Envelope Theorem, Euler, and Bellman Equations without Differentiability*

Ramon Marimon[†] Jan Werner[‡]

February 15, 2015

Abstract

We extend the envelope theorem, the Euler equation, and the Bellman equation to dynamic constrained optimization problems where binding constraints can give rise to non-differentiable value functions. The envelope theorem – an extension of Milgrom and Segal (2002) theorem for concave functions – provides a generalization of the Euler equation and establishes a relation between the Euler and the Bellman equation. For example, we show how solutions to the standard Bellman equation may fail to satisfy the respective Euler equations, in contrast with solutions to the infinite-horizon problem. In standard maximisation problems the failure of Euler equations may result in *inconsistent multipliers*, but not in non-optimal outcomes. However, in problems with *forward looking* constraints this failure can result in inconsistent promises and non-optimal outcomes. We also show how the inconsistency problem can be resolved by a minimal extension of the co-state. As an application we extend the theory of recursive contracts of Marcet and Marimon (1998, 2015) to the case where solutions are not unique, resolving a problem pointed out by Messner and Pavoni (2004).

*Preliminary draft. Do not quote without authors' permission.

[†]European University Institute, UPF-Barcelona GSE, CEPR and NBER

[‡]University of Minnesota

1 Introduction

The Euler equation and the Bellman equation are the two basic tools to analyse dynamic problems. Euler equations are the first-order inter-temporal *necessary conditions* for optimal solutions and, under standard concavity-convexity assumptions, they are also *sufficient conditions*, provided that a transversality condition holds. Euler equations are second-order difference equations. The Bellman equation allows to transform an infinite-horizon optimisation problem into a ‘two-periods-value’ problem, resulting in time independent policy functions determining the actions as functions of the states. The Envelope Theorem provides the bridge between the Bellman equation and the Euler equations, confirming the necessity of the latter for the former, and allowing to use Euler equations to obtain the policy functions of the Bellman equation. Furthermore, in deriving the Euler equations from the Bellman equation, the policy function reduces the second-order difference equation system to first-order difference equations with corresponding initial conditions, *provided* that the value function is differentiable.

Differentiability provides a tight bridge between the Bellman equation and the Euler equation. When the value function is differentiable, the state provides univocal information about its derivative and, therefore, the inter-temporal change of values across states (Bellman) is uniquely associated with the change of marginal values (Euler), and it is this tight bridge that allows for the above described passage of properties between the Euler and the Bellman equations (e.g. necessity and sufficiency in one direction and the additional structure of policy functions in the other). Unfortunately, the bridge is no longer tight when the value function is not differentiable at a given state. In this case, knowing the state and its value does not provide univocal information about its derivative. Sub-differential calculus (e.g. Rockafellar (1970, 1981)) comes into play, but needs to be properly developed to characterise *the Envelope bridge between the Euler and Bellman equations, without differentiability*. This is the objective of this paper.

Recursive methods have been widely applied – particularly, in macroeconomics – in the last 25 years (e.g. since the publication of Stokey *et al.* (1989)) using the standard framework where assumptions, such as interiority of optimal paths, imply the differentiability of the value function. However, the differentiability issue cannot be ignored in a wide range of current applications. For example, even if the underlying functions describing the model (e.g. preferences and technology) are differentiable, value functions are not differentiable in many contractual and dynamic

problems when constraints are binding. Models where household, firms, or countries, may face binding constraints in equilibrium are, nowadays, more the norm than the exception: in particular, when *forward-looking* constraints (i.e. constraints involving future equilibrium outcomes) are part of the decision process. We focus our analysis on differentiability problems arising from binding constraints.

Our analysis covers the trilogy already mentioned. First, *the envelope theorem* for constrained optimisation problems without assuming differentiability of the value function or interiorness of the solutions. We generalise the envelope theorem of Milgrom and Segal (2002, Corollary 5) by dropping the concavity assumptions and weakening the Slater's condition which they assumed. Further, we extend it to multidimensional parameters. The theorem provides a characterisation of directional derivatives, as well as sufficient conditions for the differentiability of the value function; for example, If there is a unique saddle point, the value function is differentiable and the standard form of the envelope theorem holds. Furthermore, If the value function is concave or convex, the envelope theorem can be stated using the superdifferentials or the subdifferential. We provide a characterizations of the superdifferential of the concave value function and the subdifferential of the convex value function. A sufficient condition for differentiability of the concave value function is that the saddle-point multiplier be unique. This extends the well known result of Benveniste and Scheinkman (1979) which requires an interior solution implying that zero is the unique multiplier. For the convex value function, a sufficient condition for differentiability is that the solution is unique. We provide examples of applications of our results to static optimization problems. This first part, which also contributes to the static theory of constrained optimisation, is covered in Sections 2 to 5.

Second, we turn to *the Euler equation* – the first-order inter-temporal condition of the infinite-horizon problem. Unless the solution is interior, the marginal value of the constraint must be part of the Euler equation. The marginal value is well defined if and only if the value function is differentiable at the chosen, or 'variable', state. Euler equations hold for a sequence $\{x_t^*, \lambda_t^*\}_{t=0}^{\infty}$ that includes uniquely defined multipliers λ_t^* . If the value function is differentiable, Euler equations have recursive representation. In fact, since the multipliers are uniquely defined, the recursive representation is given by a policy function of the corresponding recursive problem. In the case of interior solutions the problem of the lack of recursivity of the Euler equation does not arise, but it does in the non-differentiable case. We apply the envelope theorem without differentiability to

generalize the characterization of optimal solutions using Euler equations when the value function may not be differentiable. In general, this characterization is not recursive, as it is in the variational problem, starting from the initial x_0 . In the latter, the inter-temporal Euler equations result in ‘consistent’ selections of multipliers. However, if following the optimal path $\{x_0, x_1^*, \dots, x_n^*\}$ a new optimisation is started, the state x_n^* does not provide information of the previously selected multipliers using the envelope theorem without differentiability. Nevertheless, when constraints do not involve future outcomes, this ‘inconsistency’ only affects the multipliers: the saddle-point solutions $\{x_t^*\}_{t=0}^\infty$ are optimal, even if the optimisation is re-started at some state. Section 6 contains this analysis and discussion.

Finally, we further develop *the Bellman equation*, or more precisely, *the saddle-point Bellman equation* of recursive contracts of Marcet and Marimon (1998, 2015) for dynamic optimisation problems with *forward looking* constraints. First, we show that, when the inter-temporal constraints are time-separable, using the ‘recursive contracts’ approach, for one-period *forward looking* constraints, one can have a well defined *saddle-point Bellman equation*. Second, we show why for long-term *forward looking* constraints the ‘inconsistency problem’, discussed in the previous section can have an effect on outcomes – not just multipliers – resulting in non-optimal solutions to the *saddle-point Bellman equation*. Messner and Pavoni (2004) provided an example showing that recursive contracts with non-unique outcomes could deliver non-optimal solutions (uniqueness is assumed in Marcet and Marimon (1998, 2011)), what we show here that the root of this problem lies in the ‘inconsistency problem’. Finally, and more importantly, we show that applying again the ‘recursive contracts’ method of expanding the co-state, it is possible obtain a *new saddle-point Bellman equation* where the recursive solution results in ‘consistent’ Euler equations; that is, a solution generated by this general *saddle-point Bellman equation* is a solution to the original infinite-horizon problem, even if this solution is re-started at a later (expanded) state. We conclude Section 7 with theorem the corresponding theorem that generalizes ‘recursive contracts’ to models with non-unique solutions.

Finally, it should be noticed that extensions of the envelope theorem have been developed in recent years for the use in optimization problems with incentive constraints and/or non-convexities. Rincón-Zapatero and Santos (2009) study differentiability of the value function in dynamic optimization problems under the assumption that the value function is concave. Morand, Reffett and Tarafdar (2011) study generalized differentiability of the value function and the envelope theo-

rem in non-smooth optimization problems¹. Although the Euler equation is part of the standard toolkit of dynamic optimization problems (e.g. Stokey et al. 1989), we are not aware of any discussion of the recursively problem presented here. As said, this is possibly due to the fact that most of the analysis, and computations, with Euler equations implicitly assume that the value function is differentiable². We have already explained the structure of the paper; all proves have been relegated to the Appendix.

2 The Envelope Theorem

We consider the following parametric constrained optimization problem:

$$\max_{y \in Y} f(x, y) \tag{1}$$

subject to

$$h_1(x, y) \geq 0, \dots, h_k(x, y) \geq 0. \tag{2}$$

Parameter x lies in the set $X \subset \mathbb{R}^m$. Choice variable y lies in $Y \subset \mathbb{R}^n$. Objective function f is a real-valued function on $Y \times X$. Each constraint functions h_i is a real-valued function on $Y \times X$.³

We shall impose the following conditions

- A1.** Y is convex and compact; X is convex.
- A2.** f and h_i are continuous functions of (x, y) , for every i .
- A3.** For every x and every i , there exists $\hat{y}_i \in Y$ such that $h_i(x, \hat{y}_i) > 0$ and $h_j(x, \hat{y}_i) \geq 0$ for $j \neq i$.

Assumptions A1-2 are standard. Assumption A3 essentially says that none of the inequality constraints $h_i(x, y) \geq 0$ alone can be replaced by equality constraint $h_i(x, y) = 0$. It is weaker

¹Other contributions include Bonnisseau and Le Van (1996) and Clausen and Straub (2011).

²A recent, and limited, solution to the recursive contracts' uniqueness problem has been offered by Cole and Kubler (2012).

³Note that optimization problems with equality constraints can be represented in form (1–2) by taking $h_i = -h_j$ for some i and j .

than the Slater's condition which requires that there is $\bar{y} \in Y$ such that $h_i(\bar{x}, y) > 0$ for every i . If all functions h_i are concave in y , then A3 is equivalent to the Slater's condition.

Let $V(x)$ denote the value function of the problem (1–2). The Lagrangian function associated with (1–2) is

$$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda h(x, y), \quad (3)$$

where $\lambda \in \mathbb{R}_+^k$ is a vector of (positive) multipliers⁴. It is well-known that if (y^*, λ^*) is a saddle point of $\mathcal{L}$, that is, if

$$\mathcal{L}(x, y, \lambda^*) \leq \mathcal{L}(x, y^*, \lambda^*) \leq \mathcal{L}(x, y^*, \lambda), \quad (4)$$

for every $y \in Y$ and $\lambda \in \mathbb{R}_+^k$, then y^* is a solution to (1–2). Further, the slackness condition, $\lambda^* h(x, y^*) = 0$, holds and consequently

$$V(x) = \mathcal{L}(x, y^*, \lambda^*). \quad (5)$$

The set of saddle points of $\mathcal{L}$ at x is a product of two sets and is denoted by $Y^*(x) \times \Lambda^*(x)$ where $Y^*(x) \subset Y$ and $\Lambda^*(x) \subset \mathbb{R}_+^k$, see Rockafellar (1970), Lemma 36.2. If (y^*, λ^*) is a saddle point of $\mathcal{L}$, y^* will be called saddle-point solution and λ^* a saddle-point multiplier. The slackness condition implies that if the i th constraint is not binding, that is, $h_i(x, y^*) > 0$ for some $y^* \in Y^*(x)$, then $\lambda_i^* = 0$ for every $\lambda^* \in \Lambda^*(x)$. The set of saddle-point solutions $Y^*(x)$ is a subset of the set of solutions to (1–2). If functions f and h_i are concave in y and the Slater's condition holds, then the two sets are equal. If functions f and h_i are differentiable in y , then the Kuhn-Tucker first-order conditions hold for every saddle point of L . The set of saddle-point multipliers Λ^* is a subset of the set of Kuhn-Tucker multipliers. If functions f and h_i are differentiable and concave in y , those two sets are equal.

The envelope theorem is best stated in terms of directional derivatives. We first consider one-dimensional parameter set X – a convex set on the real line. Directional derivatives are then the left- and right-hand derivatives. The right- and left-hand derivatives of the value function V at x are

$$V'(x+) = \lim_{t \rightarrow 0+} \frac{V(x+t) - V(x)}{t}, \quad (6)$$

and

$$V'(x-) = \lim_{t \rightarrow 0-} \frac{V(x+t) - V(x)}{t}, \quad (7)$$

⁴We use the product notation: $\lambda h(x, y) = \sum_{i=1}^k \lambda_i h_i(x, y)$.

if the limits exist.

We have the following result:

Theorem 1: Suppose that $X \subset \mathfrak{R}$, conditions A1-A3 hold, and partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial h_i}{\partial x}$ are continuous functions of (x, y) . Then the value function V is right- and left-hand differentiable and the directional derivatives at $x \in \text{int}X$ are

$$V'(x+) = \max_{y^* \in Y^*(x)} \min_{\lambda^* \in \Lambda^*(x)} \left[\frac{\partial f}{\partial x}(x, y^*) + \lambda^* \frac{\partial h}{\partial x}(x, y^*) \right] \quad (8)$$

and

$$V'(x-) = \min_{y^* \in Y^*(x)} \max_{\lambda^* \in \Lambda^*(x)} \left[\frac{\partial f}{\partial x}(x, y^*) + \lambda^* \frac{\partial h}{\partial x}(x, y^*) \right], \quad (9)$$

where the order of maximum and minimum does not matter.

Theorem 1 is an extension of Corollary 5 in Milgrom and Segal (2002). It is worth pointing out that differentiability of functions f and h_i with respect to the variable y is not assumed in Theorem 1. This will be important in applications to recursive dynamic problems in Section 7.

For a multi-dimensional parameter set X in $\mathfrak{R}^m$, the directional derivative of the value function V at $x \in X$ in the direction $\hat{x} \in \mathfrak{R}^m$ such that $x + \hat{x} \in X$ is defined as

$$V'(x; \hat{x}) = \lim_{t \rightarrow 0+} \frac{V(x + t\hat{x}) - V(x)}{t}.$$

If partial derivatives of V exist, then the directional derivative $V'(x; \hat{x})$ is equal to the scalar product $DV(x)\hat{x}$, where $DV(x)$ is the vector of partial derivatives, i.e., the gradient vector.

Theorem 1 can be applied to the single-variable value function $\tilde{V}(t) \equiv V(x + t\hat{x})$ for which it holds $\tilde{V}'(0+) = V'(x; \hat{x})$. If $D_x f(x, y)$ and $D_x h_i(x, y)$ are continuous functions of (x, y) , then it follows that the directional derivative of V is

$$V'(x; \hat{x}) = \max_{y^* \in Y^*(x)} \min_{\lambda^* \in \Lambda^*(x)} \left[D_x f(x, y^*) + \lambda^* D_x h(x, y^*) \right] \hat{x}, \quad (10)$$

and the order of maximum and minimum does not matter.

3 Differentiability of the Value Function

The value function V on $X \subset \mathfrak{R}$ is differentiable at x if the one-sided derivatives are equal to each other. Sufficient conditions for differentiability can be obtained from Theorem 1.

Corollary 1: Under the assumptions of Theorem 1, each of the following conditions is sufficient for differentiability of value function V at $x \in \text{int}X$:

- (i) there is a unique saddle point,
- (ii) there is a unique saddle-point solution and h_i does not depend on x for every i .
- (iii) there is a saddle-point solution with non-binding constraints, and $\frac{\partial f}{\partial x}$ does not depend on y .⁵
- (iv) there is a unique saddle-point multiplier and $\frac{\partial f}{\partial x}$ and $\frac{\partial h_i}{\partial x}$ do not depend on y , for every i .

Condition (i) holds if there is a unique saddle-point solution with non-binding constraints. Then, zero is the unique saddle-point multiplier.

Since the set of saddle-point solutions $Y^*(x)$ is a subset of solutions to the constrained optimization problem (1–2) and the set of saddle-point multipliers $\Lambda^*(x)$ is a subset of Kuhn-Tucker multipliers, the uniqueness conditions in Corollary 1 are implied by the uniqueness of solution to (1–2) and the uniqueness of Kuhn-Tucker multiplier, respectively. A sufficient condition for uniqueness of the solution to (1–2) is that f be strictly quasi-concave and h_i be quasi-concave in y . A sufficient condition for uniqueness of Kuhn-Tucker multiplier is the following standard Constrained Qualification condition:

- CQ** (1) f and h_i are differentiable functions of y ,
- (2) vectors $D_y h_i(x, y^*)$ for $i \in I(x, y^*)$ are linearly independent, where $I(x, y^*) = \{i : h_i(x, y^*) = 0\}$ is the set of binding constraints.

A weaker form of Constrained Qualification which is necessary and sufficient for uniqueness of a Kuhn-Tucker multiplier can be found in Kyparisis (1985). Note that *CQ* holds vacuously for a solution y^* with non-binding constraints.

Under condition (i) or (iv) of Corollary 1, the derivative of the value function is

$$V'(x) = \frac{\partial f}{\partial x}(x, y^*) + \lambda^* \frac{\partial h}{\partial x}(x, y^*). \quad (11)$$

⁵It is sufficient that the constraints with h_i depending on x are non-binding. Other constraints may bind.

Under condition (ii) or (iii), it holds

$$V'(x) = \frac{\partial f}{\partial x}(x, y^*) \quad (12)$$

A result related to Corollary 1 (iii) can be found in Kim (1993). The condition of $\frac{\partial f}{\partial x}$ not depending on y is essentially the additive separability of f in x and y . In particular, it holds when f does not depend on x .

For multi-dimensional parameter set X in $\mathfrak{R}^m$, the value function is differentiable if $V'(x; \hat{x}) = -V'(x; -\hat{x})$ for every $\hat{x} \in \mathfrak{R}^m$. This holds under any of the sufficient conditions of Corollary 1 with $D_x f$ and $D_x h_i$ substituted for partial derivatives in (iv) and (v). If V is differentiable, then the gradient $DV(x)$ is well defined, and the multi-dimensional counterpart of (11) is

$$DV(x) = D_x f(x, y^*) + \lambda^* D_x h(x, y^*). \quad (13)$$

Results of this section and Section 2 can be extended to minimization problems and saddle-point problems. We present an extension to saddle-point problems in Appendix A.

4 Concave and Convex Value Functions

If the value function is concave or convex, the envelope theorem can be stated using the superdifferential or the subdifferential, respectively, for a multi-dimensional parameter set. We consider the concave case first.

Sufficient conditions for V to be concave are stated in the following well known result, the proof of which is omitted.

Proposition 1. If the objective function f and all constraint functions h_i are concave functions of (x, y) on $Y \times X$, then the value function V is concave.

The superdifferential $\partial V(x)$ of the concave value function V is the set of all vectors $\phi \in \mathfrak{R}^m$ such that

$$V(x') + \phi(x - x') \leq V(x) \quad \text{for every } x' \in X.$$

We have the following

Theorem 2: Suppose that conditions A1-A3 hold, derivatives $D_x f$ and $D_x h_i$ are continuous functions of (x, y) for every i , and V is concave. If x is an interior point of X , then

$$\partial V(x) = \bigcap_{y^* \in Y^*(x)} \bigcup_{\lambda^* \in \Lambda^*(x)} \{D_x f(x, y^*) + \lambda^* D_x h(x, y^*)\}. \quad (14)$$

Sufficient conditions for differentiability of concave value function follow from Theorem 2.

Corollary 2: Under the assumptions of Theorem 2, the following hold:

- (i) If the saddle-point multiplier is unique, then value function V is differentiable at x and (13) holds for every $y^* \in Y^*(x)$.
- (ii) If h_i does not depend on x for every i , then value function V is differentiable at every x and (13) holds for every $y^* \in Y^*(x)$.

In Corollary 2 (i), it is sufficient that the multiplier is unique for the constraints with h_i depending on x . Corollary 2 (i) implies that the value function is differentiable if there is a solution with non-binding constraints - those that depend on x - for then the unique saddle-point multiplier is zero. This is the well-known result due to Benveniste and Scheinkman (1979). Saddle-point multiplier may be unique even if some constraints are binding. Examples are given in Section 5. Corollary 2 (ii) has been established by Milgrom and Segal (2002, Corollary 3).

We now provide a similar characterization for convex value functions. Sufficient conditions for V to be convex are stated without proof in the following

Proposition 2. If the objective function $f(y, \cdot)$ is convex in x for every $y \in Y$ and all constraint functions h_i are independent of x , then the value function V is convex.

If V is convex, then the subdifferential $\partial V(x)$ is the set of all vectors $\phi \in \mathbb{R}^m$ such that⁶

$$V(x') + \phi(x - x') \geq V(x) \quad \text{for every } x' \in X.$$

We have the following

⁶We use the same notation for the superdifferential and the subdifferential as it is customary in the literature.

Theorem 3: Suppose that conditions A1-A3 hold, derivatives $D_x f$ and $D_x h_i$ are continuous functions of (x, y) for every i , and V is convex. If x is an interior point of X , then

$$\partial V(x) = \bigcap_{\lambda^* \in \Lambda^*(x)} \text{co} \left(\bigcup_{y^* \in Y^*(x)} \{D_x f(x, y^*) + \lambda^* D_x h(x, y^*)\} \right), \quad (15)$$

where $\text{co}(\cdot)$ denotes the convex hull.

Sufficient conditions for differentiability of convex value function follow from Theorem 3.

Corollary 3: Under the assumptions of Theorem 3, if the saddle-point solution is unique, then value function V is differentiable at x and (13) holds for every $y^* \in Y^*(x)$.

5 Examples

Example 1 (Perturbation of constraints): Suppose that the objective function f in (1) is independent of the parameter x and constraint functions are of the form $h_i(x, y) = \hat{h}_i(y) - x_i$. This optimization problem is a perturbation of the non-parametric problem with objective function f and constraint functions $\hat{h}_i$. Rockafellar (1970) provides an extensive discussion of the concave perturbed problem.

Corollary 1 (iv) implies that if the saddle-point multiplier λ^* is unique, then the value function is differentiable and $DV(x) = -\lambda^*$ by (13). (See 29.1.3 in Rockafellar (1970) for the concave perturbed problem.) If f and $\hat{h}^i$ are concave for every i , then V is concave and the superdifferential of V is $\partial V(x) = -\Lambda^*(x)$.

Example 2 (A planner's problem): Consider the resource allocation problem in an economy with k agents. The planner's problem is

$$\max_{\{c_i\}} \sum_{i=1}^k \mu_i u_i(c_i) \quad (16)$$

$$\text{s.t. } \sum_{i=1}^n c_i \leq x, \quad (17)$$

$$c_i \geq 0, \quad \forall i,$$

where $\mu = (\mu_1, \dots, \mu_k) \in \mathfrak{R}_{++}^k$ is a vector of welfare weights and $x \in \mathfrak{R}_+^L$ represents total resources. Utility functions u_i are continuous and increasing. Let $V(x, \mu)$ be the value of (16) as function of weights μ and total resources x . It follows from Corollary 1 (iv) that V is differentiable in x if the saddle-point multiplier of constraint (17) is unique. If utility functions u_i are differentiable, then the CQ condition holds implying that the multiplier is unique. The derivative is $D_x V = \lambda^*$, where λ^* is the multiplier of the constraint (17). V is a convex function of μ . The subdifferential $\partial_\mu V$ is (by Theorem 3) the convex hull of the set of vectors $(u_1(c_1^*), \dots, u_k(c_k^*))$ over all solutions c^* to (16). V is differentiable in μ if the solution is unique.

Consider an example with $L = 1, k = 2$, and $u_i(c) = c$. Let the welfare weights be parametrized by a single parameter μ so that $\mu_1 = \mu$ and $\mu_2 = 1 - \mu$ with $0 < \mu < 1$. The value function is $V(x, \mu) = \max\{\mu, 1 - \mu\}x$. It is differentiable with respect to μ at every $\mu \neq \frac{1}{2}$ and every x . The solution c^* is unique for every $\mu \neq \frac{1}{2}$. V is not differentiable with respect to μ at $\mu = \frac{1}{2}$. The left-hand directional derivative at $\mu = \frac{1}{2}$ is $-x$ while the right-hand derivative is x in accordance with Theorem 1. V is everywhere differentiable with respect to x .

6 Dynamic Optimization and Euler Equations

In this section we develop tools for dynamic optimization based on the results of Sections 2 - 4. We first show how these results apply to a possibly non-differentiable value function of a dynamic problem. We derive *Euler equations* without assuming differentiability of the value function and discuss in detail how to generalize standard dynamic programming results to the non-differentiable case. In particular, we show that non-differentiability of value function can break the standard results on Euler equations being necessary and sufficient for an optimal paths derived from Bellman equation.

We consider the following dynamic constrained maximization problem studied in Stokey et al. (1989):

$$\begin{aligned} \max_{\{x_t\}_{t=1}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) \\ \text{s.t.} \quad & h_i(x_t, x_{t+1}) \geq 0, \quad i = 1, \dots, k, \quad t \geq 0, \end{aligned} \tag{18}$$

where $x_t \in X \subset \mathbb{R}^n$ for every t , and x_0 is given. Functions F and h_i are real-valued functions on $X \times X$. In addition to assumptions A1 - A3, we impose the following conditions:

A4. F and h_i are bounded, and $\beta \in (0, 1)$.

A5. F and h_i are concave functions of (x, y) on $X \times X$,

A6. F and h_i are increasing and continuously differentiable.

The saddle-point problem associated with (18) is

$$\max_{\{x_t\}_{t=1}^{\infty}} \min_{\{\lambda_t\}_{t=1}^{\infty}, \lambda_t \geq 0} \sum_{t=0}^{\infty} \beta^t [F(x_t, x_{t+1}) + \lambda_{t+1} h(x_t, x_{t+1})], \quad (19)$$

As it is well known⁷, if sequence $\{x_t^*\}$ is a solution to (18), then under assumptions A1-A5 there exist multipliers $\{\lambda_t^*\}$ such that $\{x_t^*, \lambda_t^*\}$ is a saddle-point of (19). Furthermore, without any assumptions, if $\{x_t^*, \lambda_t^*\}$ is a saddle-point of (19) then $\{x_t^*\}$ is a solution to (18).

Under assumptions A1-A6, the first-order necessary conditions for saddle-point $\{x_t^*, \lambda_t^*\}$ are the following *intertemporal Euler equations*

$$D_y F(x_t^*, x_{t+1}^*) + \lambda_{t+1}^* D_y h(x_t^*, x_{t+1}^*) + \beta [D_x F(x_{t+1}^*, x_{t+2}^*) + \lambda_{t+2}^* D_x h(x_{t+1}^*, x_{t+2}^*)] = 0, \quad (20)$$

for every $t \geq 0$. Equations (20) together with complementary slackness conditions

$$\lambda_{i,t+1}^* h_i(x_t^*, x_{t+1}^*) = 0, \quad (21)$$

and the constraints $h_i(x_t^*, x_{t+1}^*) \geq 0$ for every t can be considered as a system of second-order difference equations for the sequence $\{x_t^*, \lambda_t^*\}$ with $x_0^* = x_0$. For an interior solution, with $h_i(x_t^*, x_{t+1}^*) > 0$ for every $t \geq 0$, equations (20) simplify to the standard Euler equations

$$D_y F(x_t^*, x_{t+1}^*) + \beta [D_x F(x_{t+1}^*, x_{t+2}^*)] = 0. \quad (22)$$

The sufficiency of the Euler equations and a transversality condition, see Stokey *et al.* (1989), continues to hold for non-interior solutions.

⁷See, for example, Luenberger (1969) Theorems 8.3 and 8.4.2.

Proposition 3: Suppose that conditions A1-A6 hold. Let $\{x_t^*, \lambda_t^*\}$, with $x_0^* = x_0$ and $\lambda_t^* \geq 0$ for every t , satisfy the Euler equations (20) and (21). If the transversality condition

$$\lim_{t \rightarrow \infty} \beta^t [D_x F(x_t^*, x_{t+1}^*) + \lambda_{t+1}^* D_x h(x_t^*, x_{t+1}^*)] x_t^* = 0, \quad (23)$$

holds, then $\{x_t^*\}$ is a solution to (18).

Proof: see Appendix.

Let $V(x_0)$ be the value function of (18). The value function satisfies the Bellman equation

$$V(x) = \max_y \{F(x, y) + \beta V(y)\} \quad (24)$$

$$\text{s.t. } h_i(x, y) \geq 0, \quad i = 1, \dots, k \quad (25)$$

For every solution $\{x_t^*\}$ to (18), x_{t+1}^* is a solution to Bellman equation (25) at x_t^* . The converse holds as well under assumption A4, see Stokey *et al.* (1989). As in Section 2, we shall consider saddle-points of the Lagrangian associated with the maximization problem on the right-hand side of Bellman equation. We shall refer to these as *saddle-points of Bellman equation*, for short.

Value function V is concave under assumption A5. The superdifferential of V is, by Theorem 2,

$$\partial V(x) = \bigcap_{y^* \in Y^*(x)} \bigcup_{\lambda^* \in \Lambda^*(x)} [D_x F(x, y^*) + \lambda^* D_x h(x, y^*)] \quad (26)$$

where $Y^*(x)$ is the set of saddle-point solutions and $\Lambda^*(x)$ is the set of saddle-point multipliers of the constraints (25) at x .

Corollary 2 (i) implies that V is differentiable if the multiplier λ^* is unique. Note that the constrained qualification CQ cannot be used to assert uniqueness of the multiplier since the objective function in (24) may not be differentiable in y . However, if there is a solution y^* with non-binding constraints, then the unique multiplier is zero and V is differentiable at x ; this is the well known result of Benveniste and Scheinkman (1979).

We discuss now the validity of the Euler equations for a sequence of saddle-point solutions to the Bellman equation. Consider a sequence $\{x_t^*, \lambda_t^*\}$ such that $(x_{t+1}^*, \lambda_{t+1}^*)$ is a saddle-point of Bellman equation at x_t^* . If assumptions A1-A4 hold, then the first-order condition for $(x_{t+1}^*, \lambda_{t+1}^*)$ is

$$D_y F(x_t^*, x_{t+1}^*) + \lambda_{t+1}^* D_y h(x_t^*, x_{t+1}^*) + \beta \phi_{t+1}^* = 0, \quad (27)$$

for some $\phi_{t+1}^* \in \partial V(x_{t+1}^*)$, see Rockafellar (1981, Ch.5). The envelope representation (26) of $\partial V(x_{t+1}^*)$ implies that there is saddle-point multiplier $\tilde{\lambda}_{t+2}^*$ at x_{t+1}^* such that

$$\phi_{t+1}^* = D_x F(x_{t+1}^*, x_{t+2}^*) + \tilde{\lambda}_{t+2}^* D_x h(x_{t+1}^*, x_{t+2}^*) \quad (28)$$

If value function V is differentiable, then multiplier $\tilde{\lambda}_{t+2}^*$ can be taken to be equal to λ_{t+2}^* for every t . Clearly then Euler equations (20) hold for $\{x_t^*, \lambda_t^*\}$. Otherwise, if V is non-differentiable, then multiplier $\tilde{\lambda}_{t+2}^*$ in the envelope representation (28) may be different from λ_{t+2}^* , and consequently Euler equations may not be satisfied for $\{x_t^*, \lambda_t^*\}$. Yet, Euler equations are guaranteed to hold for every sequence of solutions $\{x_t^*\}$ of Bellman equation (25) and some sequence of saddle-point multipliers $\{\lambda_t^*\}$. The sequence of multipliers is selected recursively in such way that the multiplier obtained from the envelope representation of $\partial V(x_{t+1}^*)$ is paired with saddle-point solution x_{t+2}^* . This can be done because the first-order condition (27) holds for every saddle-point multiplier.

Needless to say, solving Euler equations leads, by Proposition 3, to a solution to the problem (18) and therefore a recursive sequence of solutions to Bellman equation (25). A solution to “inconsistent” Euler-like equations where date- t equation has multiplier $\tilde{\lambda}_{t+2}^*$ that may be different from multiplier λ_{t+2}^* featured in date- $t + 1$ equation need not lead to a solution to Bellman equations (25). Note that even a single inconsistency in Euler equations at any one date t may lead to a wrong solution. For example, if Euler equations are re-started at date t from x_t^* and x_{t+1}^* but without memory of previously obtained multiplier λ_{t+1}^* , then the resulting sequence of solutions (from date 1 to infinity) may not be a solution to (18). It is therefore critical that state-variables of the system described by Euler equations include saddle-point multipliers λ_t^* .

If value function V is differentiable, then equations (27) together with equation $\phi_{t+1}^* = \partial V(x_{t+1}^*)$ give rise to first-order difference equations for $\{x_t^*, \lambda_t^*\}$. Any solution to these equations is a sequence of solutions to Bellman equation (25), and under assumption A4 also a solution to (18).

In sum, the intertemporal Euler equations (20) are *necessary* first-order conditions for solution to (18) and by Proposition 3 they are also *sufficient*. The same is true for recursive sequences of solutions to the Bellman equation (24) when they result in consistent selection of multipliers which is the case if: *i*) the value function is differentiable, or *ii*) the selected multiplier applying the envelope theorem (28) (e.g. $\tilde{\lambda}_{t+2}^*$ at x_{t+1}^*) is also selected in solving next period Bellman

equation (24) (at x_{t+1}^*). The intertemporal Euler equations (20) *are necessary and sufficient* for recursive sequences $\{x_t^*\}$ of saddle-point solutions to the Bellman equation (24). However, they *are not necessary but they are sufficient* for a recursive sequences $\{x_t^*, \lambda_t^*\}$ of saddle-points of the Bellman equation (24). In particular, the Euler-like equations are “inconsistent” when date- t equation has multiplier $\tilde{\lambda}_{t+2}^*$ that may be different from multiplier λ_{t+2}^* featured in date- $t + 1$ equation. In fact, “consistency” is not a necessary condition for Bellman equation (24) solutions, since in deriving the first-order conditions in state x_t^* there is no additional information regarding selected past multipliers. Conditions (i) and (ii) place additional restrictions guaranteeing that such information is taking into account.

7 Recursive Contracts

Recursive contract theory, see Marcet and Marimon (2015), provides an extension of Bellman’s dynamic programming to saddle-point problems arising in dynamic optimization problems that fail to have a recursive structure due to the presence of *forward looking* constraints which can not be verified without knowing the future solution path. We first use this approach to solve the dynamic programming problem (18) with time-separable constraints.

7.1 Time-separable constraints

It is possible to derive a *saddle-point Bellman equation* which results in solutions with consistent selection of multipliers. We consider separable intertemporal constraints and apply Recursive Contracts to obtain a recursive formulation of the Euler equations. By appropriately expanding the co-state, we define a *saddle-point Bellman equation* that satisfies the consistency condition.

Let $h_i, i = 1, \dots, k$, satisfying assumptions A4. - A6., take the form $h_i(x, y) = h_i^1(y) - h_i^0(x)$. The Euler equations (20) take the form

$$D_y F(x_t^*, x_{t+1}^*) + \lambda_{t+1}^* D h^1(x_{t+1}^*) + \beta [D_x F(x_{t+1}^*, x_{t+2}^*) - \lambda_{t+2}^* D h^0(x_{t+1}^*)] = 0. \quad (29)$$

Further, we have feasibility and complementary slackness conditions: $h_i^1(x_{t+1}^*) - h_i^0(x_{t+1}^*) \geq 0$ and $\lambda_{t+1}^* [h_i^1(x_{t+1}^*) - h_i^0(x_{t+1}^*)] = 0$.

Following *recursive contracts* procedures, we consider γ as a co-state variable in order to express h_i^1 and h_i^0 contemporaneously and define the following *saddle-point Bellman equation*

$$\begin{aligned} W(x, \gamma) &= \min_{\lambda \geq 0} \max_y \{F(x, y) + \gamma h^1(x) \beta^{-1} - \lambda h^0(x) + \beta W(y, \tilde{\gamma})\} \\ \tilde{\gamma} &= \lambda. \end{aligned} \quad (30)$$

The corresponding envelope (see Appendix) is:

$$\partial_x W(x, \gamma) = \bigcap_{y^* \in Y^*(x, \gamma)} \bigcup_{\lambda^* \in \Lambda^*(x, \gamma)} [D_x F(x, y^*) - \lambda^* D h^0(x) + \gamma D h^1(x) \beta^{-1}]. \quad (31)$$

Notice that, when the *saddle-point Bellman equation* is solved recursively, (31), in contrast with (26), takes into account the past multiplier – say, $\gamma = \lambda_{-1}^*$ – as a predetermined co-state. The first-order condition with respect to y is

$$D_y F(x, y^*) + \beta \tilde{\phi}^* = 0,$$

for some $\tilde{\phi}^* \in \partial W(y^*, \tilde{\gamma})$. Using the envelope theorem, and after the substitution $\tilde{\gamma} = \lambda^*$, the last equation becomes

$$D_y F(x, y^*) + \lambda^* D h^1(y^*) + \beta [D_x F(y^*, z^*) - \tilde{\lambda}^* D h^0(y^*)] = 0.$$

Given $n + k$ initial conditions $(x_0, 0)$, one can derive the Euler equations using selections $\phi_{t+1}^* \in \partial W(x_{t+1}^*, \lambda_{t+1}^*)$ and obtain the system of Euler equations (29). Using the policy function (or policy selection) of the *saddle-point Bellman equation* (30) $(x_{t+1}^*, \lambda_{t+1}^*) = \varphi(x_t^*, \gamma_t^*)$, (29) becomes a $n + k$ system of first-order difference equations, as in the differentiable case.

7.2 Recursive contracts with forward-looking constraints

A canonical example of recursive contracts with *forward looking* constraints is the partnership problem with intertemporal participation constraints where partners pool their resources, but may quit the partnership at any time if their outside options are better. Partnership problems often lead to value functions that are not differentiable.

The deterministic partnership problem can take the form

$$\begin{aligned}
V_\mu(y_0) &= \max_{\{c_t\}_0^\infty} \sum_{t=0}^{\infty} \beta^t \sum_{i=1}^m \mu_i u(c_{i,t}) \\
\text{s.t. } &\sum_{i=1}^m c_{i,t} \leq \sum_{i=1}^m y_{i,t}, \\
&\sum_{n=0}^{\infty} \beta^n u(c_{i,t+n}) \geq v_i(y_{i,t}), \\
&c_{i,t} \geq 0, \text{ for all } t \geq 0,
\end{aligned} \tag{32}$$

where the vector of initial incomes y_0 is given and the sequence of incomes $\{y_t\}$ follows a law of motion $y_{t+1} = g(y_t)$ for some $g : \mathbb{R}_+^m \rightarrow \mathbb{R}_+^m$.⁸ Utility function u is assumed concave and differentiable, and β lies in $(0, 1)$.

The saddle-point problem associated with (32) can be written (see Marcet and Marimon (2015)) as

$$\begin{aligned}
&\max_{\{c_t, \lambda_t\}_{t=1}^\infty} \min_{\{\mu_t\}_{t=1}^\infty} \sum_{t=0}^{\infty} \beta^t \sum_{i=1}^m [\mu_{it} u(c_{i,t}) + (\mu_{i,t+1} - \mu_{i,t}) (u(c_{i,t}) - v_i(y_{i,t})) - \lambda_{i,t+1} (\mu_{i,t+1} - \mu_{i,t})] \\
&\text{s.t. } \sum_{i=1}^m c_{i,t} \leq \sum_{i=1}^m y_{i,t}, \\
&c_{i,t} \geq 0,
\end{aligned} \tag{33}$$

The first-order necessary conditions (with respect to $\mu_{i,t+1}$) for saddle-point $\{c_t^*, \mu_{t+1}^*, \lambda_{t+1}^*\}$ of (33) are

$$u(c_{i,t}^*) - (v_i(y_{i,t}) + \lambda_{i,t+1}^*) + \beta[v_i(y_{i,t+1}) + \lambda_{i,t+2}^*] = 0. \tag{35}$$

for every $t \geq 0$. Equation (35) are the *intertemporal Euler equations* for the partnership problem. Those equations together with first-order conditions for $c_{i,t}$ and $\lambda_{i,t}$, with complementary slackness conditions and the constraints form a system of first-order difference equations. As in section 6, Euler equations and a transversality condition are sufficient conditions for a solution to (33).

Proposition 4: *Let $\{c_t^*, \mu_t^*, \lambda_t^*\}$, with $\mu_0^* = \mu$ and $\lambda_t^* \geq 0$ for every t , satisfy the Euler equations (35). If the transversality condition*

$$\lim_{t \rightarrow \infty} \beta^t [v_i(y_{i,t}) + \lambda_{i,t+1}^*] = 0, \tag{36}$$

⁸The standard risk-sharing problem has $\{y_t\}_{t=0}^\infty$ being a Markov process.

holds, then $\{c_t^*\}$ is a solution to the partnership problem (32).

Marcet and Marimon (2015) show that the value function V_μ satisfies $V_\mu(y_0) = W(y_0, \mu)$, where function W is the value function of the following saddle-point Bellman equation

$$W(y, \mu) = \min_{\tilde{\mu}} \max_c \left\{ \sum_{i=1}^m \mu_i u(c_i) + (\tilde{\mu}_i - \mu_i) (u(c_i) - v_i(y_i)) + \beta W(\tilde{y}, \tilde{\mu}) \right\} \quad (37)$$

$$\text{s.t. } \sum_{i=1}^m c_i \leq \sum_{i=1}^m y_i \quad (38)$$

$$\tilde{\mu}_i \geq \mu_i, \quad (39)$$

$$c_i \geq 0, \text{ for all } i,$$

where $\tilde{y} = g(y)$. If the function u is concave, then the order of maximum and minimum does not matter and the common value is the saddle-value. Further, there exists a saddle-point $(\tilde{\mu}^*, c^*)$. By a solution to the saddle-point Bellman equation we always mean a saddle-point.

For every solution $\{c_t^*\}$ to (32), there exists a sequence $\{\mu_{t+1}^*\}$ such that $\{(c_t^*, \mu_{t+1}^*)\}$ is a sequence of recursive solutions to Bellman equation (37). In contrast to the dynamic programming problem of Section 6, the converse holds true only under an additional consistency condition of Marcet and Marimon (2015). We elaborate on this below.

The value function W is convex in μ . The envelope relation for saddle-point problem (69) in the Appendix implies that the subdifferential of W with respect to μ is

$$\partial_\mu W(y, \mu) = v(y) + \bigcup_{c^*} \Lambda_{c^*}^*(k, y, \mu) \quad (40)$$

where the union is taken over all solutions c^* to (37) and $\Lambda_{c^*}^*(y, \mu)$ denotes the set of saddle-point multipliers of the constraint (39) corresponding to solution c^* . Function W is differentiable with respect to μ at (y, μ) if and only if there is unique multiplier λ^* that is common to all solutions.

Consider a sequence $\{(\mu_{t+1}^*, c_t^*, \lambda_{t+1}^*)\}$ of solutions to the saddle-point Bellman equations (37). The first-order conditions with respect to $\tilde{\mu}$ state that there exists $\phi_{t+1}^* \in \partial W_\mu(y_{t+1}, \mu_{t+1}^*)$ such that

$$u(c_{i,t}^*) - (v_i(y_{i,t}) + \lambda_{i,t+1}^*) + \beta \phi_{i,t+1}^* = 0, \quad (41)$$

where $\phi_{i,t+1}^*$ is the coordinate of subgradient vector ϕ_{t+1}^* corresponding to μ_i . The envelope representation (40) of $\partial W_\mu(y_{t+1}, \mu_{t+1}^*)$ implies that there is $\tilde{\lambda}_{t+2}^* \in \Lambda_{c_{t+1}^*}^*(y_{t+1}, \mu_{t+1}^*)$ such that

$$\phi_{i,t+1}^* = v_i(y_{i,t}) + \tilde{\lambda}_{i,t+2}^* \quad (42)$$

If the value function W is differentiable, then $\tilde{\lambda}_{t+2}^*$ can be taken to be equal to λ_{t+2}^* . Then Euler equations (35) follow. If W is not differentiable, then $\tilde{\lambda}_{t+2}^*$ may be different from λ_{t+2}^* and (35) may fail to hold. In contrast to the dynamic programming problem of Section 6, the existence of some sequence of saddle-point multipliers $\{\lambda_t^*\}$ for which Euler equations hold cannot be guaranteed in the partnership problem. The reason is that the set of saddle-point multipliers $\Lambda_{c_{t+1}}^*(y_{t+1}, \mu_{t+1}^*)$ depends on next-period consumption c_{t+1}^* . Example 7.1 provides an illustration of this issue.

It follows from Theorem 4 in Marcet and Marimon (2015) that any sequence $\{c_t^*, \mu_{t+1}^*, \lambda_{t+1}^*\}$ of recursive solutions to the saddle-point Bellman equation that satisfies an *intertemporal consistency condition* and a transversality condition is a solution to the partnership problem (32). Marcet and Marimon (2015) consistency condition is

$$\phi_{i,t}^* = u(c_{i,t}^*) + \beta \phi_{i,t+1}^*. \quad (43)$$

where $\phi_{t+1}^* \in \partial W_\mu(y_{t+1}, \mu_{t+1}^*)$ satisfies the first-order condition (41). Equation (43) together with transversality condition $\lim_{t \rightarrow \infty} \beta^t \phi_t^* = 0$ imply that

$$\phi_{i,t}^* = \sum_{n=0}^{\infty} \beta^n u(c_{i,t+n}^*). \quad (44)$$

Using (42) and $\lambda_{i,t+1}^* \geq 0$, it follows that

$$\sum_{n=0}^{\infty} \beta^n u(c_{i,t+n}^*) \geq v_i(y_{i,t}). \quad (45)$$

That is, the intertemporal participation constraint holds for $\{c_t^*, \mu_{t+1}^*, \lambda_{t+1}^*\}$. Note that the consistency condition (41) is equivalent to the Euler equation (35) and the transversality $\lim_{t \rightarrow \infty} \beta^t \phi_t^* = 0$ is equivalent to (36).

In sum, a necessary and sufficient condition for sequence $\{\mu_{t+1}^*, c_t^*\}$ generated recursively by a policy function that selects a solution on the right-hand side of (37) to be a solution to the partnership problem (32) is the consistency condition (43) or equivalently that there exists a sequence of multipliers $\{\lambda_t^*\}$ such that Euler equations (35) hold. If value function W is differentiable, then every sequence of recursive solutions to (37) is a solution to (32). Then a *policy function* φ can be defined by $(\mu_{t+1}^*, c_t^*) = \varphi(y_t, \mu_t^*)$ and $\phi_{t+1}^* = \partial W_\mu(g(y_t), \varphi^1(\mu_t^*))$. This is a system of first-order difference equations in $\{c_t^*, \mu_t^*\}$, with initial condition (y_0, μ) .

Messner and Pavoni (2004) provided an example in which value function is non-differentiable and where violations of the consistency condition result in “solutions” that do not satisfy the participation constraint, see Example 7.1.

7.3 Messner and Pavoni (2004) example

Consider a simple version of the partnership problem (32) :

$$\begin{aligned}
V(\mu) = \max_{c_t \geq 0} & \sum_{t=0}^{\infty} \beta^t \sum_{i=1}^2 \mu_i c_{i,t} \\
\text{s.t. } & c_{1,t} + c_{2,t} \leq y, \\
& \sum_{j=0}^{\infty} \beta^j c_{1,t+j} \geq 0, \\
& \sum_{j=0}^{\infty} \beta^j c_{2,t+j} \geq b(1 - \beta)^{-1}, \quad t \geq 0.
\end{aligned} \tag{46}$$

The saddle-point Bellman equation (37) is

$$W(\mu_t) = \min_{\mu_{t+1}} \max_{c_t \geq 0} \left\{ \sum_{i=1}^2 \mu_{i,t+1} c_{i,t} - (\mu_{2,t+1} - \mu_{2,t})b(1 - \beta)^{-1} + \beta W(\mu_{t+1}) \right\} \tag{47}$$

$$\text{s.t. } c_{1,t} + c_{2,t} \leq y, \tag{48}$$

$$\mu_{i,t+1} \geq \mu_{i,t}, \quad i = 1, 2. \tag{49}$$

The first order equations, with respect to μ_{t+1} are

$$c_{1,t}^* - \lambda_{1,t+1}^* + \beta \phi_1(\mu_{t+1}^*) = 0 \tag{50}$$

and

$$c_{2,t}^* - (b(1 - \beta)^{-1} + \lambda_{2,t+1}^*) + \beta \phi_2(\mu_{t+1}^*) = 0, \tag{51}$$

where $\phi(\mu_{t+1}) \in \partial W(\mu_{t+1})$. The complementary slackness conditions associated with constraints (49) are

$$\lambda_{1,t+1}^*(\mu_{1,t+1}^* - \mu_{1,t}^*) = 0 \quad \text{and} \quad \lambda_{2,t+1}^*(\mu_{2,t+1}^* - \mu_{2,t}^*) = 0. \tag{52}$$

By the Envelope Theorem

$$\partial W(\mu_t) = \begin{cases} 0 & + \lambda_{1,t+1}^* \\ b(1 - \beta)^{-1} & + \lambda_{2,t+1}^* \end{cases},$$

where $\lambda_{i,t+1}^* \geq 0$ and, by monotonicity and feasibility, $\lambda_{1,t+1}^* + \lambda_{2,t+1}^* = (y - b)(1 - \beta)^{-1}$.

Let (g^μ, g^c) be a selection from the *saddle-point policy correspondence* Ψ_W of (47) (i.e. $(c_t^*, \mu_{t+1}^*) = (g^\mu(\mu_t), g^c(\mu_t)) \in \Psi_W(\mu_t)$) and $g^\lambda(\mu_t)$ be the corresponding Lagrange multiplier (i.e. $\lambda_{1,t+1}^* = g^\lambda(\mu_t)$); in particular,

$$g^\lambda(\mu_t) = \begin{cases} (y - b)(1 - \beta)^{-1} & \text{if } \mu_1 > \mu_2 \\ 0 & \text{if } \mu_1 < \mu_2 \end{cases},$$

and, correspondingly, the value function is:

$$W(\mu) = \begin{cases} (1 - \beta)^{-1}[\mu_1(y - b) + \mu_2 b] & \text{if } \mu_1 \geq \mu_2 \\ (1 - \beta)^{-1}\mu_2 y & \text{if } \mu_1 \leq \mu_2 \end{cases}.$$

Let $(v_1, v_2) = (0, b(1 - \beta)^{-1})$, then (50) - (51) can be written as

$$g^c(\mu_t) - (v + g^\lambda(\mu_t)) + \beta\phi(g^\mu(\mu_t)) = 0, \quad (53)$$

which is a first-order difference equation in μ_t that, by the *envelope theorem*, can also be written as the Euler equation

$$g^c(\mu_t) - (v + g^\lambda(\mu_t)) + \beta(v + g^\lambda(\mu_{t+1}^*)) = 0, \quad (54)$$

or, more explicitly, as

$$c_{1,t}^* - \lambda_{1,t+1}^* + \beta\lambda_{1,t+2}^* = 0 \quad (55)$$

and

$$c_{2,t}^* - (b + \lambda_{2,t+1}^*) + \beta\lambda_{2,t+2}^* = 0. \quad (56)$$

Alternatively, letting $\omega(\mu_t) \equiv \phi(\mu_t) = v + g^\lambda(\mu_t)$, (54) takes the form

$$\omega(\mu_t) = c_t^* + \beta\omega(\mu_{t+1}^*), \quad (57)$$

showing that the *consistency condition* in Marcet and Marimon (2015; Theorem 4) is the Euler equation of (47), with respect to μ . Furthermore, being derived from (54) the *transversality condition*, $\lim_{n \rightarrow \infty} \beta^n \omega(\mu_{t+n}^*) = 0$, is redundant. However, there is an *implicit intertemporal consistency condition* in the system of Euler equations (57): the solution to (47) at μ_{t+1}^* is given by (53) at μ_{t+1}^* , which not only involves the state μ_{t+1}^* and its change of state, $g^\mu(\mu_{t+1}^*)$, but also the Lagrange multiplier $g^\lambda(\mu_{t+1}^*)$, which already appears in (53) at μ_t^* .

Messner and Pavoni did not have at their disposal the envelope theorem, and the corresponding Euler equation, discussed in this paper, and considered the solution of (47) at μ_{t+1}^* (where $\mu_{1,t+1}^* = \mu_{2,t+1}^*$) not taking into account that $g^\lambda(\mu_{t+1}^*)$ was predetermined by the solution of (47) at μ_t^* . In our formulation, this would correspond to finding an arbitrary solution, such as

$$\widehat{c}_{t+1}^* - (v + \widehat{\lambda}_{t+1}^*) + \beta \widehat{\phi}(\mu_{t+2}^*) = 0,$$

which would violate (57) at μ_t , whenever $(v + \widehat{\lambda}_{t+1}^*) \neq \omega(\mu_{t+1}^*)$.

More precisely, we can solve (46) from an initial $\mu_{1,0} > \mu_{2,0}$. That is, $\lambda_{2,1}^* = 0$ and, using complementary slackness (49), we obtain $\mu_{1,1}^* = \mu_{1,0}$, while $\mu_{2,1}^* = \mu_{1,1}^*$. Using (55) -(56), we obtain

$$c_{1,0}^* = (y - b)(1 - \beta)^{-1} - \beta \lambda_{1,2}^* \quad \text{and} \quad c_{2,0}^* = b - \beta \lambda_{2,2}^* \quad (58)$$

These equations, together with the constraint $c_{1,0}^* + c_{2,0}^* = y$ (i.e. $\lambda_{1,2}^* + \lambda_{2,2}^* = (y - b)(1 - \beta)^{-1}$), determine $c_{1,0}^*, c_{2,0}^*$, given $\lambda_{2,2}^* \in [0, (y - b)(1 - \beta)^{-1}]$. In other words, there is an indeterminacy of period zero solutions parametrized by $\lambda_{2,2}^*$. As we have seen, making a selection $g^\lambda(g^\mu(\mu_0))$ resolves this indeterminacy, but also determines λ_2^* in the first-order conditions (53) at μ_1^* .

In sum, we have shown that using the appropriate envelope theorem and Euler equations the *saddle-point Bellman equation* generates the correct solution to Messner and Pavoni (2004) example. Nevertheless, this example shows the differences between the saddle-point dynamic programming without differentiability and standard dynamic programming, and with saddle-point dynamic programming with differentiability (e.g. with unique solutions). In standard dynamic programming, if there are multiple solutions between two states – say, x_t^* and x_{t+1}^* – all the selections from the *policy correspondence* have the same value. With *forward-looking constraints* this is also true for the aggregate value function W , but not for the individual value functions ω_i . Nevertheless, if W is differentiable there is no selection to be made and, therefore, the Euler

equation (57) at μ_{t+1}^* does not depend on $g^\lambda(\mu_t^*)$.

7.4 General recursive contracts

We extend *recursive contracts* to encompass non-differential value functions. To this end, we consider the following constrained dynamic maximisation problem:

$$\max_{\{x_t\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(\mu, x_t, x_{t+1}) \quad (59)$$

$$\text{s.t. } h_i^1(x_{t+1}) - h_i^0(x_t) \geq 0, \quad i = 1, \dots, k, \text{ and} \quad (60)$$

$$\sum_{n=0}^{\infty} \beta^n d_j^1(x_{t+n}) - d_j^0(x_t) \geq 0, \quad j = 1, \dots, l, \text{ for all } t \geq 0 \quad (61)$$

where $x_t \in \mathfrak{R}^n$ for every t , and μ and $x_0 > 0$ are given. The functions F is a real-valued functions on $\mathfrak{R}^{2n+l}$ and can be interpreted as a social value function where $\mu \in \mathfrak{R}_+^l$ is the vector of weights assigned to the *forward-looking constraints* (61) (e.g. *limited enforcement constraints*). Functions h_i^n and d_j^n , $n = 0, 1$, are real-valued functions on $\mathfrak{R}^n$. In particular, $d_j^1(x_t)$ can be the current utility that agent j gets from an allocation x_t and $d_j^0(x_t)$ her outside option at this state; alternatively, $d_j^1(x_t)$ can represent ‘expected dividends’ and the constraint (61) minimal expected profits requirements. In addition to assumptions A1 - A3, we impose the following conditions:

A4b. F , h_i^n and d_j^n , $n = 0, 1$, are bounded, and $\beta \in (0, 1)$.

A5b. F is a concave function of (x, y) on $X \times X$; h_i^1 and d_i^1 are concave, and h_i^0 and d_i^0 are convex, on X .

A.6b. F and h_i are increasing and continuously differentiable, and F is homogeneous of degree one in μ .

We now proceed to build the appropriate Lagrangian of (59) and its Bellman equation:

Step 1. The basic Lagrangian with respect to (60):

$$\begin{aligned} & \min_{\{\lambda_t\}_{t=1}^\infty, \lambda_t \geq 0} \max_{\{x_t\}_{t=1}^\infty} \sum_{t=0}^{\infty} \beta^t \{F(\mu, x_t, x_{t+1}) + \lambda_{t+1} (h^1(x_{t+1}) - h^0(x_t))\} \\ & \text{s.t. } \sum_{n=0}^{\infty} \beta^n d_j^1(x_{t+n}) - d_j^0(x_t) \geq 0, j = 1, \dots, l, \\ & \text{for all } t \geq 0, \text{ given } x_0 > 0. \end{aligned}$$

Step 2. The recursive Lagrangian accounting for possible non-unique multipliers of (60):

$$\begin{aligned} & \min_{\{\lambda_t\}_{t=1}^\infty, \lambda_t \geq 0} \max_{\{x_t\}_{t=1}^\infty} \sum_{t=0}^{\infty} \beta^t \{F(\mu, x_t, x_{t+1}) + \gamma_t h^1(x_t) \beta^{-1} - \lambda_{t+1} h^0(x_t)\} \\ & \text{s.t. } \sum_{n=0}^{\infty} \beta^n d_j^1(x_{t+n}) - d_j^0(x_t) \geq 0, j = 1, \dots, l, \\ & \gamma_{t+1} = \lambda_{t+1}, t \geq 0, \text{ given } \gamma_0 = 0 \text{ and } x_0 > 0. \end{aligned}$$

Step 3. The recursive Lagrangian with *forward-looking* constraints (61):

$$\begin{aligned} & \min_{\substack{\{\lambda_t, \mu_t\}_{t=1}^\infty, \\ (\lambda_t, \mu_{t+1} - \mu_t) \geq 0}} \max_{\{x_t\}_{t=1}^\infty} \sum_{t=0}^{\infty} \beta^t \{F(\mu_t, x_t, x_{t+1}) + \gamma_t h^1(x_t) \beta^{-1} - \lambda_{t+1} h^0(x_t) \\ & + \mu_t d^1(x_t) - (\mu_{t+1} - \mu_t) (d^1(x_t) - d^0(x_t))\} \\ & \gamma_{t+1} = \lambda_{t+1}, t \geq 0, \text{ given } \gamma_0 = 0, \mu_0 = \mu, \text{ and } x_0 > 0. \end{aligned}$$

Step 4. The recursive Lagrangian accounting for possible non-unique multipliers of $\mu_{t+1} \geq \mu_t$, associated with the recursive formulation of (61):

$$\begin{aligned} & \min_{\substack{\{\lambda_t, \mu_t\}_{t=1}^\infty, \\ (\lambda_t, \mu_{t+1} - \mu_t) \geq 0}} \max_{\{x_t, \eta_t\}_{t=1}^\infty} \sum_{t=0}^{\infty} \beta^t \{F(\mu_t, x_t, x_{t+1}) + \gamma_t h^1(x_t) \beta^{-1} - \lambda_{t+1} h^0(x_t) \\ & + \mu_t d^1(x_t) - (\mu_{t+1} - \mu_t) (d^1(x_t) - d^0(x_t)) + \mu_t (\eta_t - v_t \beta^{-1})\} \\ & \gamma_{t+1} = \lambda_{t+1}, v_{t+1} = \eta_{t+1}, t \geq 0, \text{ given } (\gamma_0, v_0) = 0, \mu_0 = \mu, \text{ and } x_0 > 0. \end{aligned}$$

Step 5. The corresponding *saddle-point Bellman equation* is:

$$\begin{aligned}
W^*(x, \gamma, \mu, v) &= \min_{\substack{\lambda, \tilde{\mu}, \\ \lambda \geq 0}} \max_{\substack{y, \eta, \\ \eta \geq 0}} \left\{ F(\mu, x, y) + \gamma h^1(x) \beta^{-1} - \lambda h^0(x) \right. \\
&\quad \left. + \mu d^1(x) - (\tilde{\mu} - \mu) (d^1(x) - d^0(x)) + \mu (\eta - v \beta^{-1}) + \beta W^*(y, \tilde{\gamma}, \tilde{\mu}, \tilde{v}) \right\}, \quad (62)
\end{aligned}$$

with $\tilde{\gamma} = \lambda$, $\tilde{v} = \eta$.

This allows us to generalise the sufficiency result of Marcet and Marimon (2015):

Theorem 4. Suppose that conditions A.1 - A.3 and A4b - A6b hold and that the value function of the *saddle-point functional equation* W^* is continuous in (x, γ, μ, v) and convex and homogeneous of degree one in (γ, μ, v) . Let $\{x_{t+1}^*, \lambda_{t+1}^*, \mu_{t+1}^*, \eta_{t+1}^*\}_{t=0}^\infty$ be recursively generated by W^* from the initial conditions $(x_0, \gamma_0, \mu_0, v_0) = (x_0, 0, \mu, 0)$, then $\{x_{t+1}^*\}_{t=0}^\infty$ is a solution to (59) at $(x_0, 0, \mu, 0)$. Furthermore, if a new saddle-point solution $\{\hat{x}_{t+n+1}^*\}_{n=0}^\infty$ is generated by W^* from $(x_t, \gamma_t, \mu_t, v_t)$ then $\left(\{x_{n+1}^*\}_{n=0}^{t-1}, \{\hat{x}_{t+n+1}^*\}_{n=0}^\infty\right)$ is also a solution to (59) at $(x_0, 0, \mu, 0)$.

It should be noticed that the last statement is implicit in standard results showing that the Bellman equation is sufficient to generate optimal dynamic solutions. However, it is not necessarily true when the ‘consistency condition’ is not satisfied. It is for this reason that it is made explicit here: with (62) the ‘consistency condition’ is satisfied.

Proof: The proof follows from the previous construction of the *saddle-point Bellman equation* (62), where we have shown that using the Envelope Theorem 2 the intertemporal Euler equations are satisfied when the co-state is expanded to include (γ, v) . Therefore, the assumptions and the ‘consistency condition’ of Marcet and Marimon (2015) Theorem 4 are satisfied and the results follows as a Corollary to their theorem. Alternatively, one can also prove it using the above construction of (62), showing that, as we have seen, they imply that the *forward looking constraints* (61) are satisfied and applying Proposition 3.

8 Appendix

8.1 Saddle-point problems

We extend results of Sections 2 and 3 to saddle-point problems.

Consider the following parametric saddle-point problem

$$\max_{y \in Y} \min_{z \in Z} f(x, y, z) \quad (63)$$

subject to

$$h_i(x, y) \geq 0, \quad g_i(x, z) \leq 0, \quad i = 1, \dots, k \quad (64)$$

with parameter $x \in \mathbb{R}^m$. Let $V(x)$ denote the value function of the problem (63–64). The Lagrangian function associated with (63–64) is

$$\mathcal{L}(x, y, z, \lambda, \gamma) = f(x, y, z) + \lambda h(x, y) + \gamma g(x, z), \quad (65)$$

where $\lambda \in \mathbb{R}_+^k$ and $\gamma \in \mathbb{R}_+^k$ are vectors of multipliers. A saddle point of $\mathcal{L}$ is vector $(y^*, z^*, \lambda^*, \gamma^*)$ where $\mathcal{L}$ is maximized with respect to $y \in Y$ and $\gamma \in \mathbb{R}_+^k$, and minimized with respect to $z \in Z$ and $\lambda \in \mathbb{R}_+^k$. The set of saddle points of $\mathcal{L}$ is a product of two sets M^* and N^* so that $(y^*, \gamma^*) \in M^*$ and $(z^*, \lambda^*) \in N^*$ for every saddle point $(y^*, z^*, \lambda^*, \gamma^*)$. If $(y^*, z^*, \lambda^*, \gamma^*)$ is a saddle point, then (y^*, z^*) is a solution to (63–64).

For single-dimensional parameter x , the directional derivatives the value function V at $x \in \text{int} X$ are

$$V'(x+) = \max_{(y^*, \gamma^*) \in M^*} \min_{(z^*, \lambda^*) \in N^*} \left[\frac{\partial f}{\partial x}(x, y^*, z^*) + \lambda^* \frac{\partial h}{\partial x}(x, y^*) + \gamma^* \frac{\partial g}{\partial x}(x, z^*) \right] \quad (66)$$

and

$$V'(x-) = \min_{(y^*, \gamma^*) \in M^*} \max_{(z^*, \lambda^*) \in N^*} \left[\frac{\partial f}{\partial x}(x, y^*, z^*) + \lambda^* \frac{\partial h}{\partial x}(x, y^*) + \gamma^* \frac{\partial g}{\partial x}(x, z^*) \right] \quad (67)$$

where the order of maximum and minimum does not matter. Corollary 1 can be easily extended to this case.

Suppose that multi-dimensional parameter x can be decomposed in $x = (x^1, x^2)$ so that the constraints in the saddle-point problem (63–64) can be written as

$$h_i(x^1, y) \geq 0, \quad g_i(x^2, z) \leq 0, \quad i = 1, \dots, k. \quad (68)$$

If function f is concave in x^1 and convex in x^2 , while the functions h_i are concave in x^1 and y and the functions g_i are convex in x^2 and z , then $V(x^1, x^2)$ is concave in x^1 and convex in x^2 . The super-sub-differential calculus of Section 4 can be extended to this class of saddle-point problems. For instance, the subdifferential of value function V with respect to x^2 is

$$\partial V_{x^2}(x) = \bigcap_{(z^*, \lambda^*) \in N^*} \bigcup_{(y^*, \gamma^*) \in M^*} \{D_{x^2} f(x, y^*, z^*) + \lambda^* D_{x^2} h(x, y^*) + \gamma^* D_{x^2} g(x, z^*)\}. \quad (69)$$

8.2 Proofs

We first prove the following Lemma:

Lemma 1. Under A1-A3, the sets $Y^*(x)$ and $\Lambda^*(x)$ are non-empty and compact for every $x \in X$. Further, the correspondences Y^* and Λ^* are upper-hemicontinuous on X .

Proof:

Assumptions A1 and A2 imply that the value $V(x)$ of the optimization problem (1–2) is well defined. If λ^* is a saddle-point multiplier at x , then, by the saddle-point property (4), it holds

$$f(x, \hat{y}_i) + \lambda^* h(x, \hat{y}_i) \leq V(x) \quad (70)$$

Using A3, it follows from (70) that

$$\lambda_i^* \leq \frac{V(x) - f(x, \hat{y}_i)}{h(x, \hat{y}_i)} \quad (71)$$

Using $\bar{\lambda}_i$ to denote the RHS of (71), we conclude that the domain $\mathfrak{R}_+^k$ of multipliers in the saddle-point problem (4) can be replaced by the compact set $\times_{i=1}^k [0, \bar{\lambda}_i]$. The standard argument implies now that the set of saddle points is non-empty and compact. The Maximum Theorem implies that saddle-point correspondences Y^* and Λ^* are upper hemi-continuous on X . $\square$

Proof of Theorem 1:

We shall prove that equations (8) and (9) hold for arbitrary $x_0 \in \text{int} X$. Let $\Delta f(t, y)$ denote the difference quotient of function f with respect to x at x_0 , that is

$$\Delta f(t, y) = \frac{f(x_0 + t, y) - f(x_0, y)}{t} \quad (72)$$

for $t \neq 0$. For $t = 0$, we set $\Delta f(0, y) = \frac{\partial f}{\partial x}(x_0, y)$. Assumptions of Theorem 1 imply that function $\Delta f(t, y)$ is continuous in (t, y) on $Y \times \{X - x_0\}$.

Similar notation $\Delta h_i(t, y)$ is used for each function h_i , and $\Delta L(t, y, \lambda)$ for the Lagrangian. Functions $\Delta h_i(t, y)$ are continuous in (t, y) . Note that $\Delta L(t, y, \lambda) = \Delta f(t, y) + \lambda \Delta h(t, y)$, where we use the scalar-product notation $\lambda \Delta h(t, y) = \sum_i \lambda_i \Delta h_i(t, y)$.

The saddle-point property (4) together with (5) imply that

$$V(x_0 + t) \geq \mathcal{L}(x_0 + t, y_t^*, \lambda_t^*) \quad (73)$$

and

$$V(x_0) \leq \mathcal{L}(x_0, y_0^*, \lambda_t^*) \quad (74)$$

for every $\lambda_t^* \in \Lambda^*(x_0 + t)$ and $y_t^* \in Y^*(x_0 + t)$. Subtracting (74) from (73) and dividing the result on both sides by $t > 0$, we obtain

$$\frac{V(x_0 + t) - V(x_0)}{t} \geq \Delta \mathcal{L}(t, y_0^*, \lambda_t^*) = \Delta f(t, y_0^*) + \lambda_t^* \Delta h(t, y_0^*) \quad (75)$$

Since (75) holds for every $y_0^* \in Y^*(x_0)$, we can take the maximum on the right-hand side and obtain

$$\frac{V(x_0 + t) - V(x_0)}{t} \geq \max_{y_0^* \in Y^*(x_0)} [\Delta f(t, y_0^*) + \lambda_t^* \Delta h(t, y_0^*)] \quad (76)$$

Consider function Ψ defined as

$$\Psi(t, \lambda) = \max_{y_0^* \in Y^*(x_0)} [\Delta f(t, y_0^*) + \lambda \Delta h(t, y_0^*)] \quad (77)$$

so that the expression on the right-hand side of (76) is $\Psi(t, \lambda_t^*)$. Since $Y^*(x_0)$ is compact by Lemma 1, it follows from the Maximum Theorem that Ψ is a continuous function of (t, λ) . Further, since $\lambda_t^* \in \Lambda^*(x_0 + t)$ and Λ^* is an upper hemi-continuous correspondence by Lemma 1, we obtain

$$\liminf_{t \rightarrow 0+} \Psi(t, \lambda_t^*) \geq \min_{\lambda_0^* \in \Lambda^*(x_0)} \Psi(0, \lambda_0^*) = \min_{\lambda_0^* \in \Lambda^*(x_0)} \max_{y_0^* \in Y^*(x_0)} \left[\frac{\partial f}{\partial x}(x_0, y_0^*) + \lambda_0^* \frac{\partial h}{\partial x}(x_0, y_0^*) \right] \quad (78)$$

where we used the scalar-product notation $\lambda_0^* \frac{\partial h}{\partial x} = \sum_i \lambda_{i0}^* \frac{\partial h_i}{\partial x}$. It follows from (78) and (76) that

$$\liminf_{t \rightarrow 0+} \frac{V(x_0 + t) - V(x_0)}{t} \geq \min_{\lambda_0^* \in \Lambda^*(x_0)} \max_{y_0^* \in Y^*(x_0)} \left[\frac{\partial f}{\partial x}(x_0, y_0^*) + \lambda_0^* \frac{\partial h}{\partial x}(x_0, y_0^*) \right] \quad (79)$$

We also have

$$V(x_0 + t) \leq \mathcal{L}(x_0 + t, y_t^*, \lambda_0^*) \quad (80)$$

and

$$V(x_0) \geq \mathcal{L}(x_0, y_t^*, \lambda_0^*) \quad (81)$$

which together imply

$$\frac{V(x_0 + t) - V(x_0)}{t} \leq \Delta f(t, y_t^*) + \lambda_0^* \Delta h(t, y_t^*) \quad (82)$$

for $t > 0$. Taking the minimum over $\lambda_0^* \in \Lambda^*(x_0)$ on the right-hand side of (82) results in

$$\frac{V(x_0 + t) - V(x_0)}{t} \leq \min_{\lambda_0^* \in \Lambda^*(x_0)} [\Delta f(t, y_t^*) + \lambda_0^* \Delta h(t, y_t^*)] \quad (83)$$

Consider function Φ defined as

$$\Phi(t, y) = \min_{\lambda_0^* \in \Lambda^*(x_0)} [\Delta f(t, y) + \lambda_0^* \Delta h(t, y)] \quad (84)$$

so that the expression on the right-hand side of (83) is $\Phi(t, y_t^*)$. It follows from the Maximum Theorem that Φ is a continuous function of (t, y) . Using upper hemi-continuity of correspondence Y^* (see Lemma 1), we obtain

$$\limsup_{t \rightarrow 0+} \Phi(t, y_t^*) \leq \max_{y_0^* \in Y^*(x_0)} \Phi(0, y_0^*) = \max_{y_0^* \in Y^*(x_0)} \min_{\lambda_0^* \in \Lambda^*(x_0)} \left[\frac{\partial f}{\partial x}(x_0, y_0^*) + \lambda_0^* \frac{\partial h}{\partial x}(x_0, y_0^*) \right] \quad (85)$$

It follows now from (85) and (83) that

$$\limsup_{t \rightarrow 0+} \frac{V(x_0 + t) - V(x_0)}{t} \leq \max_{y_0^* \in Y^*(x_0)} \min_{\lambda_0^* \in \Lambda^*(x_0)} \left[\frac{\partial f}{\partial x}(x_0, y_0^*) + \lambda_0^* \frac{\partial h}{\partial x}(x_0, y_0^*) \right] \quad (86)$$

It holds (see Lemma 36.1 in Rockafellar (1970)) that

$$\max_{y_0^* \in Y^*(x_0)} \min_{\lambda_0^* \in \Lambda^*(x_0)} \left[\frac{\partial f}{\partial x}(x_0, y_0^*) + \lambda_0^* \frac{\partial h}{\partial x}(x_0, y_0^*) \right] \leq \min_{\lambda_0^* \in \Lambda^*(x_0)} \max_{y_0^* \in Y^*(x_0)} \left[\frac{\partial f}{\partial x}(x_0, y_0^*) + \lambda_0^* \frac{\partial h}{\partial x}(x_0, y_0^*) \right] \quad (87)$$

It follows from (79), (86) and (87) that the right-hand side derivative $V'(x_0+)$ exists and is given by

$$V'(x_0+) = \max_{y_0^* \in Y^*(x_0)} \min_{\lambda_0^* \in \Lambda^*(x_0)} \left[\frac{\partial f}{\partial x}(x_0, y_0^*) + \lambda_0^* \frac{\partial h}{\partial x}(x_0, y_0^*) \right] \quad (88)$$

where the order of maximum and minimum does not matter. This establishes eq. (8) of Theorem

1. The proof of (9) is similar. $\square$

Proof of Theorem 2: By Theorem 23.2 in Rockafellar (1970), $\phi \in \partial V(x_0)$ if and only if $V'(x_0; \hat{x})\phi \leq \hat{x}\phi$ for every $\hat{x}$ such that $x_0 + \hat{x} \in X$. Applying (10), we obtain that $\phi \in \partial V(x_0)$ if and only if

$$\min_{\lambda_0^* \in \Lambda^*(x_0)} \left[D_x f(x_0, y_0^*) + \sum_{i=1}^k \lambda_{i0}^* D_x h_i(x_0, y_0^*) \right] \hat{x} \leq \phi \hat{x} \quad \text{for every } \hat{x}, \quad (89)$$

for every $y_0^* \in Y^*(x_0)$, where we used the fact that inequality (89) holds for every y_0^* if and only if it holds for the maximum over y_0^* . The left-hand side of (89) is the negative of the support function of the set

$$\bigcup_{\lambda_0^* \in \Lambda^*(x_0)} \left\{ D_x f(x_0, y_0^*) + \sum_{i=1}^k \lambda_{i0}^* D_x h_i(x_0, y_0^*) \right\} \quad (90)$$

Since $\Lambda^*(x_0)$ is convex and compact, the set (90) is compact and convex. Theorem 13.1 in Rockafellar (1970) implies that (89) is equivalent to

$$\phi \in \bigcup_{\lambda_0^* \in \Lambda^*(x_0)} \left\{ D_x f(x_0, y_0^*) + \sum_{i=1}^k \lambda_{i0}^* D_x h_i(x_0, y_0^*) \right\} \quad (91)$$

for every $y_0^* \in Y^*(x_0)$. Consequently, $\phi \in \partial V(x_0)$ if and only if

$$\phi \in \bigcap_{y_0^* \in Y^*(x_0)} \bigcup_{\lambda_0^* \in \Lambda^*(x_0)} \left\{ D_x f(x_0, y_0^*) + \sum_{i=1}^k \lambda_{i0}^* D_x h_i(x_0, y_0^*) \right\}. \quad (92)$$

□

Proof of Corollary 2: If $\Lambda^*(x_0)$ is a singleton set or h_i does not depend on x for every i , then

$$\bigcup_{\lambda_0^* \in \Lambda^*(x_0)} \left\{ D_x f(x_0, y_0^*) + \sum_{i=1}^k \lambda_{i0}^* D_x h_i(x_0, y_0^*) \right\} \quad (93)$$

is a singleton set for every y_0^* . The intersection of singleton sets in (14) can either be a singleton set or an empty set. Since V is concave, the subdifferential $\partial V(x_0)$ is non-empty, and hence it must be singleton. This proves that V is differentiable at x_0 . □

Proof of Theorem 3: The proof is similar to that of Theorem 2. Using (10) and Theorem 23.2 in Rockafellar (1970), we obtain that $\phi \in \partial V(x_0)$ if and only if

$$\max_{y_0^* \in Y^*(x_0)} \left[D_x f(x_0, y_0^*) + \sum_{i=1}^k \lambda_{i0}^* D_x h_i(x_0, y_0^*) \right] \hat{x} \geq \phi \hat{x} \quad \text{for every } \hat{x}, \quad (94)$$

for every $\lambda_0^* \in \Lambda^*(x_0)$. The left-hand side of (94) is the support function of the compact (but not necessarily convex) set

$$\bigcup_{y_0^* \in Y^*(x_0)} \{D_x f(x_0, y_0^*) + \sum_{i=1}^k \lambda_{i0}^* D_x h_i(x_0, y_0^*)\}. \quad (95)$$

Theorem 13.1 in Rockafellar (1970) implies that $\phi \in \partial V(x_0)$ if and only if

$$\phi \in \bigcap_{\lambda_0^* \in \Lambda^*(x_0)} co \left(\bigcup_{y_0^* \in Y^*(x_0)} \{D_x f(x_0, y_0^*) + \sum_{i=1}^k \lambda_{i0}^* D_x h_i(x_0, y_0^*)\} \right). \quad (96)$$

□

Proof of Corollary 3: The proof is analogous to that of Corollary 2.

Proof of Proposition 3:

Consider

$$\min_{\{\lambda_t\}_{t=1}^{\infty}} \max_{\{x_t\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left[F(x_t, x_{t+1}) + \sum_{i=1}^k \lambda_{i,t+1} h_i(x_t, x_{t+1}) \right]$$

which results in the following necessary intertemporal Euler equations

$$D_y F(x_t^*, x_{t+1}^*) + \sum_{i=1}^k \lambda_{i,t+1}^* D_y h_i(x_t^*, x_{t+1}^*) + \beta \left[D_x F(x_{t+1}^*, x_{t+2}^*) + \sum_{i=1}^k \lambda_{i,t+2}^* D_x h_i(x_{t+1}^*, x_{t+2}^*) \right] = 0, \quad (97)$$

together with the other first-order Kuhn-Tucker conditions for $i = 1, \dots, k$,

Sufficiency: Let

$$D = \lim_{T \rightarrow \infty} \sum_{t=0}^T \beta^t \{ F(x_t^*, x_{t+1}^*) + \lambda_{t+1}^* h(x_t^*, x_{t+1}^*) - [F(x_t, x_{t+1}) + \lambda_{t+1}^* h(x_t, x_{t+1})] \}. \quad (98)$$

By concavity the RHS is greater than or equal to

$$\begin{aligned} & \lim_{T \rightarrow \infty} \sum_{t=0}^T \beta^t \{ [D_x F(x_t^*, x_{t+1}^*) + \lambda_{t+1}^* D_x h(x_t^*, x_{t+1}^*)] (x_t^* - x_t) \\ & \quad + [D_y F(x_t^*, x_{t+1}^*) + \lambda_{t+1}^* D_y h(x_t^*, x_{t+1}^*)] (x_{t+1}^* - x_{t+1}) \} \end{aligned} \quad (99)$$

since $x_0^* = x_0$. Rearranging terms we have

$$\begin{aligned}
D &\geq \lim_{T \rightarrow \infty} \left\{ \sum_{t=0}^{T-1} \beta^t \left[D_y F(x_t^*, x_{t+1}^*) + \lambda_{t+1}^* D_y h(x_t^*, x_{t+1}^*) \right. \right. \\
&\quad \left. \left. + \beta \left[D_x F(x_{t+1}^*, x_{t+2}^*) + \lambda_{t+2}^* D_x h(x_{t+1}^*, x_{t+2}^*) \right] \right] (x_{t+1}^* - x_{t+1}) \right. \\
&\quad \left. + \beta^T \left[D_y F(x_T^*, x_{T+1}^*) + \lambda_{T+1}^* D_y h(x_T^*, x_{T+1}^*) \right] (x_{T+1}^* - x_{T+1}) \right\}.
\end{aligned}$$

Using transversality condition (23), we obtain

$$\begin{aligned}
D &\geq - \lim_{T \rightarrow \infty} \beta^{T+1} \left[D_x F(x_{T+1}^*, x_{T+2}^*) + \lambda_{T+2}^* D_x h(x_{T+1}^*, x_{T+2}^*) \right] (x_{T+1}^* - x_{T+1}) \\
&\geq - \lim_{T \rightarrow \infty} \beta^{T+1} \left[D_x F(x_{T+1}^*, x_{T+2}^*) + \lambda_{T+2}^* D_x h(x_{T+1}^*, x_{T+2}^*) \right] x_{T+1}^* \\
&= 0
\end{aligned}$$

□

References

- L. Benveniste and J. Scheinkman (1979), “Differentiability of the Value Function in Dynamic Models of Economics” *Econometrica*, 47, 727–732.
- J.-M. Bonnisseau and C. Le Van (1996), “On the Subdifferential of the Value Function in Economic Optimization Problems,” *Journal of Mathematical Economics*, 25, 55–75.
- A. Clausen and C. Straub (2011), “Envelope Theorems for Non-Smooth and Non-Concave Optimization,” mimeo.
- H. Cole and F. Kubler (2012): “Recursive Contracts, Lotteries and Weakly Concave Pareto Sets” *Review of Economic Dynamics* 15(4), 475-500.
- T. Kim (1993), “Differentiability of the Value Function,” *Seul Journal of Economics*, 6, 257–265.
- J. Kyparisis (1985), “On Uniqueness of Kuhn-Tucker Multipliers in Nonlinear Programming,” *Mathematical Programming*, 32, 242–246.
- A. Marcet and R. Marimon (2012), “Recursive Contracts,” EUI; first version EUI-ECO 1998 # 37 WP.
- P. Milgrom and I. Segal (2002), “Envelope Theorems for Arbitrary Choice Sets” *Econometrica*, 70, 583–601.
- M. Messner and N. Pavoni (2004), “On the Recursive Saddle Point Method,” IGIER Working paper 255, Università Bocconi.
- O. Morand, K. Reffett and S. Tarafdar (2011), “A Nonsmooth Approach to Envelope Theorems,” mimeo.
- J. P. Rincón-Zapatero and M. Santos (2009), “Differentiability of the Value Function without Interiority Assumptions.” *Journal of Economic Theory*, 144, 1948-1964.
- R.T. Rockafellar. *Convex analysis*. Princeton University Press, 1970.
- R.T. Rockafellar, *The Theory of Subgradients and its Applications to Problems of Optimization. Convex and Nonconvex Functions*. Heldermann Verlag, Berlin, 1981.
- N.L. Stokey, R. E. Lucas and E. C. Prescott (1989), *Recursive Methods in Economic Dynamics*. Cambridge, Ma.: Harvard University Press.