INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Envelope theorems in Banach lattices

Anna Battauz, Marzia De Donno and Fulvio Ortu

Working Paper n. 396

First Version: December 2010; This Version: May 2011

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Envelope theorems in Banach lattices

Anna Battauz
Department of Finance, and IGIER, Bocconi University

Marzia De Donno
Department of Decision Sciences, Bocconi University

Fulvio Ortu
Department of Finance and IGIER, Bocconi University

First Version: December 2010
This Version: May 2011

Abstract

We derive envelope theorems for optimization problems in which the value function takes values in a general Banach lattice, and not necessarily in the real line. We impose no restriction whatsoever on the choice set. Our result extend therefore the ones of Milgrom and Segal (2002). We apply our results to discuss the existence of a well-defined notion of marginal utility of wealth in optimal consumption-portfolio problems in which the utility from consumption is additive but possibly state-dependent and, most importantly, the information structure is *not required to be Markovian.* In this general setting, the value function is itself a random variable and, if integrable, takes values in a Banach lattice so that our general results can be applied.

1 Introduction

Envelope theorems constitute one of the genuine workhorses of economics, and their applications are truly ubiquitous. Over the years, in fact, several extensions of the traditional envelope theorems have emerged, as a response to the necessity of analyzing the behavior of the value function of optimization problems lacking the assumptions for the applicability of the standard envelope results of any graduate textbook. To our knowledge, the most general set of envelope theorems currently available in the literature is due to Milgrom and Segal (2002), who develop envelope results that do not require *any assumption on the choice set of the optimization problem.* In particular, they first show that the traditional envelope formula holds at any differentiability point of the value function, and then they establish conditions for the (left, right or full) differentiability of the value function.

The contribution of this paper is to extend Milgrom and Segal's results by allowing the objective function of the optimization problem under scrutiny to take values in a general Banach lattice, instead of simply in the set of real numbers. In doing so, we still allow the choice set to be arbitrary, while we let our parameters to belong to a general Banach space. In this setting, we show that all the results in Milgrom and Segal (2002) can be nicely extended, provided that the standard notion of differentiability for real-valued functions is suitably replaced by the more general notion of Fréchet-differentiability. Given a function from a Banach spaces into another Banach space, in fact, the Fréchet differential constitutes the right tool to analyze the incremental effect on the value of the function due to an increment in the independent variable.

The obvious question at this point is: why Banach lattices? This question can be answered from two angles. A first, more general angle involves the long standing tradition of using the lattice structure in economic theory, in particular in general equilibrium theory and in its applications to finance theory (see e.g. the references in Aliprantis, Monteiro and Tourky, 2004). In his fundamental contribution to general equilibrium theory, Mas-Colell (1986) analyzes the existence problem in economies where the commodity space is a general topological vector lattice (of which Banach Lattices are clearly special cases). The fundamental motivation for his approach lies in the mounting interest for economies with infinitely many commodities, such as Arrow-Debreu economics with infinitely-lived agents, or in the analysis of commodity differentiation, or in the arbitrage literature dating back to Black and Scholes (1973) and Harrison and Kreps (1979). The utility of employing the techniques associated with the lattice structure comes from the fact that, in many of these applications, the positive orthant would have empty interior, a fact that could produce insurmountable difficulties unless the lattice structure is taken to full bearing.

A second angle comes from a specific, interesting problem in asset pricing that we encountered when working on Battauz et al. (2011), and that actually motivated the present paper. Consider a standard, multi-period optimal consumption-portfolio problem for an agent with finite life. Let the agent's utility function be time-additive but possibly state-dependent. Crucially, depart from the standard treatments available in the literature by *not imposing the information structure to be Markovian.* More precisely, take for given the information structure modelled as a filtration of a standard probability space, and assume that prices and dividends are any stochastic processes adapted to the given, general filtration. To any time t associate a value function that, given the level of wealth accumulated up to t, gives the maximum remaining utility (continuation utility) conditional on following an optimal consumption-portfolio strategy from t on. The question is then: is the marginal utility of wealth well defined and, if so, how does it compare to the marginal utility from consumption at time t? In a Markovian framework, the usual envelope result could be invoked to obtain that, under standard conditions, the marginal utility of wealth is indeed well defined and it coincides with the marginal utility from consumption. This conclusion would easily derive from the fact that the, in a Markovian setting, the value function is a real-valued function of the wealth level (and, possibly, of other state variables). In a our general setting, however, this is not the case, since the value function itself would be a random variable, obtained by taking the 'right type' of sup over all possible controls, i.e. consumption paths and investment strategies. Upon requiring the value function to be at least integrable, the 'right type' of sup manifests itself upon recognizing that the set of integrable random variables is a nice Banach lattice, with the sup among random variables defined via the standard pointwise max operator. At this point, our general envelope results for Banach lattices can be taken to bear on this problem, to show that under a suitable set of conditions the value function is in fact (Fréchet)-differentiable, and its (Fréchet) differential equals the (Fréchet) differential of the utility function, when the latter exists.

The remainder of the paper is as follows. In the next section, we introduce the notation and definitions to set up the Banach lattice-valued optimization problem, we discuss the assumptions underlying our results and we then prove our extension of the envelope theorem to Banach lattices. In Section 3, we describe the asset pricing application of our result, that is we introduce a general optimal consumption-portfolio problem and we apply our general results from Section 2 to discuss the conditions under which the marginal utility of wealth is well defined and coincides with the marginal utility from consumption. Section 4 concludes. The appendix contains some basic material on Banach lattices useful for the problem at hand.

2 The general results

To prove their envelope theorem for arbitrary choice sets, Milgrom and Segal (2002) assume their parameter to lie in $[0, 1]$ but make clear that their approach applies to parameters lying in a more general normed vector space. Here we extend their results by allowing the objective function, and, as a consequence the value function, to take values in a Banach lattice and not just in $\mathbb{R}$, while maintaining the arbitrariness of the choice set.

In our extended framework, the usual notion of differentiability must be replaced with a notion of differentiability in Banach spaces. Several notions of differentiability can be given for functions between normed vector spaces. We recall them hereafter (see for instance Bonnans-Shapiro (2000) or the classical textbook by Luenberger (1969)). Let X, Y be normed vector spaces and G a mapping defined on a open domain $U \subset X$, with values in Y.

Definition 2.1 *We say that G admits **directional derivative** at a point $u \in U$ in a direction $x \in X$ if the limit:*

$$G'(u; x) := \lim_{h \to 0^+} \frac{G(u + hx) - G(u)}{h} \tag{2.1}$$

exists, where the limit is meant in Y-norm.

*The function G is said to be **Gateaux differentiable** at u if it is directional differentiable at u in every direction $x \in X$ and the directional derivative $G'(u; \cdot\,) : X \to Y$ is a continuous and linear operator. In this case, we denote this operator with $DG(u)$ (namely, $DG(u)(x) = G'(u; x)$) and call it the **Gateaux differential** of G at u.*

The Gateaux-differential extends the classical concept of directional derivative and it is a rather weak notion. Indeed, for instance, it does not imply continuity of the function in the differentiability points. A stronger notion which is often employed is the following:

Definition 2.2 *We say that G is **Fréchet-differentiable** at u if there exists a continuous and linear operator $DG(u) : X \to Y$ such that*

$$\lim_{\|x\|_X \to 0} \frac{\|G(u + x) - G(u) - DG(u)(x)\|_Y}{\|x\|_X} = 0 \tag{2.2}$$

*The operator $DG(u)(x)$ is called the **Fréchet differential** of G at u.*

If G is Fréchet differentiable at some point u, then it is also Gateaux differentiable at u and the two differentials coincide. For this reason, we use the same notation and will make it clear what we mean when it is not apparent from the context. Another important property of the Fréchet differentiability is that it implies continuity in the differentiability points. Precisely, if G is Fréchet differentiable at some point u, then it is continous at u.

Let now X be a Banach space and Y a order continuous Banach lattice. As it is usually done, we adjoin to Y the abstract maximal and minimal elements $\{\pm\infty\}$ and denote by $\bar{Y}$ the enlarged space.

We take an open set U in X as the set of parameters. Let Θ denote an arbitrary choice set and let $F : \Theta \times U \to Y$ be the objective function. For each parameter $u \in U$, we define the value function as:

$$V(u) = \sup_{\theta \in \Theta} F(\theta, u). \tag{2.3}$$

We set $V(u) = -\infty$ if $\Theta = \emptyset$. Moreover, $V(u)$ belongs to Y if and only if the set $(F(\theta, u))_{\theta in \Theta}$ is bounded from above.

Following Milgrom and Segal (2002), we start by showing that the envelope formula holds at any differentiability point of the value function. For any $u \in U$, we define the set of optimal choices

$$\Theta^*(u) = \{\theta \in \Theta : F(\theta, u) = V(u)\}.$$

We point out that in general $\Theta^*(u)$ can possibly be empty. In what follows, we fix $u^* \in U$ and make the following assumption:

Assumption 2.1 *The set $\Theta^*(u^*)$ is not empty.*

Theorem 2.1 *Assume that there exists some $r > 0$ such that $V(x) \in Y$ for every $x \in B(u, r)$*[1]*. Let $\theta \in \Theta^*(u^*)$. Then*

1. *if both $F(\theta, \cdot)$ and $V(\cdot)$ admits directional derivative at u^* in some direction $x \in X$, then $F'(\theta, u^*; x) \leq V'(u^*; x)$;*

2. *if both $F(\theta, \cdot)$ and $V(\cdot)$ are Gateaux-differentiable at u^*, then $DF(\theta, u^*) = DV(u^*)$;*

[1] $B(u, r)$ denotes as usual the ball centered in u and with radius r

3. *if both $F(\theta,\,\cdot\,)$ and $V(\,\cdot\,)$ are Fréchet-differentiable at u^*, then $DF(\theta,u^*)=DV(u^*)$*

Proof.

1. Let $h\in\mathbb{R}$ and $y\in X$ such that $\|hy\|_X<r$. Then:

$$F(\theta,u^*+hx)-F(\theta,u^*)\le V(u^*+hx)-V(u^*).$$

   In particular, taking h_n in $\mathbb{R}^+$, which decreases to 0 as $n\to+\infty$, and dividing both sides of the inequalities by h_n, we obtain

$$\frac{F(\theta,u^*+h_nx)-F(\theta,u^*)}{h_n}\le\frac{V(u^*+h_nx)-V(u^*)}{h_n}. \tag{2.4}$$

   If $F(\theta,\,\cdot\,)$ admits directional derivative at u^* along x, then according to (2.1), we have that

$$\frac{F(\theta,u^*+h_nx)-F(\theta,u^*)}{h_n}$$

   converges in Y-norm to $F'(\theta,u^*;x)$. Then there exists a a subsequence which converges in order to the same limit (see Remark A.1). Analogously, if V has a derivative at u^* along x, then

$$\frac{V(u^*+h_nx)-V(u^*)}{h_n}$$

   will converge in Y-norm, and, up to a subsequence, in order, to $V'(u^*;x)$. Hence

$$F'(\theta,u^*;x)\le V'(u^*;x). \tag{2.5}$$

2. If the two functions are Gateaux-differentiable then, they admit directional derivatives along all directions. In particular, they admit directional derivatives along x and $-x$ and we have $F'(\theta,u^*;x)\le V'(u^*;x)$ and $F'(\theta,u^*;-x)\le V'(u^*;-x)$. Since the Gateaux differential is homogeneous, $DF(\theta,u^*)(-x)=-DF(\theta,u^*)(x)$ and $DV(u^*)(-x)=-DV(u^*)(-x)$. Therefore,

$$DF(\theta,u^*)(x)=DV(u^*)(x)$$

   for all $x\in X$.

3. If $F(\theta,\,\cdot\,)$ and V are Fréchet differentiable at u^*, then they are a fortiori Gateaux differentiable and the differentials coincide. □

Our aim now is to determine a set of sufficient conditions for the value function to be Fréchet differentiable. As already observed by Milgrom and Siegel, the structure of the choice set is not relevant to this aim, though several versions of the envelope theorem exploit topological properties of the choice set and regularity of the objective function in the choice variable (see for instance subsection 2.1). The basic idea is to require that the objective function satisfies some properties *uniformly* with respect to the choice parameter.

We consider agains as fixed a value of the parameter $u^* \in U$ and make the following assumptions:

Assumption 2.2 *The objective function $F(\theta, \cdot)$ is Fréchet differentiable at u^* for every $\theta \in \Theta$. In particular, it is equidifferentiable, i.e.:*

$$F(\theta, u^* + x) - F(\theta, u^*) = DF(\theta, u^*)(x) + \sigma(\theta, u^*, x) \cdot \|x\|_X$$

where $|\sigma(\theta, u^*, x)| \leq \Sigma \|x\|_X$, for $\Sigma \in Y$, for all $\theta \in \Theta$, for $x \in X$ such that $u^* + x \in U$.

This assumption implies in particular that $F(\theta)$ is differentiable at u^* for any θ in the choice set. In addition, we have that

$$\frac{\|F(\theta, u^* + x) - F(\theta, u^*) - DF(\theta, u^*)(x)\|_Y}{\|x\|_X}$$

goes to 0 uniformly in θ as $\|x\|_X$ tends to 0.

Before introducing the next assumption, it is useful to recall some basic notation from functional analysis. We denote with $\mathcal{L}(X, Y)$ the vector space of all linear continuous (bounded) operators from X to Y. When well-defined, the Fréchet differential $DF(\theta, u^*)$ belongs to $\mathcal{L}(X, Y)$. We can define a norm on this space as follows: for $T \in \mathcal{L}(X, Y)$, we set

$$\|T\|_{\mathcal{L}} = \sup_{x \in X, \|x\|_X \leq 1} \|T(x)\|_Y.$$

The spirit of the next assumption is basically to require the Fréchet differential of F to be norm bounded uniformly in θ.

Assumption 2.3 *For every $x \in X$ there exists a vector $y_x \in Y$ such that*

$$|DF(\theta, u^*)(x)| \leq y_x \|x\|_X$$

for all $\theta \in \Theta$.

As a consequence, $\|DF(\theta,u^*)(x)\|_Y \leq M_x$ for all $\theta \in \Theta$, where $M_x = \|y_x\|_Y \|x\|_X$, that is, the family of operators $(DF(\theta,u^*))_{\theta\in\Theta}$ is pointwise bounded. The Banach-Steinhaus theorem, also known as the Uniform Boundedness Principle (see, for instance, Brezis (1983)), implies that the set $(DF(\theta,u^*))_{\theta\in\Theta}$ is also norm bounded, namely

$$\sup_{\theta\in\Theta} \|DF(\theta,u^*)\|_{\mathcal{L}} < +\infty.$$

In other words, there exists a constant Λ such that

$$\|DF(\theta,u^*)(x)\|_Y \leq \Lambda \|x\|_X \tag{2.6}$$

for all $\theta \in \Theta$.

A first step in the search for sufficient conditions for the value function to be differentiable at u^* is to show that V is at the minimum, continuous at the differentiability points of the objective function. This is ensured by the previous assumptions as the next proposition proves.

Proposition 2.1 *Assume that there exists $r > 0$ such that $V(x) \in Y$ for all $x \in B(u^*,r)$. If Assumptions 2.1, 2.2, 2.3 hold, the value function V is continuous in u^*.*

Proof. Let $x \in B(u^*,r)$. Then, we have:

$$\begin{aligned}
|V(x) - V(u^*)| &= \left|\sup_{\theta_1} F(\theta_1,x) - \sup_{\theta_2} F(\theta_2,u^*)\right| \\
&\leq \sup_{\theta} |F(\theta,x) - F(\theta,u^*)| \\
&\leq \sup_{\theta} |DF(\theta,u^*)(x-u^*)| \\
&\quad + \sup_{\theta} |\sigma(\theta,u^*,x-u^*)| \cdot \| x-u^*\|_X
\end{aligned}$$

where the last inequality is a consequence of Assumption 2.2. Taking the Y-norms of both sides and exploiting the triangle inequality together with Assumption 2.2 and 2.3 (in particular, inequality (2.6)), we obtain

$$\|V(x) - V(u^*)\|_Y < \Lambda \|x-u^*\|_X + \|\Sigma\|_Y \cdot \|x-u^*\|_X^2 .$$

This shows that V is continuous in u^*. $\square$

In order to obtain that V is differentiable, we will need to require the Fréchet differential of F at u^* to be continuous (in some sense better specified below). To this end, we impose the following assumption:

Assumption 2.4 *(i) There exists some $r > 0$, such that V is finite on $B(u^*, r)$ and for all $x \in B(u^*, r)$, the set $\Theta^*(x)$ is not empty;*

(ii) for $x \in B(u^*, r)$, for every $\theta_x \in \Theta^*(x), \theta \in \Theta^*(u^*)$, the Fréchet differential of F at u^* satisfies:

$$\lim_{\|x-u^*\|_X \to 0} \|DF(\theta_x, u^*) - DF(\theta, u^*)\|_{\mathcal{L}} = 0.$$

We do not require continuity of the Fréchet differential in the classical sense: in fact we do not ask F to be differentiable in other points but u^*. Instead, we ask $F(\theta, \cdot)$ to be differential at u^* for all θ (Assumption 2.2) and we require continuity of the family of differentials $(DF(\theta, u^*))_{\theta\in\Theta}$ as the choice θ approaches an optimal choice.

Exploiting this assumption, which naturally implies Assumption 2.1, together with Assumption 2.2 on the equidifferentiability of F, we are able to derive the following property, which is useful to show the differentiability of the value function.

Lemma 2.1 *Suppose that Assumptions 2.2 and 2.4 hold. Let let $\theta \in \Theta^*(u^*)$. Moreover, for some $x \in X$ with $\|x\| < r$ let $\theta_{u^*+x} \in \Theta^*(u^* + x)$. Then*

$$\lim_{\|x\|_X \to 0} \frac{\|F(\theta_{u^*+x}, u^* + x) - F(\theta_{u^*+x}, u^*) - DF(\theta, u^*)(x)\|_Y}{\|x\|_X} = 0$$

Proof. We have, thanks to Assumption 2.2, that:

$$F(\theta_{u^*+x}, u^* + x) - F(\theta_{u^*+x}, u^*) = DF(\theta_{u^*+x}, u^*)(x) + \sigma(\theta_{u^*+x}, u^*, x) \cdot \|x\|_X .$$

where $|\sigma(\theta_{u^*+x}, u^*, x)| \leq \Sigma \|x\|_X$. Hence, the following inequalities hold:

$$\begin{aligned} & \|F(\theta_x, u^* + x) - F(\theta_x, u^*) - DF(\theta, u^*)(x)\|_Y \\ \leq\ & \|DF(\theta_{u^*+x}, u^*)(x) - DF(\theta, u^*)(x)\|_Y + \|\sigma(\theta_{u^*+x}, u^*, x)\|_Y \cdot \|x\|_X \\ \leq\ & \|DF(\theta_{u^*+x}, u^*) - DF(\theta, u^*)\|_{\mathcal{L}} \cdot \|x\|_X + \|\Sigma\|_Y \|x\|_X^2 . \end{aligned}$$

Dividing by $\|x\|_X$ and taking the limit as $\|x\|_X \to 0$ we obtain the claim by using Assumption 2.4 (ii). □

We are finally ready to prove our main theorem.

Theorem 2.2 *If Assumptions 2.2, 2.4 hold, then the value function V is Fréchet-differentiable at u^* and*

$$DV(u^*) = DF(\theta, u^*)$$

for $\theta \in \Theta^*(u^*)$.

Proof. Let $\theta_{u^*} \in \Theta^*(u^*)$. Then,

$$V(u^*) = F(\theta_{u^*}, u^*) \geq F(\theta, u^*) \qquad \text{for any } \theta \in \Theta.$$

Now, let x be such that $\|x\|_X < r$ and take θ_{u^*+x} in $\Theta^*(u^* + x)$, which is not empty by Assumption 2.4 (i). Then

$$V(u^* + x) = F(\theta_{u^*+x}, u^* + x) \geq F(\theta, u^* + x) \qquad \text{for any } \theta \in \Theta.$$

In particular, $V(u^*) \geq F(\theta_{u^*+x}, u^*)$ and $V(u^* + x) \geq F(\theta_{u^*}, u^* + x)$. Thus, we can write:

$$F(\theta_{u^*}, u^* + x) - F(\theta_{u^*}, u^*) \leq V(u^* + x) - V(u^*) \leq F(\theta_{u^*+x}, u^* + x) - F(\theta_{u^*+x}, u^*).$$

Substracting the differential $DF(\theta_{u^*}, u^*)(x)$ and dividing by $\|x\|_X$, we obtain the following inequalities:

$$\begin{aligned} & \frac{F(\theta_{u^*}, u^* + x) - F(\theta_{u^*}, u^*) - DF(\theta_{u^*}, u^*)(x)}{\|x\|_X} \\ \leq \; & \frac{V(u^* + x) - V(u^*) - DF(\theta_{u^*}, u^*)(x)}{\|x\|_X} \\ \leq \; & \frac{F(\theta_{u^*+x}, u^* + x) - F(\theta_{u^*+x}, u^*) - DF(\theta_{u^*}, u^*)(x)}{\|x\|_X} \end{aligned}$$

Take now the limit as $\|x\|_X \to 0$. Since the first and the last term converge to 0 in Y-norm, the middle term must converge to 0 as well. This implies that V is Fréchet-differentiable at u^* and $DV(u^*) = DF(\theta, u^*)$. $\square$

2.1 The concavity assumption

In this section we want to investigate sufficient condition for the value function to be differentiable when a concavity assumption is made on the objective function in the spirit of Benveniste and Scheinkman (1979). For vector-valued concave functions, analogous results on differentiablity holds as in the real-valued case. For ease of the reader, we collect in Appendix B the main results on concavity and differentiability for functions on vector spaces.

Throughout this section, we make the following assumption:

Assumption 2.5 *The sets Θ and U are convex and the objective function F is concave with respect to both θ and u.*

The following result is immediate:

Lemma 2.2 *The value function V is concave.*

Proof. For any $\lambda \in [0,1]$, $u_1, u_2 \in X$, $\theta_1, \theta_2 \in \Theta$, we have the following inequalities:

$$\begin{aligned} V(\lambda u_1 + (1-\lambda)u_2) &\geq F(\theta_1 + (1-\lambda)\theta_2, \lambda u_1 + (1-\lambda)u_2) \\ &\geq \lambda F(\theta_1, u_1) + (1-\lambda)F(\theta_2, u_2). \end{aligned}$$

As a consequence,

$$V(\lambda u_1 + (1-\lambda)u_2) \geq \lambda \sup_{\theta_1 \in \Theta} F(\theta_1, u_1) + (1-\lambda) \sup_{\theta_2 \in \Theta} F(\theta_2, u_2) = \lambda V(u_1) + (1-\lambda)V(u_2). \qquad \square$$

We now fix as usual $u^* \in U$ and work under Assumption 2.1. The next two results show that under the concavity assumption the differentiability of the objective function implies the differentiability of the value function without further assumptions on the differential of F.

Theorem 2.3 *Let $F(\theta)$ be Gateaux differentiable and continuous at u^* for some $\theta \in \Theta^*(u^*)$. Then V is continuous and Gateaux differentiable at u^* and*

$$DV(u^*) = DF(\theta, u^*).$$

Proof. By Lemma 2.2, we know that V is concave. Moreover, $V(u) \geq F(\theta, u)$ for all $u \in U$. Therefore V is continuous at u^* thanks to Proposition B.1. It follows that $\partial V(u^*)$ is non-empty, where $\partial V(u^*)$ is the superdifferential set of V at u^* (Proposition B.3). Take $L \in \partial V(u)$: then,

$$Lx \geq V(u^* + x) - V(u^*) \geq F(\theta, u^* + x) - F(\theta, u^*),$$

namely $L \in \partial F(\theta, u^*)$. This means that $\partial V(u^*) \subset \partial F(\theta, u^*) = \{DF(\theta, u^*)\}$, where the last equality is a consequence of Proposition B.4. Since $\partial V(u^*)$ is non-empty, it must be necessarily

$$\partial V(u^*) = \{DF(\theta, u^*)\},$$

hence by Proposition B.4, we can say that V is Gateaux-differentiable at u^* and $DV(u^*) = DF(\theta, u^*)$. □

Corollary 2.1 *Let $F(\theta)$ be Fréchet differentiable at u^* for some $\theta \in \Theta^*(u^*)$. Then V is Fréchet differentiable at u^* and*

$$DV(u^*) = DF(\theta, u^*).$$

Proof. If $F(\theta)$ is Fréchet differentiable at u^*, then it is continuous and Gateaux differentiable at u^*. In virtue of the previous theorem, V is continuous and Gateaux-differentiable at u^* and

$$DV(u^*) = DF(\theta, u^*).$$

Moreover, since V is concave, the differential is a superdifferential, hence the following inequalities hold for all $x \in X$:

$$DV(u^*)(x) \geq V(u^* + x) - V(u^*) \geq F(\theta, u^* + x) - F(\theta, u^*)$$

or, equivalently,

$$0 \geq V(u^* + x) - V(u^*) - DV(u^*)(x) \geq F(\theta, u^* + x) - F(\theta, u^*) - DF(\theta, u^*)(x).$$

The inequalities are clearly reversed when taking the absolute values, that is:

$$0 \leq |V(u^* + x) - V(u^*) - DV(u^*)(x)| \leq |F(\theta, u^* + x) - F(\theta, u^*) - DF(\theta, u^*)(x)|$$

Passing to the norms and dividing by $\|x\|_X$ one obtains:

$$0 \leq \frac{\|V(u^* + x) - V(u^*) - DV(u^*)(x)\|_Y}{\|x\|_X} \leq \frac{\|F(\theta, u^* + x) - F(\theta, u^*) - DF(\theta, u^*)\|_Y}{\|x\|_X}.$$

One can then take the limit as $\|x\|_X$ goes to 0: the right-hand term goes to 0 because of the Fréchet differentiability of $F(\theta)$. As a consequence, the middle term goes to 0, which implies that V is Fréchet differentiable. □

3 Envelope results for general asset pricing models

We consider a frictionless security market in which J assets are traded over the investment horizon $\mathcal{T} = \{0, 1, \ldots, T\}$. Asset prices and cash-flows are denominated in units of the single good consumed in the economy. We assume that investors can freely dispose of the good. To describe the stochastic evolution of asset prices and cash-flows we take as given a filtered probability space $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t=0}^{T})$,[2] and denote by $d_j(t)$ the $\mathcal{F}_t$−measurable cash flow distributed by asset j at date t and by $S_j(t)$ the $\mathcal{F}_t$−measurable date t price of asset j *net of the current cash flow*. Given $p \in [1, +\infty[$, we assume that $S_j(t)$, $d_j(t) \in L^p(\mathcal{F}_t)$ for all t. Without loss of generality, we assume that the assets distribute no cash flow at date 0 and a liquidating one at date T, that is $d_j(0) = S_j(T) = 0$ almost surely.

A *dynamic investment strategy* is a sequence $\theta = \{\theta(t)\}_{t=0}^{T-1}$ of J-dimensional, $\mathcal{F}_t$−measurable random variables, that is $\theta(t) = \{\theta_1(t), \theta_2(t), \ldots, \theta_J(t)\}$, where $\theta_j(t)$ represents the position (in number of units) in assets j taken at date t and liquidated at date $t+1$. We denote by $V_\theta = \{V_\theta(t)\}_{t=0}^{T}$ the *value process* of the dynamic investment strategy θ, namely $V_\theta(t)$ is the date t value of a dynamic investment strategy, defined as the cost of establishing the positions in the J assets at their net-of-cash-flow prices, if t precedes the last trading date, and, at T, as the payoff from the final liquidation of θ. Formally:

$$V_\theta(t) = \begin{cases} \theta(t) \cdot S(t) & t < T \\ \theta(T-1) \cdot d(T) & t = T. \end{cases}$$

At any date t, a dynamic investment strategy θ produces a cash flow $x_\theta(t)$, generated by the difference between the resources obtained from liquidating the positions taken at $t-1$ at the cum-cash flow prices $S(t) + d(t)$, and the cost to establish the new positions at the net-of-cash flow prices $S(t)$. The cash-flow $x_\theta(t)$ is therefore related to the value $V_\theta(t)$ as follows:

$$x_\theta(t) = \begin{cases} -V_\theta(0) & t = 0 \\ \theta(t-1) \cdot [S(t) + d(t)] - V_\theta(t) & t = 1, \ldots, T-1 \\ V_\theta(T) & t = T. \end{cases} \tag{3.1}$$

Henceforth, we call the sequence $x_\theta = \{x_\theta(t)\}_{t=0}^{T}$ the *cash-flow process* of θ.

[2] As usual, we assume that $\mathcal{F}$ is augmented with P−null sets, $\mathcal{F}_0$ is the trivial sigma-algebra $\{\varnothing, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$.

Definition 3.1 *We call **admissible** any dynamic investment strategy θ such that $V_\theta(t)$, $x_\theta(t) \in L^p(\mathcal{F}_t)$ for $t = 0, 1, \ldots, T$. We denote with Θ the set of all admissible dynamic investment strategies.*

An agent in this market is identified by an *initial endowment* $e_0 \geq 0$ of the single consumption good and a *complete and transitive preference relation* on the set $\mathcal{C} = \prod_{t=0}^{T} L^p(\mathcal{F}_t)$ of consumption sequences $c = (c(0), c(1), \ldots, c(T))$, with $c(t) \in L^p(\mathcal{F}_t)$ for all t. In choosing the optimal intertemporal consumption and asset allocation, each agent $(e_0, \succeq)$ in $\mathcal{A}$ faces the budget constraint

$$B(e_0) = \{c \in \mathcal{C} \mid c(0) \leq x_\theta(0) + e_0,\ c(t) \leq x_\theta(t)\ \forall\ t > 0 \quad \textit{for some } \theta \in \Theta\}.$$

In particular, we consider the class of agents whose preferences have a time-additive von Neumann-Morgenstern representation, such that the period-utilities are allowed to depend on the state ω. In details, we assume that the preference $U(c)$ of an agent takes the following form

$$U(c) = \sum_{t=0}^{T} \int_\Omega u_t(c(t, \omega), \omega) dP(\omega) = \sum_{t=0}^{T} E\left[u_t\left(c(t)\right)\right] \tag{3.2}$$

where for all $t < T$,the period utilities $u_t : \Re \times \Omega \to \Re$ are assumed to satisfy the following conditions:[3]

(i) for all t, the function $u_t(c, \omega) : \Re \times \Omega \to \Re$ is measurable with respect to the product σ-algebra $\mathcal{B}(\Re) \otimes \mathcal{F}_t$ (where $\mathcal{B}(\Re)$ denotes the Borel σ-algebra)[4];

(ii) for all $c \in B(e_0)$, the integrals in (3.2) $\int_\Omega u_t(c(t, \omega), \omega) dP(\omega)$ are well defined[5] and either are finite or take the value $-\infty$; as a consequence, $U(c) < +\infty$ for all $c \in B(e_0)$.

(iii) for all t, the function $u_t(\,\cdot\,, \omega) : \Re \to \Re$ is real-valued and strictly increasing for almost every ω.

[3] For a discussion of period utilities that depend directly also on the state of nature ω see for instance Berrier, Rogers and Tehranchi (2007) or Frittelli, Maggis (2011).

[4] This condition guarantees that for every $\mathcal{F}_t$-measurable random vector $(c(t))$, the function $u_t\left(c(t, \omega), \omega\right)$ (defined on Ω with values in $\Re$) is $\mathcal{F}_t$-measurable.

[5] As usual for a random variable Z the integral $\int_\Omega Z(\omega) dP(\omega)$ is well defined and finite if both $\int_\Omega Z^+(\omega) dP(\omega) < +\infty$ and $\int_\Omega Z^-(\omega) dP(\omega) < +\infty$. We set $\int_\Omega Z(\omega) dP(\omega) = -\infty$ if $\int_\Omega Z^-(\omega) dP(\omega) = +\infty$ and $\int_\Omega Z^+(\omega) dP(\omega) < +\infty$. We set $\int_\Omega Z(\omega) dP(\omega) = +\infty$ if $\int_\Omega Z^+(\omega) dP(\omega) = +\infty$ and $\int_\Omega Z^-(\omega) dP(\omega) < +\infty$. Otherwise the integral is not defined.

An *optimal consumption-portfolio choice* for such an agent is a couple $(c^*, \theta^*) \in \mathcal{C} \times \Theta$ such that $c^*(0) \le x_{\theta^*}(0) + e_0$, $c^*(t) \le x_{\theta^*}(t)$ for $t = 1, \ldots, T$ and $U(c^*) \ge U(c)$ for all $c \in \mathcal{C}$ such that $c(0) \le x_\theta(0) + e_0$, $c(t) \le x_\theta(t)$ for $t = 1, \ldots, T$ for some $\theta \in \Theta$. We make the following assumption:

Assumption 3.1 : *There exists an optimal solution to the consumption-portfolio problem for an agent with preferences as in (3.2) and initial endowment e_0.*

It can be easily shown, as a consequence of the strict monotonicity of the period-utilities, that the constraints will be binding at the optimum, namely

$$\begin{aligned} &c^*(0) = x_{\theta^*}(0) + e_0 \\ &c^*(t) = x_{\theta^*}(t) \qquad\qquad \text{for } t = 1, \ldots, T. \end{aligned}$$

To any optimal consumption-portfolio choice (c^*, θ^*) for an agent with preferences as in (3.2) and initial endowment e_0, we associate the optimal intertemporal wealth $W^* = \{W^*(t)\}_{t=0}^T$ generated by θ^*, that is

$$W^*(t) = \begin{cases} e_0 & t = 0 \\ \theta^*(t-1) \cdot [S(t) + d(t)], & t = 1, \ldots, T. \end{cases}$$

Note that $W^*(t) = x_{\theta^*}(t) + V_{\theta^*}(t) = c^*(t) + V_{\theta^*}(t)$ for $t = 1, \ldots, T-1$.

Fix now $t \in \{0, 1, \ldots, T-1\}$, and let W be an $\mathcal{F}_t$-measurable random variable. We define a random variable $H(t, W)$ which represents the maximum remaining utility (or continuation utility) at time t for an agent whose current level of wealth is W:

$$\begin{aligned} H(t, W) \equiv \operatorname*{ess\,sup}_{(c,\theta) \in \mathcal{C} \times \Theta} &\sum_{s=t}^{T} E_t\left[u_s(c(s))\right] \\ s.t. \ &\begin{cases} c(t) + V_\theta(t) \le W \\ c(s) \le x_\theta(s) & s = t+1, \ldots, T \end{cases} \end{aligned} \tag{3.3}$$

for $t = 0, 1, \ldots, T$, where $E_t[\ \cdot\ ]$ denotes the conditional expectation with respect to $\mathcal{F}_t$. We assume that the integrals $E\left[u_s(c(s))\right]$ (and hence the conditional expectations in (3.3)) are well defined, and, for all consumptions satisfying the budget constraint at time t, are either finite or take the value $-\infty$ (in which case we set $E_t\left[u_s(c(s))\right] = -\infty$). In particular, for $W = W^*(t)$, we have the maximum remaining utility $H(t, W^*(t))$, given the optimal wealth level and the optimal past consumption.

We remark that the necessity to define the maximum remaining utility H via an essential sup, instead of a normal sup over real numbers, is due to the fact that we do not impose the information structure of the model to be Markovian.

In Battauz et al. (2011, Proposition 1), it is shown that at the optimum H is well-defined and finite, and that it satisfies the dynamic programming principle, that is:

$$H\left(t,W^{*}(t)\right)=u_{t}(c^{*}(t))+E_{t}\left[H\left(t+1,W^{*}(t+1)\right)\right];$$

or, equivalently

$$H\left(t,W^{*}(t)\right)=\sum_{s=t}^{T}E_{t}\left[u_{s}\left(c^{*}(s)\right)\right]. \tag{3.4}$$

In what follow, we consider the time t as fixed: therefore, for the sake of simplicity, we will let $H(W)=H(t,W)$ for some $W\in L^{p}(\mathcal{F}_{t})$. Following now the approach and the notation introduced in Section 2, we define the function:

$$F(\theta,W)=u_{t}(W-V_{\theta}(t))+E_{t}\left[\sum_{s=t+1}^{T}u_{s}\left(x_{\theta}(s)\right)\right].$$

Since all the period-utilities are not satiated, the constraints in (3.3) will be binding, so we can write the optimization problem (3.3) as

$$H(W)\equiv\operatorname*{ess\,sup}_{\theta\in\Theta_{t}}F(\theta,W)$$

where the choice set Θ_t is the set of admissible strategies at time t, namely the set of sequences $\theta=\{\theta\left(s\right)\}_{s=t}^{T-1}$ of J-dimensional, $\mathcal{F}_s-$measurable random variables such that $V_{\theta}(s),\ x_{\theta}(s)\in L^{p}(\mathcal{F}_{s})$ for $s=t,\ldots,T$. Note that Θ_t is a convex set. This property becomes relevant when a concavity assumption is made on the period-utilities. Assume that $F(\theta,W)$ takes value in $L^{1}(\mathcal{F}_{t})$ in a neighbourhood of the optimal wealth $W^{*}(t)$. We have thus reduced our initial problem to a problem of the form (2.3), where the parameter W lies in the Banach space $L^{p}(\mathcal{F}_{t})$ and the objective function takes values in the Banach lattice $L^{1}(\mathcal{F}_{t})$.

Remark 3.1 *Note that if we fix a strategy* $\theta \in \Theta_t$*, and two wealth levels* $W_1, W_2 \in L^p(\mathcal{F}_t)$ *and define* $c_i(t) = W_i - V_\theta(t)$ *for* $i = 1, 2$ *we have that*

$$\begin{aligned} F(\theta, W_1) - F(\theta, W_2) &= u_t(W_1 - V_\theta(t)) - u_t(W_2 - V_\theta(t)) \\ &= u_t(c_1(t)) - u_t(c_2(t)). \end{aligned}$$

In particular, if u_t *is Fréchet-differentiable at some point* $c(t) = W - V_\theta(t) \in L^p(\mathcal{F}_t)$*, the function* $F(\theta, \cdot\,)$ *is Fréchet-differentiable at* W *and*

$$DF(\theta, W) = Du_t(W - V_\theta(t)) = Du_t(c(t)). \tag{3.5}$$

As in the previous section, we denote with $\Theta^*(W)$ the set of optimal choices given the parameter W, namely the set of optimal admissible strategies (from time t up to time $T-1$), given the wealth level W at time t:

$$\Theta^*(W) = \{\theta \in \Theta_t : F(\theta, W) = H(W)\}.$$

Under Assumption 3.1, the set $\Theta^*(W^*(t))$ is not empty, as it is easily shown by exploting the Dynamic Programming Principle (3.4).

Proposition 3.1 *Let Assumption 3.1 hold and assume that the time t-period utility* $u_t(c)$ *is Fréchet-differentiable at the optimal consumption* $c^*(t)$ *and the time* t *value function* $H(W)$ *is Fréchet-differentiable at the optimal wealth* $W^*(t) = c^*(t) + V_{\theta^*}(t)$*. Then*

$$Du_t(c^*(t)) = DH(W^*(t)).$$

Proof. The optimal strategy θ^* belongs to the set $\Theta^*(W^*(t))$. From Remark 3.1, we deduce that $F(\theta^*, \cdot\,)$ is Fréchet-differentiable in $W^*(t)$. The claim then follows from Theorem 2.1 and equality (3.5). □

The Frechet differentials of H and u allow to define a notion of marginal utilities of wealth and consumption respectively. Following the argument in Battauz et alii. (2011), given the Frechet differential of H, one can define a linear and continuous functional $\mathcal{E}_{\mathcal{H}} : L^p(\mathcal{F}_t) \to \Re$ via

$$\mathcal{E}_{\mathcal{H}}(Y) = E\left[DH(W^*(t))(Y)\right]$$

for all $Y \in L^p(\mathcal{F}_t)$. By the Riesz representation theorem, there exists a unique $H_W \in L^q(\mathcal{F}_t)$ such that

$$\mathcal{E}_{\mathcal{H}}(Y) = E\left[DH(W^*(t))(Y)\right] = E\left[H_W Y\right] \qquad \text{for all } Y \in L^p(\mathcal{F}_t).$$

We call H_W the time t-**marginal utility of optimal wealth**. Analogously, we can uniquely find a random variable $u_c \in L^q(\mathcal{F}_t)$ such that

$$E\left[Du_t(c^*)(Y)\right] = E\left[u_c Y\right] \qquad \text{for all } Y \in L^p(\mathcal{F}_t)$$

which we call time t-**marginal utility of optimal consumption**. Then Proposition 3.1 implies that, if they exist, the time t- marginal utilities of optimal consumption and of optimal wealth coincide because of the uniqueness of the Riesz representation. If, in particular, u_t is not state-dependent, namley $u_t : \Re \to \Re$, and it is differentiable, then $Du_t(c^*)(Y) = u_t'(c^*)Y$. Therefore, in this case, $u_c = u_t'(c^*)$ which is the *standard* marginal utility of consumption.

In general, conditions are given on the utility functions u_s but not on the value function. It becomes important then to understand which assumptions on the period utilities guarantee the Fréchet differentiability of the value function. Benveniste and Scheinkman (1979) give sufficient condition for the case $t = 0$, when the choice set is convex and the objective function is concave. Milgrom and Segal showed that this result can be seen as a particular case of their envelope theorem. Following their lines, we will now exploit the results in Section 2 to obtain an envelope condition for general asset pricing models in which the utility function is allowed to be state-dependent and the information structure is not required to be markovian. To do so, we first translate the Assumptions introduced in the previous section in terms of the period utility u_t of our general asset pricing model.

First of all, we observe that Assumption 3.1 implies that there exists an optimal consumption-portfolio choice for an agent whose wealth at time t is $W^*(t)$. Hence Assumption 2.1 is satisfied.

We define the set of admissible maximal consumptions at time t which can be obtained with the wealth W^*:

$$\begin{aligned} \mathcal{C}^*(W^*) &= \{c \in L^p(\mathcal{F}_t) : c = W^* - V_\theta(t) \text{ for some } \theta \in \Theta_t\} \\ &= \{c \in L^p(\mathcal{F}_t) : c = c^*(t) + V_\theta(t) \text{ for some } \theta \in \Theta_t\}. \end{aligned}$$

Assumptions 2.2 and 2.3 becomes respectively a sort of equidifferentiability and uniform boundedness on the set of admissible consumptions. In particular, Assumption 3.2 implies that u_t is Fréchet-differentiable in $c^*(t)$.

Assumption 3.2 *For every $X \in L^p(\mathcal{F}_t)$ with a sufficiently small norm:*

$$u_t(c+X) - u_t(c) = Du_t(c)(X) + \sigma_t(c,X) \cdot \|X\|_{L^p}$$

with

$$\underset{c \in \mathcal{C}^*(W^*)}{\text{ess sup}} |\sigma_t(c,X)| \leq \Sigma \|X\|_{L^p} ,$$

for some integrable random variable Σ

Assumption 3.3 *For every $X \in L^p(\mathcal{F}_t)$ there exists an integrable random variable Λ_X such that*

$$\underset{c \in \mathcal{C}^*(W^*)}{\text{ess sup}} |Du_t(c)(X)| < \Lambda_X \|X\|_{L^p} \ .$$

This two assumptions, together with Assumption 3.1, imply the continuity of the value function, as an immediate application of Proposition 2.1:

Proposition 3.2 *Under Assumptions 3.1, 3.2, 3.3, the value function H is finite in a neighborhood of the optimal wealth and is continuous at the optimum $W^*(t)$.*

Finally Assumption 3.4 requires a continuity of the Fréchet differential of u as an admissible consumption approaches the optimal consumption c^*, provided that after perturbing the optimal wealth the agent is still able to find an optimal consumption-portfolio pair and that the optimal consumption is continuous as a function of the optimal wealth.

Assumption 3.4 *There exists a neighbourhood $\mathcal{I}^*$ of W^* such that:*

(i) for each $W \in \mathcal{I}^*$ the set $\Theta^*(W)$ is not empty, namely there exists an optimal consumption-portfolio choice for every level of wealth $W \in \mathcal{I}^*$;

(ii) if $W \in \mathcal{I}^*$ tends to W^* in $L^p(\mathcal{F}_t)$, then the corresponding optimal consumption at time t, $c_W(t)$ converges to $c^*(t)$ in $L^p(\mathcal{F}_t)$;

(iii) the Fréchet differential Du_t is continuous in $c^*(t)$

$$\lim_{\|X\|_{L^p}\to 0} \|Du_t(c^*(t)+X) - Du_t(c^*(t))\|_{\mathcal{L}} = 0.$$

Remark 3.2 *It is not difficult to verify that if Assumption 3.4 holds for u_t, then F satisfies Assumption 2.4. Indeed, let (c_W, θ_W) be the optimal consumption-portfolio pair for some $W \in \mathcal{I}^*$: this means in particular that $c_W(t) + V_{\theta_W}(t) = W$. Moreover*

$$\begin{aligned} DF(\theta_W, W^*) - DF(\theta^*, W^*) &= Du_t(W^* - V_{\theta_W}(t)) - Du_t(W^* - V_{\theta^*}(t)) \\ &= Du_t(c^*(t) + V_{\theta^*}(t) - V_{\theta_W}(t)) - Du_t(c^*(t)). \end{aligned}$$

If we denote $X = V_{\theta^*}(t) - V_{\theta_W}(t) = (W^* - W) - (c_W(t) - c^*(t))$, it is evident that, by Assumption 3.4 (ii), if W tends to W^* in L^p then X tends to 0 in L^p. Therefore Assumption 3.4 (iii) implies Assumption 2.4 *(ii).*

We are now ready to obtain the *envelope theorem for state-dependent utilities and for a general (i.e. not required to be Markov) information structure* as a consequence of our general Theorem 2.2:

Proposition 3.3 *If the time t period utility u_t satisfies Assumptions 3.2,3.4, then the value function $H(t, W(t))$ is Fréchet-differentiable at the optimal wealth $W^*(t)$ and*

$$Du_t(c^*(t)) = DH(W^*(t)) \tag{3.6}$$

In the literature, it is often assumed that the period-utilities are concave. As a consequence of Corollary 2.1, we see that in this case, Assumptions 3.2, 3.3, 3.4 are not necessary to obtain the differentiability of the value function. More precisely, the following result holds:

Proposition 3.4 *Assume that the period utilities u_s are concave for all $s = t, \dots, T$. If u_t is continuous and Gateaux-differentiable (resp. Fréchet differentiable) at the optimal consumption $c^*(t)$, then the value function $H(t, W(t))$ is continuous and Gateaux-differentiable (resp. Fréchet differentiable) at the optimal wealth $W^*(t)$ and equality* (3.6) *holds.*

3.1 The habit-formation case

In this section we extend the previous results to a class of period-utilities which may depend on past consumptions, so to include the habit formation case. In details, consider an agent with initial endowment e_0 whose preferences $U(c)$ over consumption sequences $c \in \mathcal{C}$ take the form

$$U(c) = \sum_{t=0}^{T} \int_{\Omega} u_t(\gamma(t,\omega),\omega) dP(\omega) = \sum_{t=0}^{T} E\left[u_t\left(\gamma(t)\right]\right.$$

where for all $t < T$, $\gamma(t) = (c(0), c(1), \ldots, c(t))$ is the collection of consumptions up to time t. Coherently to the previous section, we assume that the period utilities $u_t : \Re^{t+1} \times \Omega \to \Re$ satisfy the following conditions:

(i) for all t, the function $u_t(\gamma,\omega) : \Re^{t+1} \times \Omega \to \Re$ is measurable with respect to the product σ-algebra $\mathcal{B}(\Re^{t+1}) \otimes \mathcal{F}_t$;

(ii) for all $c \in B(e_0)$, the integrals in (3.2) $\int_{\Omega} u_t(\gamma(t,\omega),\omega) dP(\omega)$ are well defined and either are finite or take the value $-\infty$;

(iii) for every t, the function $u_t(\,\cdot\,,\omega) : \Re^{t+1} \to \Re$ is real-valued and strictly increasing[6] for almost every ω.

One can define an optimal-consumption portofolio choice as above, as well as the optimal intertemporal wealth. We assume that there exists an agent who solves her optimal consumption-portfolio problem and denote with $(\gamma^*(t))_{0 \le t \le T}$ the stream of optimal consumptions and, as usual, with θ^* and W^* respectively the optimal strategy and the optimal wealth. For this agent, the maximum remaining utility at time t, given the stream of optimal past consumptions $\gamma^*(t-1)$ and a $\mathcal{F}_t$-measurable level of wealth W, is defined as:

$$\begin{aligned} H(t,\gamma^*(t-1),W) \equiv \operatorname*{ess\,sup}_{(c,\theta)\in\mathcal{C}\times\Theta} & \sum_{s=t}^{T} E_t\left[u_s(\gamma^*(t-1), c(t), \ldots, c(s))\right] \\ s.t. & \begin{cases} c(t) + V_\theta(t) \le W & \\ c(s) \le x_\theta(s) & s = t+1, \ldots, T \end{cases} \end{aligned} \tag{3.7}$$

As it was proved in Battauz et alii. (2011), the dynamic programming principle (3.4) is still valid for this class of preferences. Moreover, the strict monotonicity of the period-

[6] The function $u : \Re^{t+1} \to \Re$ is strictly increasing if $u(c_0, \ldots, c_t) > u(\tilde{c}_0, \cdots, \tilde{c}_t)$ for every pair $(c_s)_{0 \le s \le t}, (\tilde{c}_s)_{0 \le s \le t}$ such that $c_s \ge \tilde{c}_s$ for all s and $c_{\bar{s}} > \tilde{c}_{\bar{s}}$ for at least one $\bar{s}$.

utilities forces the constraints to be binding, so that (3.7) can be written as:

$$H(t,\gamma^*(t),W) \equiv \underset{\theta\in\Theta_t}{\text{ess sup}}\; F(t,\gamma^*,\theta,W)$$

where

$$F(t,\gamma^*,\theta,W) = u_t(\gamma^*(t-1),W-V_\theta(t))+E_t\left[\sum_{s=t+1}^{T} u_s\left(\gamma^*(t-1),W-V_\theta(t),x_\theta(t+1),\dots,x_\theta(s)\right)\right]$$

and Θ_t is the set of admissible strategies at time t. In what follows, we omit the dependence of F on t, since the time is fixed.

Proposition 3.5 *Assume that the functions $\tilde{u}_s(\,\cdot\,) = u_s(\gamma^*(t-1),\,\cdot\,,x_\theta(t+1),x_\theta(s))$ are Fréchet differentiable at some point $W - V_\theta(t)$ for all $s = t,\dots,T$. Then the function $F(\gamma^*,\theta,\,\cdot\,)$ is Fréchet differentiable at W and, in this case,*

$$\begin{aligned} DF(\gamma^*,\theta,W)(X) &= \sum_{s=t}^{T} E_t\left[D\tilde{u}_s(W-V_\theta(t))(X)\right] \\ &= \sum_{s=t}^{T} E_t\left[Du_s(\gamma^*(t-1),W-V_\theta(t),x_\theta(t+1),\dots,x_\theta(s))\right] \end{aligned}$$

This proposition is an easy consequence of the following lemma:

Lemma 3.1 *Let $g : L^p \to L^1$ be Fréchet differentiable at some point $W \in L^p$. Then the function*

$$G(W) = E_t\left[g(W)\right]$$

is Fréchet differentiable at W and $DG(W)(X) = E_t\left[Dg(W)(X)\right]$.

Proof. Since g is is Fréchet differentiable at W, we have that

$$\lim_{\|X\|_{L^p}\to 0} \frac{E\left[|g(W+X)-g(W)-Dg(W)(X)|\right]}{\|X\|_{L^p}} = 0.$$

On the other hand, exploiting the properties of conditional expectation and, in particular, Jensen's inequality, we obtain that:

$$\begin{aligned} E\left[|G(W+X)-G(W)-E_t\left[Dg(W)(X)\right]|\right] &= E\left[|E_t\left[g(W+X)-g(W)-Dg(W)(X)\right]|\right] \\ &\le E\left[E_t\left[|g(W+X)-g(W)-Dg(W)(X)|\right]\right] \\ &= E\left[|g(W+X)-g(W)-Dg(W)(X)|\right]. \end{aligned}$$

It follows immediately that

$$\lim_{\|X\|_{L^p}\to 0} \frac{E\left[\left|G(W+X)-G(W)-E_t\left[\mathcal{D}g(W)(X)\right]\right|\right]}{\|X\|_{L^p}} = 0$$

which implies that G is Fréchet differentiable at W and $DG(W) = E_t\left[Dg(W)\right].$ □

In light of this result, it is immediate to extend Proposition 3.1 to the habit-formation case.

Proposition 3.6 *Assume that all the time s-period utilities* $\tilde{u}_t(\cdot) = u_s(\gamma^*(t-1),\, \cdot\,, c^*(t+1),\dots,c^*(s))$ *are Fréchet-differentiable at the optimal consumption* $c^*(t)$ *and the time* t *value function* $H(t,\gamma^*(t-1),\,\cdot\,)$ *is Fréchet-differentiable at the optimal wealth* $W^*(t) = c^*(t)+V_{\theta^*}(t)$*. Then*

$$\begin{aligned} DH(W^*(t))(X) &= \sum_{s=t}^{T} E_t\left[D\tilde{u}_s(c^*(t))(X)\right] \qquad (3.8)\\ &= \sum_{s=t}^{T} E_t\left[Du_s(\gamma^*(t-1),c^*(t),c^*(t+1),\dots,c^*(s))\right]. \end{aligned}$$

Proposition 3.3 can also be reformulated for this case, provided that all the period-utilities satisfy the regularity assumptions. For ease of the reader we write the equivalent of Assumptions 3.2,3.3,3.4 for the case where utilities depend on past consumption.

We first observe that the set of admissible maximal consumptions at time t, given the optimal wealth W^* in the current framework is:

$$\begin{aligned} \mathcal{C}^*(W^*) &= \left\{ c \in \prod_{s=t}^{T} L^p(\mathcal{F}_s) : c(t) = W^* - V_\theta(t), c(s) = x_\theta(s), \text{ for } s = t\dots,T, \text{ for some } \theta\in\Theta_t \right\}\\ &= \left\{ c \in \prod_{s=t}^{T} L^p(\mathcal{F}_s) : c(t) = c^*(t) + V_\theta(t), c(s) = c^*(s) + x_\theta(s) \text{ for some } \theta\in\Theta_t \right\}. \end{aligned}$$

Fix now $s\in\{t,\dots,T\}$. For sake of notation, we denote $c_t^s = (c(t),\dots,c(s))$.

Assumption 3.5 *Given* $c\in\mathcal{C}^*(W^*)$*, the period-utility* $u_s(\gamma^*(t-1),\,\cdot\,,c_{t+1}^s)$ *is Fréchet differentiable at* $c(t)$ *and for every* $X\in L^p(\mathcal{F}_t)$ *with a sufficiently small norm:*

$$\begin{aligned} u_s(\gamma^*(t-1),c(t)+X,c_{t+1}^s) - u_s(\gamma^*(t-1),c(t),c_{t+1}^s) &= Du_s(\gamma^*,\,\cdot\,,c_{t+1}^s)(c(t))(X)\\ &+ \sigma_s(c,X)\cdot\|X\|_{L^p} \end{aligned}$$

with

$$\underset{c \in \mathcal{C}^*(W^*)}{ess\ sup} |\sigma_s(c, X)| \leq \Sigma \|X\|_{L^p},$$

for some integrable random variable Σ*.*

Assumption 3.6 *For every* $X \in L^p(\mathcal{F}_t)$ *there exists an integrable random variable* Λ_X *such that*

$$\underset{c \in \mathcal{C}^*(W^*)}{ess\ sup} \left| Du_s(\gamma^*, \cdot, c^s_{t+1})(c(t))(X) \right| < \Lambda_X \|X\|_{L^p}.$$

Assumption 3.7 *There exists a neighbourhood* $\mathcal{I}^*$ *of* W^* *such that:*

(i) for each $W \in \mathcal{I}^*$ *the set* $\Theta^*(W)$ *is not empty, namely there exists an optimal consumption-portfolio choice for every level of wealth* $W \in \mathcal{I}^*$*;*

(ii) if $W \in \mathcal{I}^*$ *tends to* W^* *in* $L^p(\mathcal{F}_t)$*, then the corresponding stream of optimal consumptions after time t,* $(c_W(u))_{t \leq u \leq T}$ *converges to* $(c^*(u))_{t \leq u \leq T}$ *in* $\prod_{u=t}^T L^p(\mathcal{F}_u)$*;*

(iii) the Fréchet differential $Du_s(\gamma^*, \cdot, c^s_{t+1})(c(t))$ *is continuous in* $c^*(t)$ *in the sense that*

$$\lim_{\|W - W^*\|_{L^p} \to 0} \|Du_s(\gamma^*, \cdot, (c_W)^s_{t+1})(c_W(t)) - Du_s(\gamma^*, \cdot, (c^*)^s_{t+1})(c^*(t))\|_{\mathcal{L}} = 0.$$

We summarize in the next proposition all the results for the habit-formation case:

Proposition 3.7 *1. If Assumptions 3.1 holds and the period utilities* u_s *satisfy Assumptions 3.5 and 3.6 for all* $s = t \ldots, T,$ *, then the value function* $H(t, \gamma^*(t-1), \cdot)$ *is finite in a neighborhood of the optimal wealth and continuous at* $W^*(t)$*,*

2. If the period utilities u_s *satisfy Assumptions 3.5 and 3.7 for all* $s = t \ldots, T$*, then the value function* $H(t, \gamma^*(t-1), \cdot)$ *is Fréchet-differentiable at the optimal wealth* $W^*(t)$ *and equality* (3.8) *holds.*

3. If the period utilities u_s *are concave and Fréchet-differentiable at the optimal consumption* $c^*(t)$*, for all* $s = t \ldots, T$*, then the value function is concave and Fréchet-differentiable at the optimal wealth* $W^*(t)$ *and equality* (3.8) *holds.*

4 Conclusions

In this paper we have extended the most general class of envelope results, i.e. those due to Milgrom and Segal (2002), to the case in which the objective function takes values in a general Banach Lattice, and not necessarily the real line. Employing the concept of Fréchet-differentiability, our main results consists in identifying a set of assumptions under which the value function is Fréchet-differentiable, and its Fréchet differential coincides with the Fréchet differential of the objective function, seen as a function of the parameters.

We then apply our general result to the consumption-portfolio problem of an agent with time additive but possibly state-dependent utility, in a context in which the information structure is not required to be Markovian. In this setting, at any time t the value function (maximum remaining utility) is in fact a random variable itself, and not just a real-valued function defined on a set of state variables. To investigate if the value function for this problem has a well-defined marginal utility of wealth, defined as the Fréchet differential of the value seen as a function of wealth levels accumulated up to time t, we recognize that the value function takes values in L^1, the space of integrable random variables, and that L^1 is indeed a Banach lattice. This allows us to bring to full bearing our general results to identify a set of conditions under which the marginal utility of wealth is well defined and coincides with the marginal utility consumption, when the last one exists.

A Banach lattices

In this appendix, we recall the main definition and results on Banach lattices which are needed in our paper. We will mainly refer to Aliprantis and Border (1999).

Let $(X, C, \geq)$ be a partially ordered vector space, where C is a pointed convex cone which induces the order on X defined as:

$$x \geq y \iff x - y \in C.$$

Given $x, y, z \in X$, we say that z is the **supremum** of x, y, and denote $z = \sup\{x, y\} = x \vee y$, if:

1. $z \geq x$ and $z \geq y$;

2. if $u \geq x$ and $u \geq y$, then $u \geq z$.

The **infimum** of two elements is defined similarly and denoted $z = \inf\{x, y\} = x \wedge y$.

Definition A.1 *A partially ordered set X is called a **lattice** if each pair of elements has a supremum or an infimum.*
*A partially ordered vector space that is also a lattice is called a **Riesz space**.*

Definition A.2 *A subset A of a Riesz space X is **order bounded from above** if there is a vector $u \in X$ (called an upper bound of A) that dominates each element of A. A nonempty subset of a Riesz spaces has a **supremum** if there is an upper bound u such that $v \geq x$ for all $x \in A$ implies $v \geq u$. The supremum, if it exists, is unique.*

The definition of **order bounded from below** and **infimum** are analogous.
It is clear that A is order bounded from above if and only if inf $-A$ is order bounded from below. Moreover, we say that A is **order bounded** if it is order bounded both from above and below.

Definition A.3 *A Riesz space is **Archimedean** if $0 \leq nx \leq y$ for all $n \in \mathbb{N}$ and some $y \in X^+$ implies $x = 0$.*

*A Riesz space is called **order complete** if every nonempty susbset that is order bounded from above has a supremum.*

Every order complete Riesz space is Archimedean, but the converse is not true.

Definition A.4 *A net (x_α) **converges in order** to some $x \in X$ ($x_\alpha \xrightarrow{o} x$) if there is a net (y_α) such that $y_\alpha \downarrow 0$ and $|x_\alpha - x| \leq y_\alpha$ for each α.*

A net can have at most one order limit.

A Riesz space can be equipped with a norm. A lattice norm $\|\cdot\|$ has the property that $|x| \leq |y|$ implies $\|x\| \leq \|y\|$.

Definition A.5 *A complete normed Riesz space is called a **Banach lattice***

Examples

1. The Euclidean space $\mathbb{R}^n$ with the Euclidean norm.

2. The space $C(K)$ of all continuous functions on a compact space K with the sup norm (it is Archimedean but not order complete).

3. The spaces $L^p(\mu)$ $(1 \leq p \leq \infty)$ with the usual L^p-norm. Order convergence coincide with μ-almost-sure convergence.

Definition A.6 *A lattice norm $\|\cdot\|$ on a Riesz space is **order continuous** if $x_\alpha \downarrow 0$ implies $\|x_\alpha\| \downarrow 0$.*

All reflexive Banach lattices are order continuous. This property is important because a Banach lattice with order continuous norm is order complete. For instance, $L^p(\mu)$ is order continuous for $1 \leq p < \infty$ but not for $p = \infty$.

Remark A.1 *In a Banach space, the norm convergence is equivalent to **relative uniform star convergence**, namely a sequence x_n converges in norm to x if and only if for every subsequence x_{n_k} there exist a subsequence $x_{nk}(l)$ and an element $y \in X$ such that $|x_{nk}(l) - x| \leq y/l$ for $l = 1, 2, \ldots$.*

In an Archimedean vector lattice, relative uniform convergence implies order convergence. As a consequence, in an order complete Banach lattice, if a sequence x_n converges in norm to x, then there exist a subsequence which is order convergent to x.

B Concavity and differentiability

In this section we summarize the main definition and results for cone-concave functions on vector spaces and, in particular, on the relation between concavity and differentiability. The results which follow can be found in Valadier (1972), Borwein (1982), Papageorgiou (1983). Usually, definition and results are stated for convex function. Since we work under a concavity assumption, we reformulated them in the appropriate form for concave functions.

Let X be a topological vector space and $(Y, C, \geq)$ a order complete Banach lattice. We adjoin to Y the abstract maximal elements $\pm\infty$ and denote the new object by $\bar{Y}$.

Definition B.1 *A function $F : X \to Y$ is C**-concave** (or simply **concave**) if for all $x, y \in X$, $\lambda \in [0, 1]$*

$$F(\lambda x + (1 - \lambda)y) \geq \lambda F(x) + (1 - \lambda)F(y),$$

namely, $F(\lambda x + (1-\lambda)y) - \lambda F(x) + (1-\lambda)F(y) \in C$.

The sets of points at which F is finite is called the *essential domain* of F and denoted by $domF$. The algbraic interior of F is denoted $coreF$.

Proposition B.1 *(Proposition 2.3 in Borwein (1982)) Let* $G : X \to \bar{Y}$ *be concave. Assume that there exists a function* $F : X \to \bar{Y}$ *such that* $G(x) \geq F(x)$ *for all* $x \in X$*. If* F *is continuous at some point* $x_0 \in X$*, then* G *is continuous at* x_0*.*

Let now $\mathcal{L}(X,Y)$ denote the set of continous and linear operators between X and Y and let F be a concave function form X to $\bar{Y}$.

Definition B.2 *An operator* $L \in \mathcal{L}(X,Y)$ *is called a* ***superdifferential*** *for* F *at* x_0 *if for all* $x \in X$

$$L(x) \geq F(x_0 + x) - F(x_0).$$

The ***superdifferential set*** *is denoted by* $\partial F(x_0)$*.*

Proposition B.2 *(Proposition 3.2 (a) and Proposition 3.7 (a) in Borwein (1982)) If* $F : X \to \bar{Y}$ *is concave, with* $x_0 \in coreF$*, then*

$$F^{>}(x_0, x) = \sup_{h>0} \frac{F(x_0 + hx) - F(x_0)}{h}$$

exists and is everywhere finite and superlinear.

Proposition B.3 *(Proposition 4 and Théorème 6 in Valadier (1972)) If* $F : X \to \bar{Y}$ *is concave and* $x_0 \in coreF$ *then:*

(i) $L \in \mathcal{L}(X,Y)$ *is a superdifferential for* F *at* x_0 *if and only if* $L(x) \geq F^{>}(x_0, x)$ *for all* $x \in X$*;*

(ii) if in addition F *is continuous at* x_0*, then* $\partial F(x_0)$ *is non-empty, convex and equicontinuous in* $\mathcal{L}(X,Y)$ *and*

$$F^{>}(x_0, x) = \min\{L(x), L \in \partial F(x_0)\}.$$

Proposition B.4 *(Theorem 4.6 in Papageorgiou (1983)) Let* $F : X \to \bar{Y}$ *be a concave function. If* F *is continuous at* x_0*, then* F *is Gateaux-differentiable at* x_0 *if and only if* $\partial F(x_0)$ *is a singleton.*

References

[1] Aliprantis C.D., Border K.C. (1999), *Infinite Dimensional Analysis - A Hitchhiker's Guide*, Springer, Berlin.

[2] Aliprantis, C.D., Florenzano, M.., Tourky, R. (2004), General equilibrium analysis in ordered topological vector spaces, *J.Math. Econom. 40: 247-269.*

[3] Aliprantis, C.D., Monteiro, P.K., Tourky, R. (2004), Non-marketed options, non-existence of equilibria, and non-linear prices, *J.Econom.Theory 114: 345-357.*

[4] Battauz, A., De Donno, M., Ortu, F. (2011), Intertemporal Asset Pricing and the the Marginal Utility of Wealth, *Journal of Mathematical Economics,* forthcoming, doi:10.1016/j.jmateco.2011.02.005

[5] Benveniste, L.M., Scheinkman, J.A. (1979), On the differentiability of the value function in dynamic models of economics , *Econometrica 47, no.3: 727-732.*

[6] Berrier, F., Rogers, L.C.G., Tehranchi., M.R. (2009), A Characterization of Forward Utility Functions , mimeo, Statistical Laboratory, University of Cambridge.

[7] Black, F., Scholes, M. (1973) The pricing of options and corporate liabilities, *Journal of Political Economy 81(3): 637-654.*

[8] Bonnans, J.F., Shapiro, A. (2000), *Perturbation Analysis of Optimization Problems,* New York: Springer-Verlag.

[9] Borwein, J.M. (1982), Continuity and differentiability properties of convex operators, *Proc. London Math. Soc. 44: 420-444.*

[10] Brezis, H. (1983), *Analyse Fonctionelle - Théorie et applications*, Masson Editeur, Paris.

[11] Clausen, A., Strub, C. (2011), Envelope theorem for non-smooth and non-concave optimization. Preprint.

[12] Duffie, D. (2001), *Dynamic Asset Pricing Theory,* 3rd Edition, Princeton University Press, Princeton, New Jersey.

[13] Follmer, H., Schied, A. (2002) *Stochastic Finance – an Introduction in Discrete Time*, de Gruyter.

[14] Frittelli, M., Maggis, M. (2011) Conditional Certainty Equivalent, *International Journal of Theoretical and Applied Finance,* forthcoming

[15] Gale, D., Mas-Collel, A. (1975), An equilibrium existence theorem for a general model without ordered preferences , *Journal of Mathematical Economics 2:* 9-15*.*

[16] Grossman, S.J., Laroque, G. (1990), Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods , *Econometrica 58, no.1: 25-51.*

[17] Harrison, M. and D. Kreps (1979), "Martingales and arbitrage in multiperiod securities markets", *Journal of Economic Theory 20 : 381-408.*

[18] Luenberger, D.G. (1969), *Optimization by vector space methods.* New York: John Wiley and Sons.

[19] Mass-Colell, A. (1986), The Price Equilibrium Existence Problem in Topological Vector Lattices , *Econometrica 54, no.5: 1039-1053.*

[20] Mass-Colell, A., Whinston, M., Green, J. (1995) *Microeconomic Theory*, Oxford University Press

[21] Milgrom, P., Segal, I. (2002), Envelope Theorems for Arbitrary Choice Sets , *Econometrica 70, no.2: 583-601.*

[22] Papageorgiou, N. (1983), Nonsmooth analysis on partially ordered vector spaces: Part I - Convex case *Pacific Journal of Mathematics, 109: 463-495*

[23] Rockafellar, R.T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey.

[24] Valadier, M. (1972), Sous différentiabilité de fonctions convexes à valeurs dans un espace vectoriel ordonné, *Math. Scand., 30: 65-74.*