathbb{R}^2$*

$$\text{(18)} \qquad \begin{cases} \dfrac{\partial \theta}{\partial t} + L(t,\theta,D\theta,D^2\theta) = 0 \\ \theta(t,x,y) - g(x,y) = 0 \qquad (t,x,y) \in \partial\mathcal{O}. \end{cases}$$

$\theta \in C(\overline{\mathcal{O}})$ is said to be a viscosity sub- (resp. super-) solution of (18) if for any function $\varphi \in C^{1,2}(\overline{\mathcal{O}})$, taken any $(\bar{t},\bar{x},\bar{y}) \in \overline{\mathcal{O}}$, which is a global maximum point for $\theta - \varphi$, we have

$$\text{(19)} \qquad \begin{cases} \dfrac{\partial \varphi}{\partial t}(t,\bar{x},\bar{y}) + L(\bar{t},\bar{x},\bar{y},\theta(\bar{t},\bar{x},\bar{y}),D\varphi(\bar{t},\bar{x},\bar{y}),D^2\varphi(\bar{t},\bar{x},\bar{y})) \leq (\text{resp. } \geq) \, 0 \\ \theta(\bar{t},\bar{x},\bar{y}) - g(\bar{x},\bar{y}) \leq (\text{resp. } \geq) 0 \qquad \text{whenever} \quad (\bar{t},\bar{x},\bar{y}) \in \partial\mathcal{O}. \end{cases}$$

θ is said to be a solution of (18) if it is both a viscosity sub and super-solution.

Remark 3.10 : By the previous proposition, we have that the function $\theta(t,x,y) = V_t^{t,x,y}$ is indeed continuous in $[0,T] \times \mathbb{R} \times \mathbb{R}$, answering the first condition of viscosity solution.

Theorem 3.11 : *Under Assumptions 3.7, $\theta(t,x,y)$ is a viscosity solution of the PDE problem in $[0,T] \times [0,\infty) \times [y_0 e^{-\alpha T}, \infty)$,*

$$\text{(20)} \qquad \begin{cases} \dfrac{\partial \theta}{\partial t} + \dfrac{\sigma^2(t,x)}{2}\dfrac{\partial^2 \theta}{\partial x^2} + \mu(t,x)\dfrac{\partial \theta}{\partial x} + (\nu\theta + \eta x - \alpha y)\dfrac{\partial \theta}{\partial y} - u(x,y) + \beta(t)\theta = 0 \\ \theta(T,x,y) = g(x,y). \end{cases}$$

Proof: We remark that by construction, the processes $c_s^{t,x}, Y_s^{t,x,y}$ and $V_s^{t,x,y}$ have all continuous paths and they are adapted with respect to the filtration generated by the Brownian motion. Therefore by the Markov property and the pathwise uniqueness of the solution, it is possible to show that actually $V_s^{t,x,y} = \theta(s, c_s^{t,x}, Y_s^{t,x,y})$ a.s.

To show our statement, we need to prove that θ is both a sub and a super-solution of (20). As a matter of fact, we show only the sub-solution inequality, since the proof of the other goes along the same lines.

Let us consider a point $(t,x,y) \in [0,T] \times \mathbb{R} \times \mathbb{R}$ and a function φ such that

$$0 = \theta(t,x,y) - \varphi(t,x,y)$$

is a global maximum for $\theta - \varphi$ (without loss of generality we can assume this maximum to be zero).

This means that for any stopping time, necessarily

$$\theta(\tau, c_\tau^{t,x}, Y_\tau^{t,x,y}) - \varphi(\tau, c_\tau^{t,x}, Y_\tau^{t,x,y}) \leq 0. \tag{21}$$

For ease of writing, from now on we omit the superscripts of c, Y and V. Applying Itô's formula to φ in the interval $[t,\tau]$, because of the equations for c and Y, we have

$$\begin{aligned}
\varphi(\tau, c_\tau, Y_\tau) &= \varphi(t,x,y) + \int_t^\tau \sigma(r,c_r)\frac{\partial \varphi}{\partial x}(r,c_r,Y_r)dW_r \\
&+ \int_t^\tau \left[\frac{\partial \varphi}{\partial t}(r,c_r,Y_r) + \frac{\sigma^2}{2}(r,c_r)\frac{\partial^2 \varphi}{\partial x}(r,c_r,Y_r) + \mu(r,c_r)\frac{\partial \varphi}{\partial x}(r,c_r,Y_r)\right] dr \\
&+ \int_t^\tau \left[(\nu V_r + \eta c_r - \alpha Y_r)\frac{\partial \varphi}{\partial y}(r,c_r,Y_r)\right] dr.
\end{aligned}$$

On the other hand, because of (7), using the martingale representation theorem, we have

$$\begin{aligned}
\theta(t,x,y) = V_t &= V_\tau + \int_t^\tau [u(c_r,Y_r) - \beta(r)V_r]dr - \int_t^\tau Z_r dW_r \\
&= \theta(\tau, c_\tau, Y_\tau) + \int_t^\tau [u(c_r,Y_r) - \beta(r)V_r]dr - \int_t^\tau Z_r dW_r.
\end{aligned}$$

Substituting these last two equalities in (21), we obtain

$$\begin{aligned}
0 &\geq \theta(\tau, c_\tau, Y_\tau) - \varphi(\tau, c_\tau, Y_\tau) \\
&= \theta(t,x,y) - \varphi(t,x,y) + \int_t^\tau \left[\frac{\partial \varphi}{\partial t}(r,c_r,Y_r) + \frac{\sigma^2}{2}(r,c_r)\frac{\partial^2 \varphi}{\partial x}(r,c_r,Y_r)\right] dr \\
&+ \int_t^\tau \left[\mu(r,c_r)\frac{\partial \varphi}{\partial x}(r,c_r,Y_r) + (\nu V_r + \eta c_r - \alpha Y_r)\frac{\partial \varphi}{\partial y}(r,c_r,Y_r) - u(c_r,Y_r) + \beta(r)V_r\right] dr \\
&+ \int_t^\tau \left[Z_r - \sigma(r,c_r)\frac{\partial \varphi}{\partial x}(r,c_r,Y_r)\right] dW_r.
\end{aligned}$$

By the uniqueness of paths, we know that $V_r = \theta(r,c_r,Y_r)$, therefore substituting in the former expression we obtain

$$\begin{aligned}
&\int_t^\tau \left[\frac{\partial \varphi}{\partial t}(r,c_r,Y_r) + \frac{\sigma^2(r,c_r)}{2}\frac{\partial^2 \varphi}{\partial x}(r,c_r,Y_r) + \mu(r,c_r)\frac{\partial \varphi}{\partial x}(r,c_r,Y_r)\right. \\
&\qquad \left. + (\nu\theta(r,c_r,Y_r) + \eta c_r - \alpha Y_r)\frac{\partial \varphi}{\partial y}(r,c_r,Y_r) - u(c_r,Y_r) + \beta(r)\theta(r,c_r,Y_r)\right] dr \\
&+ \int_t^\tau \left[Z_r - \sigma(r,c_r)\frac{\partial \varphi}{\partial x}(r,c_r,Y_r)\right] dW_r \leq 0.
\end{aligned}$$

Taking expectations, the martingale part gives no contribution and we can summarize the inequality by writing

$$E\left(\int_t^T \Sigma(r, c_r, Y_r)dr\right) \le 0, \tag{22}$$

where $\Sigma(\cdot,\cdot,\cdot) = \frac{\partial \varphi}{\partial t} + L(\cdot,\cdot,\cdot,\theta(\cdot,\cdot,\cdot),\varphi(\cdot,\cdot,\cdot))$ and

$$\begin{aligned} L(t,x,y,\theta(t,x,y),\varphi(t,x,y)) &= \frac{1}{2}\sigma^2(t,x)\frac{\partial^2\varphi}{\partial x^2}(t,x,y) + \mu(t,x)\frac{\partial \varphi}{\partial x}(t,x,y) \\ &+ (\nu\theta(t,x,y) + \eta x - \alpha y)\frac{\partial \varphi}{\partial y}(t,x,y) - u(x,y) + \beta(t)\theta(t,x,y) \end{aligned}$$

To say that θ is a subsolution of (20) means that we must verify that $\Sigma(t,x,y) \le 0$. By contradiction we assume there exists an $\varepsilon_0 > 0$ such that $\Sigma(t,x,y) > \varepsilon_0$ and we define the stopping time

$$\tau_1 = \inf\{s > t : \Sigma(s, c_s, Y_s) \le \frac{\varepsilon_0}{2}\} \wedge T.$$

Since $\Sigma(t,x,y) > \varepsilon_0$, we have $\tau_1 > t$ a.s. Inequality (22) holds for any stopping time, therefore also for τ_1 and we have

$$0 < \frac{\varepsilon_0}{2}(\tau_1 - t) < E\left(\int_t^{\tau_1} \Sigma(s, c_s, Y_s)ds\right) \le 0$$

which is a clear contradiction, hence we proved that θ is a subsolution of (20). Analogously we can prove that θ is a viscosity super-solution of (20) and complete the proof. □

In [Citti, Pascucci, Polidoro, 1998] it is shown that if $\nu\theta_x + \eta \neq 0$ in $\mathcal{O}$ and σ μ are constant then θ is C^∞. If σ, μ, u are $C^{1,\alpha}$ then θ_{xx} is $C^{1,\alpha}$ and θ_y is C^α. About the sign of the partial derivatives of θ we have that at least for a compact subset $\mathcal{O}$ with T small enough then the sign of the partial derivatives of g is confirmed for the partial derivatives of θ and therefore the above condition on the partial derivative of θ with respect to x is satisfied.

4 Optimal Consumption and Equilibrium Analysis

In this Section, we present the optimal consumption problem and the equilibrium analysis for a representative agent characterized by a BFSDU.

The optimal consumption problem (maximization of U over the set of the admissible consumption-portfolio policies of Definition 2.1) can be handled via dynamic optimization techniques or via the martingale method (see [Cox and Huang, 1989 and 1991]). Here we follow the second approach which seems more appropriate for the BFSDU. To simplify the analysis, we assume β constant.

The optimal consumption problem of the representative agent is equivalent to the following constrained static maximization problem:

$$\max_{C,\pi} \; U(C) \qquad \text{under the constraint} \tag{23}$$

$$E^*\left(\int_0^T e^{-\int_o^t r_s ds} c_t dt + e^{-\int_0^T r_s ds} X_T\right) \leq E^*\left(\int_0^T e^{-\int_o^t r_s ds} e_t dt\right). \tag{24}$$

E^* denotes expectation under the equivalent martingale measure nested in the complete financial market model, that is $E^*(\cdot) = E(\psi_T \cdot)$, where ψ is defined by (1), while $U(C) = V_0$, the unknown initial value of the backward component in the system (4)-(5) or (7)-(8).

We further specify our setting assuming that the endowment process is given by

$$de_t = \mu_t^e dt + \sigma_t^e dW_t, \tag{25}$$

with Lipschitz and predictable coefficients μ_t^e and σ_t^e in $\mathbb{R}$ and constant initial condition $e_0 > 0$.

When considering the Additive Expected Utility or the Backward Stochastic Differential one, the constrained maximization problem is solved by exploiting the first order necessary conditions for optimality and the concavity of U. In the AEU setting (see for example [Duffie, 1996, p. 205-208]), the problem is solved through the associated Lagrangean. The consumption plan obtained from the first order necessary conditions associated with the Lagrangean is parametrized with respect to the Lagrange multiplier, the multiplier is determined by imposing that the consumption plan satisfies the budget constraint. Thanks to the Inada conditions on the utility function a unique Lagrange multiplier is determined and therefore the optimal consumption plan is defined. For the BSDU basically the same procedure can be followed. [Duffie and Epstein and Skiadas 1992] prove that the concavity of the aggregator implies the concavity of the Utility with respect to c, this fact is used to prove that the first order conditions are sufficient conditions for the optimum. Again, the constrained maximization problem is solved through the associated Lagrangean by means of a standard saddle point theorem and the Lagrange multiplier satisfying the budget constraint is determined exploiting the Inada conditions.

When habit formation is introduced in the stochastic differential Utility, existence of a solution for the consumption problem is equivalent to showing existence of the solution of a Backward-Forward SDE and some conditions on the regularity of the inverse of the marginal utility and/or restrictions on the state price process are needed, see [Detemple and Zapatero, 1992]. We want to extend these techniques to solve the consumption problem for an agent characterized by a BFSDU. Proposition 3.6 guarantees the concavity of the BFSDU. We address the existence of a solution satisfying the first order necessary conditions and the characterization of the equilibrium prices process in the next two Sections. Here we focus our attention on the representation of the utility gradient of the BFSDU. As explained in [Duffie and Skiadas, 1994], provided the optimal consumption exists, the Arrow-Debreu equilibrium price process can be characterized by means of the Gateaux derivative of $U(C)$ and its Riesz representation evaluated along the endowment process.

Given a reference pair of cumulative consumption and trading strategy, $(\overline{\pi}, \overline{C})$, and a set F of feasible directions, the Gateaux derivative of $U(C)$ at $(\overline{\pi}, \overline{C})$ is defined as the functional

$$\nabla U(\overline{C}; C) = \lim_{\alpha \to 0} \frac{U(\overline{C} + \alpha C) - U(\overline{C})}{\alpha}, \quad C \in F.$$

We say that $\nabla U(\overline{C}; C)$ admits a Riesz representation if there exists a process γ_t such that

$$\nabla U(\overline{C}; C) = E(\int_0^T (c_t - \overline{c}_t)\gamma_t dt).$$

In [Duffie and Skiadas, 1994, Proposition 2] it is shown that γ_t represents the Arrow-Debreu price process if $\overline{C}$ is the optimal consumption policy and it coincides with the endowment process. In the same paper, conditions for the existence of the Gateaux derivative are provided and the Riesz representation is computed for some utility functions, but in general, those conditions do not apply to BFSDU.

The linear case, treated in the next section, presents no difficulties, since the hypotheses will determine an explicit Riesz representation. In the nonlinear case, differentiating formally (7) and (8), we obtain that the the Gateaux derivatives $\nabla V_t(\overline{C}; C)$ and $\nabla Y_t(\overline{C}; C)$ have to verify the system

$$\begin{aligned} \nabla V_t(\overline{C}; C) &= E(g_x(c_T, Y_T)c_T + g_y(c_T, Y_T)\nabla Y_T(\overline{C}; C) \\ &+ \int_t^T (u_x(c_s, Y_s)c_s + u_y(c_s, Y_s)\nabla Y_s(\overline{C}; C) - \beta_s \nabla V_s(\overline{C}; C))ds|\mathcal{F}_t), \end{aligned} \tag{26}$$

$$\nabla Y_t(\overline{C}; C) = \int_0^t [\nu \nabla V_s(\overline{C}; C) + \eta c_s - \alpha \nabla Y_s(\overline{C}; C)]ds. \tag{27}$$

The existence of this pair is assured under the same assumptions as in Proposition 3.4. To study the equilibrium price by means of the Riesz representation of the Gateaux derivative, we would need an explicit representation of the solution of linear BFSDE's (26)-(27), but such closed formula is not present in the literature. Nevertheless, for this linear model we are able to obtain an explicit representation of the Gateaux derivative at time 0, while in the nonlinear setting we recover an almost explicit representation of the gradient by exploiting the functional link between V and c, Y. This will be shown in detail in Section 6.

Finally, we want to recall the equilibrium analysis results, when using the Additive Expected Utility, since those will serve as our main reference. Setting $\xi_t = e^{\int_0^t \beta - r_u du}\psi_t$, the Arrow-Debreu price process adjusted by the preference discount factor, the first order necessary conditions for the AEU evaluated along the endowment process provide the following normalized Arrow-Debreu price process ξ_s (we set the Lagrange multiplier equal to 1)

$$u'(e_s) = \xi_s \quad \text{and} \quad v'(X_T) = \xi_T.$$

To ensure that the price process ξ belongs to $\mathcal{L}^2_+$, it is enough to assume that the endowment process is bounded away from zero and that the utility function satisfies the standard Inada conditions, see [Duffie and Zame, 1989]. The price process ξ_t contains many interesting pieces of information about the economy. In particular, the equilibrium interest rate is given by the negative expected growth rate of ξ_t, while the market prices of risk are the negative of the volatility in the growth of ξ_t. Therefore one obtains that the equilibrium interest rate and the market price of risk are

$$\begin{aligned} r_t &= \beta - (\xi_t)^{-1}(u''(e_t)\mu_t^e + \frac{1}{2}u'''(e_t)\sigma_t^{e2}), \\ \mu_t - r_t\mathbf{1} &= -\sigma_t \frac{u''(e_t)}{u'(e_t)}\sigma_t^e = -\beta_t^e Cov(dG_t/S_t, de_t), \end{aligned}$$

where $\beta_t^e = \frac{u''(e_t)}{u'(e_t)}$.

5 Linear Backward-Forward SDU

In this section we consider a linear BFSDU, namely

$$\begin{aligned} V_t &= E(\int_t^T [u(c_s) - \gamma Y_s - \beta V_s]ds + \Gamma|\mathcal{F}_t) \\ Y_t &= y_0 e^{-\alpha t} + \delta \int_0^t e^{-\alpha(t-s)}[\mu V_s + (1-\mu)c_s]ds, \end{aligned}$$

where $\alpha, \beta, \gamma, \delta, \mu$ are all positive constants. Similarly to [Constantinides, 1990], where the istantaneous utility at time t is a function of the difference between consumption and the habit at time t, we assume that the habit affects the instantaneous utility from consumption negatively and linearly. To simplify the notation, we take $\Gamma \equiv 0$, $\nu = \delta\mu$ and $\eta = \delta(1-\mu)$, so we have

$$V_t = E(\int_t^T [u(c_s) - \gamma Y_s - \beta V_s] ds | \mathcal{F}_t) \tag{28}$$

$$Y_t = y_0 + \int_0^t [\nu V_s + \eta c_s - \alpha Y_s] ds. \tag{29}$$

Existence of the utility process is assured if Assumption 3.1 and the Assumption in Proposition 3.2 are verified by $u(c) - \gamma y$. From (28)-(29), it is possible to find an explicit expression of V_0.

For the time being let us treat Y_T as given, so we may rewrite the above system in backward form

$$\begin{pmatrix} V_t \\ Y_t \end{pmatrix} = E\left(\int_t^T \left[A \begin{pmatrix} V_s \\ Y_s \end{pmatrix} + \begin{pmatrix} u(c_s) \\ -\eta c_s \end{pmatrix} \right] ds + \begin{pmatrix} 0 \\ Y_T \end{pmatrix} \middle| \mathcal{F}_t \right),$$

where the matrix $A = \begin{pmatrix} -\beta & -\gamma \\ -\nu & \alpha \end{pmatrix}$ is made up of constants. The solution V, Y can be explicitly written in terms of c, Y_T as

$$\begin{pmatrix} V_t \\ Y_t \end{pmatrix} = E\left(\int_t^T \mathrm{e}^{A(s-t)} \begin{pmatrix} u(c_s) \\ -\eta c_s \end{pmatrix} ds + \mathrm{e}^{A(T-t)} \begin{pmatrix} 0 \\ Y_T \end{pmatrix} \middle| \mathcal{F}_t \right),$$

where the matrix $\mathrm{e}^{A(s-t)}$ is intended to be

$$\mathrm{e}^{A(s-t)} = \sum_{n=0}^{\infty} \frac{((s-t)A)}{n!}.$$

Therefore we have

$$V_t = E(\int_t^T (\mathrm{e}_{11}^{A(s-t)} u(c_s) - \mathrm{e}_{12}^{A(s-t)} \eta c_s) ds + \mathrm{e}_{12}^{A(T-t)} Y_T | \mathcal{F}_t)$$

$$Y_t = E(\int_t^T (\mathrm{e}_{21}^{A(s-t)} u(c_s) - \mathrm{e}_{22}^{A(s-t)} \eta c_s) ds + \mathrm{e}_{22}^{A(T-t)} Y_T | \mathcal{F}_t),$$

where e_{ij}^{At} denotes the $ij-$th element $(i, j = 1, 2)$ of the matrix e^{At}. Solving the last equation in $t = 0$ and recalling that $Y_0 = y_0$, we obtain

$$E(Y_T) = \frac{y_0}{e_{22}^{AT}} - E\left(\int_0^T \frac{\mathrm{e}_{21}^{As} u(c_s) - \mathrm{e}_{22}^{As} \eta c_s}{\mathrm{e}_{22}^{AT}} ds \right), \tag{30}$$

provided that the denominator is different from zero. Hence substituting $E(Y_T)$ in the expression of V_0 we get

$$U(C) = V_0 = E\left(\int_0^T (\mathrm{e}_{11}^{As} u(c_s) - \mathrm{e}_{12}^{As} \eta c_s - \mathrm{e}_{12}^{AT} \frac{\mathrm{e}_{21}^{As} u(c_s) - \mathrm{e}_{22}^{As} \eta c_s}{\mathrm{e}_{22}^{AT}}) ds + \frac{e_{12}^{AT}}{e_{22}^{AT}} y_0 \right).$$

We can compute the Gateaux derivative of $U(C)$ at a consumption process c_s along a feasible direction and we can find its Riesz representation γ_s, given by

$$\gamma_s = \frac{e_{11}^{As}e_{22}^{AT} - e_{12}^{AT}e_{21}^{As}}{e_{22}^{AT}} u'(c_s) + \frac{e_{12}^{AT}e_{22}^{As} - e_{12}^{As}e_{22}^{AT}}{e_{22}^{AT}} \eta.$$

Setting

$$H_s = \frac{e_{11}^{As}e_{22}^{AT} - e_{12}^{AT}e_{21}^{As}}{e_{22}^{AT}}, \qquad K_s = \frac{e_{12}^{AT}e_{22}^{As} - e_{12}^{As}e_{22}^{AT}}{e_{22}^{AT}},$$

we can briefly rewrite

$$\gamma_s = H_s u'(c_s) + \eta K_s.$$

H and K are differentiable functions of the time, by h and k we denote their derivatives. In Appendix A.1 we analyze the coefficients H, K and their derivatives. In particular, given our parameter conditions, we show that $H_s > 0$ and $K_s < 0$, $\forall s \in [0, T]$, while the reverse is true for the derivatives, $h_s < 0$ and $k_s > 0$, $\forall s \in [0, T]$.

To ensure existence of the optimal consumption policy and of a well behaved Arrow-Debreu price process we impose the following conditions, see [Detemple and Zapatero, 1991, Detemple and Zapatero, 1992].

Assumption 5.1 *The following conditions are satisfed:*

- $u(\cdot) : [0, \infty) \to (0, \infty)$, *is three times continuously differentiable, strictly increasing and strictly concave,* $\lim_{c\to 0} u'(c) = \infty$ *and* $\lim_{c\to\infty} u'(c) = 0$*;*
- *In equilibrium* ($c_t = e_t$, $\forall t \in [0, T]$) *we have*

$$\frac{e_{11}^{As}e_{22}^{AT} - e_{12}^{AT}e_{21}^{As}}{e_{22}^{AT}} u'(e_s) + \frac{e_{12}^{AT}e_{22}^{As} - e_{12}^{As}e_{22}^{AT}}{e_{22}^{AT}} \eta > 0, \ \forall s \in [0, T];$$

- $e_s >> 0$, $\forall s \in [0, T]$.

The first condition is the standard concavity-Inada conditions. The second and the third conditions ensure that the Arrow-Debreu price process belongs to $\mathcal{L}_+^2$. The last condition of Assumption 5.1 can be weakened to the following

- For each fixed $n \in \mathbb{N}$, the function u' is Lipschitz with constant k_n on $[\frac{1}{n+1}, \frac{1}{n}]$ such that $k_n \leq k_{n+1} \leq \ldots \to +\infty$ and

$$\int_0^T \sum_{n=1}^{+\infty} (\frac{k_n}{n})^2 P(\frac{1}{n+1} \leq e_s \leq \frac{1}{n}) ds < +\infty.$$

Therefore, even if e might be not bounded away from zero, this condition still implies that $E(\int_0^T |u'(e_s)|^2 ds) < +\infty$ and that $\xi \in \mathcal{L}^2$.

Thanks to Assumption 5.1, as proven in Proposition 3.6, we know that U is concave in C and also that the inverse of the marginal utility is well defined for every value, therefore we can apply a procedure similar to the one employed for the AEU to prove existence of the optimal solution. Given γ and fixed a positive Lagrange multiplier ρ, the consumption policy parametrized by ρ is obtained through the inverse of the marginal utility, then the Lagrange multiplier ρ is determined through the budget constraint. A unique positive solution for ρ is obtained thanks to the Inada conditions on the utility function.

Assuming market equilibrium, the optimal consumption must coincide with the endowment, $e_s = c_s^*$, $\forall s \in [0,T]$, so we obtain the following characterization of the Arrow-Debreu price process for the one consumer economy:

$$\text{(31)} \qquad \mathrm{e}^{-\int_0^s r_u du}\psi_s = \mathrm{e}^{-\beta s}\xi_s = H_s u'(e_s) + \eta K_s,$$

where we set the Lagrange multiplier equal to 1, by rescaling the price process. The equilibrium price is made up of two components. The first one is related to the istantaneous marginal utility, the second one to η. The equilibrium price process obtained in this setting can be compared to the one obtained with an AEU. Being $H_s e^{\beta s} \le 1$ and $K_s < 0 \; \forall s \in [0,T]$, we have that ceteris paribus, i.e. for a given the istantaneous utility function u and endowment process, the equilibrium prices process with a linear BFSDU is smaller than the equilibrium price process with the AEU.

Differentiating both sides of (31) we obtain the following

Proposition 5.2 *Let Assumptions 3.1 and 5.1 be satisfied then the interest rate of equilibrium has the following expression:*

$$r_t = -(e^{-\beta t}\xi_t)^{-1}[h_t u'(e_t) + H_t(\,\mu_t^e u''(e_t) + \frac{1}{2}(\sigma_t^e)^2 u'''(e_t)\,) + \eta k_t];$$

the market prices of risk is the following:

$$\text{(32)} \qquad \mu_t - r_t \mathbf{1} = -\sigma_t (e^{-\beta t}\xi_t)^{-1} H_t u''(e_t)\sigma_t^e.$$

We observe that this procedure can be applied also when the matrix A is time varying, but still deterministic. We would like to remark the similarity of these results with those for the standard AEU, which is in fact included by our model when we take $\gamma = \nu = 0$.

A (single factor) Consumption CAPM similar to the one associated with the AEU is obtained

$$\mu_t - r_t\mathbf{1} = -\beta_t^e Cov(dG_t/S_t, de_t),$$

where $\beta_t^e = \frac{H_t u''(e_t)}{e^{-\beta t}\xi_t}$.

Given the istantaneous utility function, the right hand side in (32) and therefore the risk premium results higher than for the AEU. This is easily seen, since for a given endowment process e_s and istantaneous utility function u we have

$$\frac{H_s u''(e_s)}{H_s u'(e_s) + \eta K_s} \leq \frac{u''(e_s)}{u'(e_s)}, \quad \forall s \in [0, T],$$

being $K_s < 0$. So we have shown that if we replace the time additivity with a linear backward-forward habit, then the risk premium goes up providing us with a solution for the equity premium puzzle.

Let us analyze the interest rate of equilibrium. As for the AEU, this consists of three components. The first one comes from the agent discount factor, which is simply β in the AEU framework and $-e^{\beta t}(\frac{h_t}{\xi_t}u'(e_t) + \eta\frac{k_t}{\xi_t})$ assuming the presence of the habit. The second component is related to the expected growth rate in consumption (the interest rate is positively related to it), the introduction of the habit amplifies this term, since $\frac{H_s}{H_s u'(e_s) + \eta K_s} > \frac{1}{u'(e_s)}$. The last component is related to the expected variance of consumption growth (the interest rate is negatively related to it if $u'''(e_t) > 0$). Again the habit formation has a magnifying effect on this term.

6 The $V_t = \theta(t, c_t, Y_t)$ case

In this section we solve the optimal consumption problem and we develop the equilibrium analysis when $u(c, y)$ is not linear in y. To simplify the analysis we assume again α, β, δ and μ to be constant. Let Assumption 3.7 hold in this section. Under this Assumption, the system (7)-(8) has a solution such that $V_t = \theta(t, c_t, Y_t)$ and θ has bounded first order partials, thus the Gateaux derivative of V is

$$\nabla V_t(\overline{C}; C) = \theta_x(t, \overline{c}_t, Y_t(\overline{C}))c_t + \theta_y(t, \overline{c}_t, Y_t(\overline{C}))\nabla Y_t(\overline{C}; C). \tag{33}$$

To give a complete characterization of the utility gradient we need to identify ∇Y_t. Since Y_t is the solution of

$$Y_t = y_0 + \int_0^t [\nu\theta(s, c_s, Y_s) + \eta c_s - \alpha Y_s] ds \tag{34}$$

we obtain that ∇Y_t has to verify

$$\nabla Y_t = \int_0^t [\nu\theta_x(s,\overline{c}_s,\overline{Y}_s)c_s + \nu\theta_y(s,\overline{c}_s,\overline{Y}_s)\nabla Y_s + \eta c_s - \alpha\nabla Y_s]ds.$$

(for simplicity we wrote $\overline{Y} = Y(\overline{C})$ and we omitted the argument $(\overline{C};C)$). Because of the boundedness of the derivatives of θ, this equation is well defined and has solution

$$\nabla Y_t = \int_0^t \exp\{\int_s^t (\nu\theta_y(r,\overline{c}_r,\overline{Y}_r) - \alpha)dr\}(\nu\theta_x(s,\overline{c}_s,\overline{Y}_s) + \eta)c_s ds.$$

On the other hand, from (7) we have that the Gateaux derivative of V_t satisfies

$$\nabla V_t = E(g_x(\overline{c}_T,\overline{Y}_T)c_T + g_y(\overline{c}_T,\overline{Y}_T)\nabla Y_T + \int_t^T (u_x(\overline{c}_s,\overline{Y}_s)c_s + u_y(\overline{c}_s,\overline{Y}_s)\nabla Y_s - \beta_s\nabla V_s)ds|\mathcal{F}_t),$$

whose solution is given by

$$\nabla V_t = E\left(e^{-\int_t^T \beta_r dr}[g_x(\overline{c}_T,\overline{Y}_T)c_T + g_y(\overline{c}_T,\overline{Y}_T)\nabla Y_T] + \int_t^T e^{-\int_t^s \beta_r dr}[u_x(\overline{c}_s,\overline{Y}_s)c_s + u_y(\overline{c}_s,\overline{Y}_s)\nabla Y_s]ds|\mathcal{F}_t\right).$$

In $t=0$ we have

$$\nabla V_0 = E\left(e^{-\int_0^T \beta_r dr}[g_x(\overline{c}_T,\overline{Y}_T)c_T + g_y(\overline{c}_T,\overline{Y}_T)\nabla Y_T] + \int_0^T e^{-\int_0^s \beta_r dr}[u_x(\overline{c}_s,\overline{Y}_s)c_s + u_y(\overline{c}_s,\overline{Y}_s)\nabla Y_s]ds\right)$$

and substituting (34) in the previous expression, we obtain

$$\begin{aligned}\nabla V_0 = E\Big(& e^{-\int_0^T \beta_r dr} g_x(\overline{c}_T,\overline{Y}_T)c_T \\ &+ e^{-\int_0^T \beta_r dr} g_y(\overline{c}_T,\overline{Y}_T)\int_0^T \exp\{\int_s^T (\nu\theta_y(r,\overline{c}_r,\overline{Y}_r) - \alpha)dr\}(\nu\theta_x(s,\overline{c}_s,\overline{Y}_s) + \eta)c_s ds \\ &+ \int_0^T e^{-\int_0^r \beta_v dv}\int_0^r \exp\{\int_s^r (\nu\theta_y(v,\overline{c}_v,\overline{Y}_v) - \alpha)dv\}(\nu\theta_x(s,\overline{c}_s,\overline{Y}_s) + \eta)c_s ds u_y(\overline{c}_r,\overline{Y}_r)dr \\ &+ \int_0^T e^{-\int_0^s \beta_r dr} u_x(\overline{c}_s,\overline{Y}_s)c_s ds\Big).\end{aligned}$$

Applying Fubini's theorem, we can conclude that the Gateaux derivative of $U(C)$ at the reference consumption process $\overline{C}$ is

$$\begin{aligned}\nabla U(\overline{C};C) = E\Big(& e^{-\int_0^T \beta_r dr} g_x(\overline{c}_T,\overline{Y}_T)c_T \\ &+ e^{-\int_0^T \beta_r dr} g_y(\overline{c}_T,\overline{Y}_T)\int_0^T e^{\int_s^T (\nu\theta_y(r,\overline{c}_r,\overline{Y}_r)-\alpha)dr}(\nu\theta_x(s,\overline{c}_s,\overline{Y}_s) + \eta)c_s\, ds \\ &+ \int_0^T\int_s^T e^{\int_0^r -\beta_v dv + \int_s^r (\nu\theta_y(v,\overline{c}_v,\overline{Y}_v)-\alpha)dv} u_y(\overline{c}_r,\overline{Y}_r)\, dr(\nu\theta_x(s,\overline{c}_s,\overline{Y}_s) + \eta)c_s ds \\ &+ \int_0^T e^{\int_0^s -\beta_r dr} u_x(\overline{c}_s,\overline{Y}_s)c_s ds\Big).\end{aligned}$$

When $g = 0$ and β is constant, the Riesz representation of $\nabla U(\overline{C}; C)$ is therefore

$$\gamma_s = e^{-\beta s} u_x(\overline{c}_s, \overline{Y}_s) + (\nu\theta_x(s, \overline{c}_s, \overline{Y}_s) + \eta) E(\int_s^T \mathrm{e}^{-\beta r} \mathrm{e}^{-\alpha(r-s)} \mathrm{e}^{\int_s^r \nu\theta_y(v, \overline{c}_v, \overline{Y}_v) dv} u_y(\overline{c}_r, \overline{Y}_r) dr | \mathcal{F}_s).$$

It remains to prove existence of the optimal consumption. To this end we apply the two steps procedure illustrated above and adapted to habit formation utility functions in [Detemple and Zapatero, 1992]. Being $\overline{c}_s, \overline{Y}_s$ generic, we denote them simply by c_s, Y_s. From Proposition 3.6, we know that $U = V_0$ is concave in C, therefore it is enough to prove that a consumption plan exists satisfying the first order necessary conditions and the budget constraint.

We denote by ζ_s, the monetary cost of marginal consumption augmented by the expected incremental impact on future utilities

$$\zeta_s \;=\; \rho\gamma_s - (\nu\theta_x(s, c_s, Y_s) + \eta) E(\int_s^T \mathrm{e}^{-\beta r + \int_s^r (\nu\theta_y(v, c_v, Y_v) - \alpha) dv} u_y(c_r, Y_r) dr | \mathcal{F}_s)$$

where ρ is a positive constant. We define $I(s, z, y)$ the inverse of $e^{-\beta s} u_x(x, y)$ with respect to the first argument, that is to say $e^{-\beta s} u_x(I(s, z, y), y) = z$, and $I^+(s, z, y) = \max\{0, I(s, z, y)\}$. By the concavity of u, we say that $c_s^* = I^+(\zeta_s, Y_s)$ is optimal if (ζ_s, Y_s) is the solution of

$$\zeta_s \;=\; \rho\gamma_s - (\nu\theta_x(s, I^+(s, \zeta_s, Y_s), Y_s) + \eta) \times \tag{35}$$
$$E(\int_s^T \mathrm{e}^{-\beta r + \int_s^r (\nu\theta_y(v, I^+(v, \zeta_v, Y_v), Y_v) - \alpha) dv} u_y(I^+(r, \zeta_r, Y_r), Y_r) dr | \mathcal{F}_s)$$

$$Y_s \;=\; y_0 \mathrm{e}^{-\alpha s} + \int_0^s \mathrm{e}^{-\alpha(s-r)} [\nu\theta(r, I^+(r, \zeta_r, Y_r), Y_r) + \eta I(r, \zeta_r, Y_r)] dr \tag{36}$$

and the budget constraint (24) is satisfied.

The system (35)-(36) is again a BFSDE and we have to prove existence and uniqueness of the solution for any given $\rho > 0$. We first state

Lemma 6.1 *Let $f, g, h, b(t, x, z)$ be uniformly Lipschitz functions with constant k, f and g are uniformly bounded by a constant M and $h(s, 0, 0)$, $b(s, 0, 0)$ are in $L^2([0, T])$. Let J_t be a process in $\mathcal{L}^2$.*

If T is small enough, depending on M and k, then there exists a unique adapted solution in $\mathcal{L}^2 \times \mathcal{L}^2$ of the BFSDE

$$x_t \;=\; J_t - f(t, x_t, z_t) E(\int_t^T \mathrm{e}^{\int_t^s g(u, x_u, z_u) du} h(s, x_s, z_s) ds | \mathcal{F}_t)$$
$$z_t \;=\; z_0 + \int_0^t b(s, x_s, z_s) ds$$

Proof: The proof of this Lemma is similar to that of Proposition 3.2, based on Doob's inequality and the smallness of the time interval, so to guarantee that the operator induced by the system will be a contraction on $\mathcal{L}^2 \times \mathcal{L}^2$. □

If we assume

Assumption 6.2 *$u(c,y)$ is such that $\lim_{c\to 0} u_x(c,y) \leq \infty \ \forall y > 0$ and $\lim_{c\to\infty} u_x(c,y) = 0 \ \forall y > 0$. Let $I^+(s,z,y)$ and $u_y(I^+(s,z,y),y)$ be bounded and uniformly Lipschitz in z and y, uniformly for $y \geq y_0 e^{-\alpha T}$, $z > 0$ and $s \geq 0$.*

Then we can set $(x_t, z_t) = (\zeta_t, Y_t)$ and

$$\begin{aligned} J_t &= \rho\xi_t, \qquad f(t,\zeta,y) = (\nu\theta_x(t,I(\zeta,y),y) + \eta), \ g(t,\zeta,y) = \nu\theta_y(t,I(\zeta,y)) - \alpha), \\ z_0 &= y_0 e^{-\alpha t}, \ h(t,\zeta,y) = \mathrm{e}^{-\beta t} u_y(I(\zeta,y),y), \ b(t,\zeta,y) = \nu\theta(t,I(\zeta,y),y) + \eta I(\zeta,y) - \alpha\zeta, \end{aligned}$$

and apply the Lemma. The regularity conditions required in the Lemma are guaranteed by Assumptions 3.7, 6.2 and by the results proved in [Citti, Pascucci, Polidoro, 1998].

For each positive Lagrange multiplier ρ we find the corresponding triple consumption, monetary cost, habit $c_{.}(\rho), \zeta_{.}(\rho), Y_{.}(\rho)$. As in [Detemple and Zapatero, 1992, Assuption 3.5] we assume that the composition map $I(s, \zeta_s(\rho), Y_s(\rho))$ is continuous in ρ, a.e. ω and all s, and satisfies the condition $\lim_{\rho\downarrow 0} I(s,\zeta_s(\rho),Y_s(\rho)) = \infty$ a.e. ω and all s. This guarantees that the equation obtained from the budget constraint (24)

$$E^*[\int_0^t \exp(-\int_0^t r_s ds)(c_t(\rho) - e_t)dt] = 0$$

admits a positive solution for ρ yielding the optimal solution c^*.

In equilibrium ($c_t = e_t$, $\forall t \in [0,T]$), we have that γ_t becomes the Arrow-Debreu price process:

$$\xi_s = u_x(e_s, Y_s^e) + (\nu\theta_x(s,e_s,Y_s^e) + \eta)E(\int_s^T \mathrm{e}^{-(\beta+\alpha)(r-s)} \mathrm{e}^{\int_s^r \nu\theta_y(v,e_v,Y_v^e)dv} u_y(e_r, Y_r^e)dr|\mathcal{F}_s),$$

where Y^e is the habit corresponding to e. Let the following Assumption hold.

Assumption 6.3 : *For all $s \in [0,T]$*

$$u_x(e_s, Y_s^e) + (\nu\theta_x(s,e_s,Y_s^e) + \eta)E(\int_s^T \mathrm{e}^{-(\beta+\alpha)(r-s)} \mathrm{e}^{\int_s^r \nu\theta_y(v,e_v,Y_v^e)dv} u_y(e_r, Y_r^e)dr|\mathcal{F}_s) > 0.$$

This hypothesis ensures that the Arrow-Debreu price process is strictly positive and guarantees the uniform properness of preferences. The boundedness of the derivatives implies that this process is certainly in $\mathcal{L}^2$.

As for the utility functional proposed in [Detemple and Zapatero, 1991], the process ξ_s is made up of two components. The first one (positive) is connected to the marginal utility from instantaneous consumption, the second one (negative being $u_y < 0$ and $\theta_x > 0$) is related to the future disutility of consumption due to an increase today in the agent's habit. This means that an increase in consumption determines a positive increase in the instantaneous utility $u_x(e_s, \overline{Y}_s)$ and a decrease in all future utilities given by

$$(\nu\theta_x(s, \overline{c}_s, \overline{Y}_s) + \eta)E\Big(\int_s^T e^{-(\beta+\alpha)(r-s)} e^{\int_s^r \nu\theta_y(u,\overline{c}_u,\overline{Y}_u)du}\ u_y(\overline{c}_r, \overline{Y}_r)\ dr|\mathcal{F}_s\Big).$$

We point out two main differences with respect to the pure backward habit analyzed in [Detemple and Zapatero, 1991]. First, the effect on the habit caused by a marginal increase in consumption is not δ as in [Detemple and Zapatero, 1991], but it is given by this factor multiplied by a convex combination of 1 and of θ_x. A convex combination that takes into account the backward-forward feature of the utility function. Second, the discount factor for future disutility is no longer $e^{-(\beta+\alpha)(r-s)}$, there is also the integral of θ_y. Therefore the effect of our characterization of the habit process instead of the classical habit on the equilibrium price process depends on the sign of θ_y and of $\theta_x - 1$. As stressed above, at least for small T we have that $\theta_y < 0$ and therefore this component leads to a higher equilibrium price than in the case of a pure backward habit. Nothing can be said about $\theta_x - 1$. If $\theta_x > 1$ then all the future utilities will be affected negatively with a magnitude higher than in the pure backward case, if $\theta_x < 1$ the opposite effect is obtained. However, ceteris paribus, the equilibrium price process for a BFSDU is smaller than the one obtained with an AEU.

In the following Proposition, whose proof is in Appendix A.2, we characterize the equilibrium interest rate and the market price of risk when the endowment process is

$$de_t = e_t(\bar{\mu}_t^e dt + \bar{\sigma}_t^e dW_t) \quad e_o = 1 \tag{37}$$

with deterministic coefficients $\bar{\mu}_t^e$ and $\bar{\sigma}_t^e$. Paying a higher computational cost, the results can be extended to the case of stochastic coefficients. In what follows we denote by $u(t), \theta(t)$ the functions evaluated along the endowment process e_t, the same notation is employed for their partial derivatives. The equilibrium analysis is summarized in the following

Proposition 6.4 *Let Assumptions 3.7, 6.2, 6.3 hold and assume that u is three times continuously differentiable, then the market interest rate of equilibrium is given by*

$$r_t = \beta - \xi_t^{-1}\left[\frac{1}{2}\bar{\sigma}_t^{e2} e_t^2 u_{xxx}(t) + +\bar{\mu}_t^e e_t u_{xx}(t) + (\nu\theta(t) + \eta e_t - \alpha Y_t)u_{xy}(t) - (\nu\theta_x(t) + \eta)u_y(t)\right]$$

$$-\xi_t^{-1}E\left(\int_t^T e^{-\beta(s-t)}e^{\int_t^s(\nu\theta_y(r)-\alpha)dr}u_y(s)ds|\mathcal{F}_t\right)\times$$
$$[-\nu\theta_x(t)\bar{\mu}_t^e-(\nu\theta_x(t)+\eta)(2\nu\theta_y(t)-\alpha)+\beta\eta+\nu(u_x(t)-e_t\bar{\sigma}_t^{e2}\theta_{xx}(t))]$$
$$-\xi^{-1}(t)\nu e^{-\int_0^t(\nu\theta_y(r)-\alpha-\beta)dr}e_t\bar{\sigma}_t^{e2}\theta_{xx}(t)E(A_T-A_t|\mathcal{F}_t).$$

The market price of risk is

$$\mu_t - r_t\mathbf{1} = -\sigma_t(\xi_t)^{-1}\bar{\sigma}_t^e\left[e_t u_{xx}(t)+\nu\theta_{xx}(t)e_tE\left(\int_t^T e^{-\beta(s-t)}e^{\int_t^s(\nu\theta_y(r)-\alpha)dr}u_y(s)ds|\mathcal{F}_t\right)\right.$$
$$\left.+(\nu\theta_x(t)+\eta)e^{-\int_0^t(\nu\theta_y(r)-\alpha-\beta)dr}E(A_T-A_t|\mathcal{F}_t)\right].$$

$A_T - A_t$ is given by

$$\int_t^T e^{-\beta r+J_r}[\nu u_y(r)\int_t^r(\theta_{xy}(u)e_u+\theta_{yy}(u)\int_t^u e^{J_u-J_v}(\nu\theta_x(v)+\eta)e_v dv)du+u_{xy}(r)e_r]dr$$
$$+\int_t^T e^{-\beta r+J_r}[u_{yy}(r)\int_t^r e^{J_r-J_u}(\nu\theta_x(u)+\eta)e_u du]dr$$

and $j_u = \nu\theta_y(u)-\alpha,\ J_t=\int_0^t j_u du.$

Setting $\nu = 0$ we obtain the formula obtained in [Detemple and Zapatero, 1991].

Some of the components of the equilibrium interest rate are similar to those obtained with an AEU or with the classical habit, we refer to [Detemple and Zapatero, 1991] for the general equilibrium interpretation. Ceteris paribus the equilibrium interest rate is positively related to the expected growth in consumption and negatively related (when $u_{xxx} > 0$) to the variance in consumption. There is a component which associates marginal utility from consumption to the habit, its effect depends on the sign of u_{xy} and on the time derivative of the habit. If the utility function exhibits strong complementarity ($u_{xy} < 0$) and the habit is going up then we have an increase in the interest rate, on this point see [Detemple and Zapatero, 1991]. The last two terms come from the disutility associated with the marginal disutility of future standard of living. It is difficult to assess their sign and their size, under reasonable parameters values the first one should be negative.

Being the coefficients of the endowment process deterministic, we have that a single beta consumption CAPM holds as in [Detemple and Zapatero, 1991]:

$$\mu_t - r_t\mathbf{1} = -\beta^e cov(dG_t/S_t, de_t).$$

Assuming a more general Itô process then we will obtain a two Beta consumption CAPM as pointed out in [Detemple and Zapatero, 1991]. The second Beta will be related to changes in marginal disutility of future standards of living induced by stochastic shifts in the coefficients of the model.

Without further specifying the utility functional, no results can be stated about the magnitude of the risk premium. However, as it is not possible to evalaute the partial derivatives of θ, it seems to be difficult to establish a result about the resolution of the equity premium puzzle with a BFSDU as it is done with the classical habit formation process, see [Detemple and Zapatero, 1991].

7 Conclusions

In this paper we have proposed a utility process obtained as the solution of a backward-forward stochastic differential equation. The backward-forward feature is due to a habit formation process which is given by the weighted average of past consumption and of past conditional expected utility. The peculiarity of our utility function is that the habit of the agent is influenced not only by past cosnumption but also by a smoothed average of the conditional expected utility that the agent experienced in the past.

In our analysis we have addressed some of the key points related to this utility function. Existence of the utility process, existence of the optimal consumption plan, characterization of the Arrow-Debreu price process, of the equilibrium interest rate and of the market prices of risk. The results obtained with the classical habit formation process are confirmed only in part. As a general result we have that ceteris paribus the price process is smaller than the one obtained with an additive expected utility, the risk premium is higher than the one obtained with an additive expected utility only when the system is linear (istantaneous utility linear in the habit), considering the general nonlinear system it is difficult to state conclusive results about the risk premium and the consumption process.

A Appendix

A.1 The Linear BFSDU

Given the matrix $A = \begin{pmatrix} -\beta & -\gamma \\ -\nu & \alpha \end{pmatrix}$, there are two real eigenvalues:

$$\lambda_1 = \frac{\alpha - \beta - \sqrt{(\alpha+\beta)^2 + 4\gamma\nu}}{2}$$

$$\lambda_2 = \frac{\alpha - \beta + \sqrt{(\alpha+\beta)^2 + 4\gamma\nu}}{2}.$$

Then we have the following:

$$e^{As} = \begin{pmatrix} \frac{(\beta+\lambda_2)e^{\lambda_1 s} - (\beta+\lambda_1)e^{\lambda_2 s}}{\lambda_2 - \lambda_1} & \frac{\gamma(e^{\lambda_1 s} - e^{\lambda_2 s})}{\lambda_2 - \lambda_1} \\ \frac{\nu(e^{\lambda_1 s} - e^{\lambda_2 s})}{\lambda_2 - \lambda_1} & \frac{(\beta+\lambda_2)e^{\lambda_2 s} - (\beta+\lambda_1)e^{\lambda_1 s}}{\lambda_2 - \lambda_1} \end{pmatrix}.$$

Given the assumptions done in our model we have two real eigenvalues with $\lambda_1 < 0 < \lambda_2$. About the sign of the elements of e^{As} it is easy to show that:

$$e_{21}^{As} < 0, \quad e_{12}^{As} < 0, \quad e_{11}^{As} > 0, \quad e_{22}^{As} > 0, \quad \forall s \in [0,T].$$

We observe that

$$H_s = \frac{e_{11}^{As} e_{22}^{AT} - e_{12}^{AT} e_{21}^{As}}{e_{22}^{AT}} =$$

$$\frac{[(\beta+\lambda_2)e^{\lambda_1 s} - (\beta+\lambda_1)e^{\lambda_2 s}][(\beta+\lambda_2)e^{\lambda_2 T} - (\beta+\lambda_1)e^{\lambda_1 T}] - \gamma\nu(e^{\lambda_1 T} - e^{\lambda_2 T})(e^{\lambda_1 s} - e^{\lambda_2 s})}{(\lambda_2 - \lambda_1)((\beta+\lambda_2)e^{\lambda_2 T} - (\beta+\lambda_1)e^{\lambda_1 T})} =$$

$$\frac{((\beta+\lambda_2)^2 + \gamma\nu)e^{\lambda_1 s + \lambda_2 T} + ((\beta+\lambda_1)^2 + \gamma\nu)e^{\lambda_1 T + \lambda_2 s}}{(\lambda_2 - \lambda_1)((\beta+\lambda_2)e^{\lambda_2 T} - (\beta+\lambda_1)e^{\lambda_1 T})} > 0$$

and

$$K_s = \frac{e_{12}^{AT} e_{22}^{As} - e_{12}^{As} e_{22}^{AT}}{e_{22}^{AT}} =$$

$$\gamma \frac{(e^{\lambda_1 T} - e^{\lambda_2 T})((\beta+\lambda_2)e^{\lambda_2 s} - (\beta+\lambda_1)e^{\lambda_1 s}) - (e^{\lambda_1 s} - e^{\lambda_2 s})((\beta+\lambda_2)e^{\lambda_2 T} - (\beta+\lambda_1)e^{\lambda_1 T})}{(\lambda_2 - \lambda_1)((\beta+\lambda_2)e^{\lambda_2 T} - (\beta+\lambda_1)e^{\lambda_1 T})} =$$

$$\gamma \frac{e^{\lambda_1 T}e^{\lambda_2 s} - e^{\lambda_2 T}e^{\lambda_1 s}}{(\beta+\lambda_2)e^{\lambda_2 T} - (\beta+\lambda_1)e^{\lambda_1 T}} < 0.$$

About the time derivatives of K_s and H_s, k_s and h_s respectively we have the following:

$$h_s = \frac{\lambda_1((\beta+\lambda_2)^2 + \gamma\nu)e^{\lambda_1 s + \lambda_2 T} + \lambda_2((\beta+\lambda_1)^2 + \gamma\nu)e^{\lambda_1 T + \lambda_2 s}}{(\lambda_2 - \lambda_1)((\beta+\lambda_2)e^{\lambda_2 T} - (\beta+\lambda_1)e^{\lambda_1 T})} < 0$$

$$k_s = \gamma \frac{\lambda_2 e^{\lambda_1 T}e^{\lambda_2 s} - \lambda_1 e^{\lambda_2 T}e^{\lambda_1 s}}{(\beta+\lambda_2)e^{\lambda_2 T} - (\beta+\lambda_1)e^{\lambda_1 T}} > 0.$$

A.2 Proof of Proposition 6.4

With the notation employed in Section 6, the Arrow-Debreu price process ξ_t is characterized as follows

$$\xi_t = u_x(t) + (\nu\theta_x(t) + \eta)E(\int_t^T e^{-\beta(s-t)+\int_t^s(\nu\theta_y(u)-\alpha)du} u_y(s)\, ds|\mathcal{F}_t).$$

To prove our statement we have to differentiate both sides.

Assuming market equilibrium $c_t = e_t$, the consumption process and the habit formation satisfy

$$\begin{aligned} de_t &= e_t(\bar{\mu}^e_t dt + \bar{\sigma}^e_t dW_t) \qquad e_0 = 1 &(38)\\ dY_t &= (\nu\theta(t) + \eta e_t - \alpha Y_t)dt, \qquad y_0 > 0 &(39)\end{aligned}$$

Since $\bar{\mu}^e, \bar{\sigma}^e$ are bounded and deterministic and θ with its first partial derivatives are bounded and continuous, $e_t, Y_t \in \mathbb{L}^{1,2}$ and we can compute explicitly their Malliavin derivatives. Denoting by $\mathcal{D}$ the Malliavin derivative for $r \leq t$ we have

$$\begin{aligned} \mathcal{D}_r e_t &= \mathcal{D}_r(\exp\{\int_0^t (\bar{\mu}^e_s - \frac{1}{2}\bar{\sigma}^e_s)ds + \int_0^t \bar{\sigma}^e_s dW_s\}) = e_t\bar{\sigma}^e_r,\\ \mathcal{D}_r Y_t &= \int_r^t \mathcal{D}_r(\nu\theta(s) + \eta e_s - \alpha Y_s)ds\\ &= \int_r^t [(\nu\theta_x(s) + \eta)\mathcal{D}_r e_s + (\nu\theta_y(s) - \alpha)\mathcal{D}_r Y_s]ds\\ &= \int_r^t [(\nu\theta_x(s) + \eta)\sigma^e_r e_s + (\nu\theta_y(s) - \alpha)\mathcal{D}_r Y_s]ds, \end{aligned}$$

the latter can be solved explicitly obtaining

$$\mathcal{D}_r Y_t = \bar{\sigma}^e_r \int_r^t e^{\int_s^t(\nu\theta_y(u)-\alpha)du}(\nu\theta_x(s) + \eta)e_s ds.$$

Let us rename

$$\begin{aligned} j_u &= \nu\theta_y(u) - \alpha, \qquad J_t = \int_0^t j_u du\\ M_t &= E(\int_0^T e^{-\beta s + J_s} u_y(e_s, Y_s)\, ds|\mathcal{F}_t), \qquad G_t = \int_0^t e^{-\beta s + J_s} u_y(s)\, ds \end{aligned}$$

then $E(\int_t^T e^{-\beta(s-t)} \exp\{\int_t^s (\nu\theta_y(u) - \alpha)du\}\, u_y(s)\, ds|\mathcal{F}_t) = e^{-J_t+\beta t}(M_t - G_t)$ and we can write the Arrow Debreu price as

$$\xi_t = u_x(t) + (\nu\theta_x(t) + \eta)e^{-J_t+\beta t}(M_t - G_t). \qquad (40)$$

Differentiating the left side we have

$$d\xi_s = -\xi_s((r_s - \beta)ds + \lambda_s dW_s),$$

while the right side gives

$$\begin{aligned}
&du_x(t) + (\nu\theta_x(t) + \eta)d(\mathrm{e}^{-J_t+\beta t}(M_t - G_t)) + \nu\mathrm{e}^{-J_t+\beta t}(M_t - G_t)d\theta_x(t) + d[\nu\theta_x, \mathrm{e}^{-J+\beta}(M-G)]_t \\
&= du_x(t) - (\nu\theta_x(t) + \eta)(j_t - \beta)\mathrm{e}^{-J_t+\beta t}(M_t - G_t)dt \\
&\quad + (\nu\theta_x(t) + \eta)\mathrm{e}^{-J_t+\beta t}d(M_t - G_t) + \nu\mathrm{e}^{-J_t+\beta t}(M_t - G_t)d\theta_x(t) + \nu\mathrm{e}^{-J_t+\beta t}d[\theta_x, M]_t
\end{aligned}$$

being G continuous and of finite variation. By using Itô's Lemma and (38)-(39) we have

$$du_x(t) = \left[e_t\bar{\mu}_t^e u_{xx}(t) + (\nu\theta(t) + \eta e_t - \alpha Y_t)u_{xy}(t) + \frac{1}{2}e_t^2\bar{\sigma}_t^{e2}u_{xxx}(t)\right]dt + e_t\bar{\sigma}_t^e u_{xx}(t)dW_t.$$

Similarly we may compute

$$d\theta_x(t) = \left[\theta_{tx}(t) + e_t\bar{\mu}_t^e\theta_{xx}(t) + (\nu\theta(t) + \eta e_t - \alpha Y_t)\theta_{xy}(t) + \frac{1}{2}e_t^2\bar{\sigma}_t^{e2}\theta_{xxx}(t)\right]dt + e_t\bar{\sigma}_t^e\theta_{xx}(t)dW_t.$$

The last expression can be simplified, recalling that θ is the solution of the PDE (20) and that $\bar{\mu}^e$ and $\bar{\sigma}^e$ do not depend on x. Differentiating with respect to x we have

$$\begin{aligned}
&\theta_{tx}(t) + e_t\bar{\mu}_t^e\theta_{xx}(t) + (\nu\theta(t) + \eta e_t - \alpha Y_t)\theta_{xy}(t) + \frac{1}{2}e_t^2\bar{\sigma}_t^{e2}\theta_{xxx}(t) \\
= \quad & -(\beta + \bar{\mu}_t^e)\theta_x(t) - (\nu\theta_x(t) + \eta)\theta_y(t) + u_x(t) - e_t\bar{\sigma}_t^{e2}\theta_{xx}(t).
\end{aligned}$$

So we obtain

$$d\theta_x(t) = [-(\beta + \bar{\mu}_t^e)\theta_x(t) - (\nu\theta_x(t) + \eta)\theta_y(t) + u_x(t) - e_t\bar{\sigma}_t^{e2}\theta_{xx}(t)]dt + e_t\bar{\sigma}_t^e\theta_{xx}(t)dW_t.$$

Clearly $dG_t = \mathrm{e}^{-\beta t + J_t}u_y(t)dt$, so it remains to evaluate dM_t. By the Clark Ocone formula we have

$$M_t = E[F] + \int_0^t E[\mathcal{D}_s F|\mathcal{F}_s]dW_s,$$

where $F = \displaystyle\int_0^T \mathrm{e}^{-\beta r + J_r}u_y(r)dr$. Our hypotheses still guarantees that $F \in \mathbb{D}^{1,2}$, so from Malliavin calculus we have

$$\begin{aligned}
\mathcal{D}_s F &= \int_s^T \mathrm{e}^{-\beta r}\mathcal{D}_s(\mathrm{e}^{J_r}u_y(r))\, dr = \int_s^T \mathrm{e}^{-\beta r + J_r}[u_y(r)\mathcal{D}_s J_r + \mathcal{D}_s u_y(r)]dr \\
&= \int_s^T \mathrm{e}^{-\beta r + J_r}[u_y(r)\mathcal{D}_s(\int_0^r (\nu\theta_y(u) - \alpha)du)) + \mathcal{D}_s u_y(r)]dr \\
&= \int_s^T \mathrm{e}^{-\beta r + J_r}[u_y(r)\int_s^r \nu\mathcal{D}_s\theta_y(u)du) + \mathcal{D}_s u_y(r)]dr
\end{aligned}$$

where we used the definition of J_r. Applying the chain rule and the expressions of the Malliavin derivatives of e and Y, we obtain

$$
\begin{aligned}
\mathcal{D}_s F &= \int_s^T \mathrm{e}^{-\beta r+J_r}[u_y(r)\int_s^r \nu(\theta_{xy}(u)\mathcal{D}_s e_u + \theta_{yy}(u)\mathcal{D}_s Y_u)du + u_{xy}(r)\mathcal{D}_s e_r + u_{yy}(r)\mathcal{D}_s Y_r]dr \\
&= \bar{\sigma}_s^e \int_s^T \mathrm{e}^{-\beta r+J_r}[u_y(r)\int_s^r \nu\theta_{xy}(u)e_u du + u_{xy}(r)e_r]dr \\
&+ \int_s^T \mathrm{e}^{-\beta r+J_r}[u_y(r)\int_s^r \nu\theta_{yy}(u)\mathcal{D}_s Y_u du + u_{yy}(r)\mathcal{D}_s Y_r]dr
\end{aligned}
$$

which becomes

$$
\begin{aligned}
\mathcal{D}_s F &= \bar{\sigma}_s^e \Big\{ \int_s^T \mathrm{e}^{-\beta r+J_r}[u_y(r)\int_s^r \nu\theta_{xy}(u)e_u du + u_{xy}(r)e_r]dr \\
&+ \int_s^T \mathrm{e}^{-\beta r+J_r}\Big[u_y(r)\int_s^r \nu\theta_{yy}(u)\int_s^u \mathrm{e}^{J_u-J_v}(\nu\theta_x(v)+\eta)e_v dv du \\
&\qquad + u_{yy}(r)\int_s^r \mathrm{e}^{J_r-J_u}(\nu\theta_x(u)+\eta)e_u du\Big]\Big\} dr
\end{aligned}
$$

Summarizing $dM_t = \bar{\sigma}_t^e E(A_T - A_t|\mathcal{F}_t)dW_t$, where $A_T - A_t$ is given by

$$
\begin{aligned}
&\int_t^T \mathrm{e}^{-\beta r+J_r}[\nu u_y(r)\int_t^r (\theta_{xy}(u)e_u + \theta_{yy}(u)\int_t^u \mathrm{e}^{J_u-J_v}(\nu\theta_x(v)+\eta)e_v dv)du + u_{xy}(r)e_r]dr \\
+ &\int_t^T \mathrm{e}^{-\beta r+J_r}[u_{yy}(r)\int_t^r \mathrm{e}^{J_r-J_u}(\nu\theta_x(u)+\eta)e_u du]dr.
\end{aligned}
$$

Putting everything together we conclude that

$$
\begin{aligned}
&\xi_t(r_t-\beta) = -[\frac{1}{2}\bar{\sigma}_t^{e2}e_t^2 u_{xxx}(t) + e_t\bar{\mu}_t^e u_{xx}(t) + (\nu\theta(t)+\eta e_t - \alpha Y_t)u_{xy}(t) - (\nu\theta_x(t)+\eta)u_y(t)] \\
&-E(\int_t^T \mathrm{e}^{-\beta(s-t)+\int_t^s(\nu\theta_y(r)-\alpha)dr}u_y(s)ds|\mathcal{F}_t)[-\nu\theta_x(t)(\beta+\bar{\mu}_t^e) - (\nu\theta_x(t)+\eta)(2\nu\theta_y(t)-\alpha-\beta) \\
&+\nu(u_x(t) - e_t\bar{\sigma}_t^{e2}\theta_{xx}(t))] - \nu\mathrm{e}^{-\int_0^t(\nu\theta_y(r)-\alpha-\beta)dr}e_t\bar{\sigma}_t^{e2}\theta_{xx}(t)E(A_T - A_t|\mathcal{F}_t).
\end{aligned}
$$

With analogous considerations observing that $-\xi\lambda = \sigma_\xi$ we may conclude

$$
\begin{aligned}
\mu_t - r_t\mathbf{1} &= -\sigma_t(\xi_t)^{-1}\bar{\sigma}_t^e\Big[e_t u_{xx}(t) + \nu\theta_{xx}(t)e_t E\Big(\int_t^T \mathrm{e}^{\int_t^s(\nu\theta_y(r)-\alpha-\beta)dr}u_y(s)ds|\mathcal{F}_t\Big) \\
&\qquad +(\nu\theta_x(t)+\eta)\mathrm{e}^{-\int_0^t(\nu\theta_y(r)-\alpha-\beta)dr}E(A_T - A_t|\mathcal{F}_t)\Big].
\end{aligned}
$$

References

[Antonelli, 1993] Antonelli, F. (1993). Backward-forward stochastic differential equations. *Annals of Applied Probability*, 3:777–793.

[Breeden, 1979] Breeden, D. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7:265–296.

[Citti, Pascucci, Polidoro, 1998] Citti, G., Pascucci, A., Polidoro, S. (1998) Private communication.

[Constantinides, 1990] Constantinides, G. (1990). Habit formation: a resolution of the equity premium puzzle. *Journal of Political Economy*, 98:519–543.

[Cox and Huang, 1989] Cox, J. and Huang, C.-F. (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49:33–83.

[Cox and Huang, 1991] Cox, J. and Huang, C.-F. (1991). A variational problem arising in financial economics. *Journal of Mathematical Economics*, 20:465–487.

[Cox et al., 1985] Cox, J., Ingersoll, J., and Ross, S. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica*, 53:363–384.

[Cvitanic and Ma, 1996] Cvitanic, J. and Ma, J. (1996). Forward-Backward SDE's with reflection. *Prob. Th. and Rel. Fields*, .

[Detemple and Giannikos, 1996] Detemple, J. and Giannikos, C. (1996). Asset and commodoty prices with multi-attribute durable goods. *Journal of Economic Dynamics and Control*, 20:1451–1504.

[Detemple and Zapatero, 1991] Detemple, J. and Zapatero, F. (1991). Asset prices in an exchange economy with habit formation. *Econometrica*, 59:1633–1657.

[Detemple and Zapatero, 1992] Detemple, J. and Zapatero, F. (1992). Optimal consumption-portfolio policies with habit formation. *Mathematical Finance*, 2:251–274.

[Duffie, 1996] Duffie, D. (1996). *Dynamic Asset Pricing Theory.* Princeton University Press, Princeton.

[Duffie and Epstein and Skiadas 1992] Duffie, D. Epstein, L. and Skiadas, C. (1992). Stochastic differential utility. *Econometrica*, 60:353–394.

[Duffie, Geoffard, Skiadas, 1994] Duffie, D., Geoffard, J. and Skiadas, C. (1994). Efficient and equilibrium allocations with stochastic differential utility. *Journal of Mathematical Economics*, 23:133–146.

[Duffie and Lions, 1992] Duffie, D. and Lions, P. (1992). Pde solutions of stochastic differential utility. *Journal of Mathematical Economics*, 21:577–606.

[Duffie and Skiadas, 1994] Duffie, D. and Skiadas, C. (1994). Continuous-time security pricing. a utility gradient approach. *Journal of Mathematical Economics*, 23:107–131.

[Duffie and Zame, 1989] Duffie, D. and Zame, W. (1989). The consumption-based capital asset pricing model. *Econometrica*, pages 1279–1297.

[Epstein, 1987] Epstein, L. (1987). The global stability of efficient intertemporal allocations. *Econometrica*, 55:329–355.

[Epstein, 1988] Epstein, L. (1988). Risk aversion and asset prices. *Journal of Monetary Economics*, 22:179–192.

[Epstein and Wang, 1994] Epstein, L. and Wang, T. (1994). Intertemporal asset pricing under knightian uncertainty. *Econometrica*, 62:283–322.

[Epstein and Wang, 1995] Epstein, L. and Wang, T. (1995). Uncertainty, risk-neutral measures and security price booms and crashes. *Journal of Economic theory*, pages 40–82.

[Epstein and Zin, 1989] Epstein, L. and Zin, S. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. *Econometrica*, 57:937–969.

[Grossman and Laroque, 1990] Grossman, S. and Laroque, G. (1990). Asset pricing and optimal portfolio choice in the presence of illiquid durable consumption goods. *Econometrica*, 58:25–51.

[Hansen and Singleton, 1982] Hansen, L. and Singleton, K. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econbometrica*, 50:1269–1286.

[Hindy and Huang, 1993] Hindy, A. and Huang, C. (1993). Optimal consumption and portfolio rules with durability and local substitution. *Econometrica*, 61:85–121.

[Hu and Peng, 1995] Hu, Y. and Peng, S. (1995). Solution of forward -backward stochastic differential equations. *Probability Theory and Related Fields*, 103:273–283.

[Lucas, 1978] Lucas, R. (1978). Asset pricing in an exchange economy. *Econometrica*, 46:1429–1445.

[Ma et al., 1994] Ma, J., Protter, P., and Yong, J. (1994). Solving forward-backward stochastic differential equations explicitly -a four step scheme. *Probability Theory and Related Fields*, 98:339–359.

[Mehra and Prescott, 1985] Mehra, R. and Prescott, E. (1985). The equity premium: a puzzle. *Journal of Monetary Economics*, 15:145–161.

[West, 1988] West, K. (1988). Dividend innovations and stock price volatility. *Econometrica*, 56:37–61.