J Optim Theory Appl
https://doi.org/10.1007/s10957-018-1241-5

Generalized Envelope Theorems: Applications to Dynamic Programming

Olivier Morand[1] · **Kevin Reffett**[2] · **Suchismita Tarafdar**[3]

Received: 6 May 2017 / Accepted: 5 February 2018

Abstract We show in this paper that the class of Lipschitz functions provides a suitable framework for the generalization of classical envelope theorems for a broad class of constrained programs relevant to economic models, in which nonconvexities play a key role, and where the primitives may not be continuously differentiable. We give sufficient conditions for the value function of a Lipschitz program to inherit the Lipschitz property and obtain bounds for its upper and lower directional Dini derivatives. With strengthened assumptions we derive sufficient conditions for the directional differentiability, Clarke regularity, and differentiability of the value function, thus obtaining a collection of generalized envelope theorems encompassing many existing results in the literature. Some of our findings are then applied to decision models with discrete choices, to dynamic programming with and without concavity, to the problem of existence and characterization of Markov equilibrium in dynamic economies with nonconvexities, and to show the existence of monotone controls in constrained lattice programming problems.

Communicated by Michel Théra.

✉ Suchismita Tarafdar
suchismita.tarafdar@snu.edu.in

Olivier Morand
olivier.morand@uconn.edu

Kevin Reffett
kevin.reffett@asu.edu

1 Department of Economics, University of Connecticut, Storrs, CT, USA

2 Department of Economics, WP Carey School of Business, Arizona State University, Tempe, AZ, USA

3 Department of Economics, Shiv Nadar University, Greater Noida, India

Published online: 22 February 2018

Keywords Envelope theorem · Lipschitz function · Clarke regularity · Supermodularity · Lipschitz dynamic programming · Recursive dynamic program

1 Introduction

Envelope theorems for constrained optimization problems have been an important tool for both microeconomic and macroeconomic analyses. In its standard or "classical" form, an envelope theorem is simply an equality between the derivative of the value function and the derivative of the objective evaluated at the optimum, ignoring the indirect effects due to changes in the optimal solution. In once continuously differentiable ("smooth" hereafter) and convex programs, envelopes are typically standard derivatives giving precise information about the rate of growth of the value function in all directions at a given point and are essential for comparative statics in many economic models.

Programs with nonsmooth or nonconvex objectives and constraints are, however, very common in economics, appearing, for instance, in dynamic growth models, in constrained lattice programming problems, in incentive constrained dynamic programs, and in "bi-level"/Stackelberg games. In the absence of smoothness and convexity, one cannot expect envelopes to be simple derivatives, since Lagrange multipliers may not be unique and traditional derivatives (or even subgradients of convex analysis) may not be defined. In these instances, classical envelope theorems are usually of little help.

Nevertheless, generalizations of the classical envelope theorem have been derived for economic models represented by nonconvex or nonsmooth programs. Amir et al. [1], for instance, gave sufficient conditions for the Gâteaux differentiability of the value function in a large class of nonconvex growth models. Bonnisseau and LeVan [2] developed some fundamental results on the subdifferentiability of the value function for convex, but not necessarily smooth programs, and Askri and LeVan [3] pioneered the use of Clarke gradient to prove the differentiability of the value function in economic models with continuous- and compact-valued choice domains under interiority of solutions. Milgrom and Segal [4] derived various results for unconstrained Gâteaux differentiable programs, subsequently extended to convex and smooth programs, and, more recently, Rincon-Zapatero and Santos [5] proved the continuous differentiability of the value function in smooth concave dynamic programming without interiority of solutions.

This variety of results highlights the need for a comprehensive and systematic way of deriving generalized envelope theorems for nonconvex and/or nonsmooth programs. An important motivation for this paper is to show that the class of Lipschitz functions provides an environment well suited for the generalization of classical envelope theorems and for the unification of existing findings in the economic literature. We extend the results of Morand et al. [6] to settings, in which both constraints and objectives are not necessarily smooth or concave, but at least Lipschitz, and put special emphasis on Clarke regularity. Our paper seeks to contribute to the theoretical literature on nonsmooth optimization with illustrations in dynamic economic models.

Generically not everywhere differentiable, Lipschitz functions have some very important features. First, the Lipschitz property is a verifiable hypothesis satisfied in a wide variety of settings, in particular in the presence of convexity or continuous differentiability. Second, Lipschitz functions are at the core of a well-developed theory of differentiability extending that of convex functions (see Clarke [7]) and as such have played an important role in nonexpected utility (see Chatterjee and Krishna [8]) and in stochastic dynamic programming without concavity (see Maroto and Moran [9]). Lipschitz functions on a compact domain are also absolutely continuous and therefore equal to the integral of their almost everywhere derivative, a property exploited, for instance, by Milgrom and Segal [4]. Finally, the set of (upper) Clarke regular Lipschitz functions should be of particular interest to economists since it contains all convex functions.

We demonstrate, in Sect. 2 of this paper, that the Lipschitz property is preserved by maximization under relatively weak hypothesis. One of these hypotheses is a nonsmooth constraint qualification related to the work of Hiriart Urruty [10] and Auslender [11] and easily checked in many applications. The other, taken from the work of Clarke [7], imposes restrictions on the choice domain and the objective, which are satisfied under some very general conditions in many economic models.

Having established that the value function is Lipschitz, in Sect. 3 of the paper we give lower and upper bound estimates of its Dini derivatives, a result useful in computational work and for proving the absolute continuity of the value function (a critical step in the proof of its supermodularity in a large class of models, as shown in Sect. 4). The rest of Sect. 3 consists in narrowing these bounds to sharpen the characterization of the rates of growth of the value function in specific directions. Gâteaux differentiability is the next step, since it permits comparative statics in all directions at a specific point, followed by Clarke regularity (a step short of continuous differentiability), which is shown to be preserved under maximization under some conditions. Because continuously differentiable functions are necessarily Lipschitz and, both upper and lower Clarke regular, while convex functions are upper Clarke regular, classical envelope theorems are just special cases of our more general results.

In Sect. 4, our results are applied to dynamic programming with and without concavity, to decision models with discrete choices, and to a proof of existence of equilibrium in a large class of dynamic models with nonconvexities and nonsmooth primitives.

2 Lipschitz Programs

Since neither the set of convex functions nor the set of continuously differentiable functions are best suited for the study of economic models with nonconvexities or nonsmooth primitives, we focus on the larger set of Lipschitz continuous real-valued functions and consider Lipschitz programs of the form:

$$\max_{a \in D(s)} f(a, s), \tag{1}$$

in which $f : A \times S \to \mathbb{R}$ is the objective function and $D : S \rightrightarrows A$ is the feasible correspondence defined as:

$$D(s) = \{a : g_i(a, s) \leq 0, \quad i = 1, \dots, p \text{ and } h_j(a, s) = 0, \quad j = 1, \dots, q\},$$

where $g_i : A \times S \to \mathbb{R}$, $i = 1, \dots, p$, and $h_j : A \times S \to \mathbb{R}$, $j = 1, \dots, n$.

The choice set A and the parameter space S are both open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and functions f, g_i and h_j are at first only assumed to be Lipschitz at every $(a, s) \in A \times S$. Note that the continuity of g_i and h_j implies that the feasible correspondence D is closed at any s. When all functions are continuously differentiable, we refer to program (1) as a smooth program.[1]

Function $V : S \to \mathbb{R}$, defined as $V(s) = \max_{a \in D(s)} f(a, s)$, is the value function, and $A^* : S \rightrightarrows A$ is the optimal solution correspondence, defined as $A^*(s) = \arg\max_{a \in D(s)} f(a, s)$. The Lagrangian associated with the above program is:

$$L(a, s, \lambda, \mu) = f(a, s) - \lambda g(a, s) - \mu h(a, s),$$

where λ and μ are vectors in $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively. If A is closed, then the abstract constraint $a \in A$ induces an additional term in the Lagrangian (see, for instance, Clarke [7], Chapter 6).

2.1 Constraint Qualifications

We first recall the definition of a Karush–Kuhn–Tucker (KKT) point, which requires the existence of a vector of multipliers satisfying a specific "multiplier rule" stated in terms of Clarke gradients.

Definition 2.1 Given $s \in S$, $a \in D(s)$ is a *KKT* point of Program (1), if there exists a vector $(\lambda, \mu) \in \mathbb{R}^p_+ \times \mathbb{R}^q$ such that:

$$0 \in \partial_a \left(f - \sum_{i=1}^{p} \lambda_i g_i - \sum_{j=1}^{q} \mu_j h_j \right) (a, s)$$

and $\lambda_i g_i(a, s) = 0$ for all $i = 1, \dots, p$.

Denoting by $K(a, s)$ the closed, convex, but possibly empty set of vectors (λ, μ) satisfying the above "multiplier rule" at (a, s), one role of a constraint qualification (CQ) is to eliminate the trivial cases in which $K(a, s)$ is empty. Alternatively, a CQ can be seen as a set of sufficient conditions for the existence of nontrivial Fritz-Jones multipliers.

2.1.1 Smooth Programs

A standard CQ for smooth programs is the Mangasarian–Fromovitz CQ (hereafter, denoted by MFCQ), which when satisfied implies the existence of a direction belonging

[1] We refer the reader to Appendices for a brief summary of the properties of correspondences and of Lipschitz functions.

to both the cone of inward directions of active inequality constraints and the set of tangent directions of equality constraints.[2]

Gauvin [12] proved in smooth programs that the MFCQ at $a^*(s) \in A^*(s)$ is equivalent to the compactness of $K(a^*(s), s)$. This result was sharpened by Kyparisis [13], who showed that the strict MFCQ, a slightly less general condition than MFCQ, is both necessary and sufficient for the set $K(a^*(s), s)$ to be a singleton in smooth programs (see also Bonnans and Shapiro [14]) for each $a^*(s) \in A^*(s)$.

Both CQs, the latter treating active inequality constraints for which multipliers are strictly positive ("binding constraints") as equality constraints, are stated next.

Definition 2.2 The *MFCQ is satisfied at* $a^*(s) \in A^*(s)$ if there exists $y \in \mathbb{R}^n$ such that:

$$\begin{aligned} &\nabla_a g_i(a^*(s), s) \cdot y < 0, \quad i \in I(a^*(s), s), \\ &\nabla_a h_j(a^*(s), s) \cdot y = 0 \quad j = 1, \ldots, q, \end{aligned}$$

where $I(a^*(s), s)$ is the set of indexes of the active inequality constraints (those for which the constraints, $g_i(a^*(s), s) = 0$), and the matrix $\nabla_a h(a^*(s), s)$ has full rank. The strict Mangasarian–Fromovitz constraint qualification (SMFCQ) is satisfied at optimal point $a^*(s) \in A^*(s)$ if there exists $y \in \mathbb{R}^n$ such that:

$$\begin{aligned} &\nabla_a g_i(a^*(s), s) \cdot y < 0, \quad i \in I_s(a^*(s), s) \\ &\nabla_a g_i(a^*(s), s) \cdot y = 0, \quad i \in I_b(a^*(s), s) \\ &\nabla_a h_j(a^*(s), s) \cdot y = 0 \quad j = 1, \ldots, q, \end{aligned}$$

where $I_s(a^*(s), s) = \{i \in I(a^*(s), s), \lambda_i = 0\}$, $I_b(a^*(s), s) = \{i \in I(a^*(s), s), \lambda_i > 0\}$, and the vectors $\nabla_a g_i(a^*(s), s)$, $i \in I_b(a^*(s), s)$, $\nabla_a h_j(a^*(s), s)$, $j = 1, \ldots, q$, are linearly independent.

2.1.2 Lipschitz Programs

Classical gradients generically do not exist for Lipschitz functions, so we rely on a generalization of the MFCQ (referred to as the "generalized MFCQ," or GMFCQ) introduced by Hiriart Urruty [10] and stated in terms of Clarke gradients. We denote by $\overline{g}(a^*(s), s)$ the vector of active inequality constraints at $a^*(s)$ ($\overline{g} : A \times S \to \mathbb{R}^{\overline{p}}$, where $\overline{p} = Card(I(a^*(s), s)))$.

Definition 2.3 The *GMFCQ is satisfied at* $a^*(s) \in A^*(s)$ if there exists $y \in \mathbb{R}^n$ such that:

$$\forall (\gamma, \upsilon) \in \partial_a(\overline{g}, h)(a^*(s), s), \;\; \gamma \cdot y < 0 \text{ and } \upsilon \cdot y = 0$$

and $\partial_a h(a^*(s), s)$ is of maximal rank.

[2] Such geometrical feature is critical to permit some form of sensitivity analysis in parameterized optimization problems and points at another important role of CQs.

We note Hiriart Urruty [10] proved that the GMFCQ at $a^*(s) \in A^*(s)$ implies the nonemptiness of $K(a^*(s), s)$, as well as:

$$\partial_a \overline{g}(a^*(s), s) \subset \prod_{i \in I(a^*(s),s)} \partial_a g_i(a^*(s), s).$$

So this version of the GMFCQ is slightly more general than that used in Auslender [11].

2.2 Lipschitz Value Functions

The Lipschitz properties of the primitives of program (1) are, of course, not sufficient for the value function V to even be continuous. Additional restrictions are therefore needed to ensure that V is Lipschitz, so we adopt the following general hypothesis made in Clarke [7] (Hypothesis 6.5.1, page 241).

Criterion 2.1 (Clarke's hypothesis) *$V(s)$ is finite, and there exists a compact set Λ and $\varepsilon_0 > 0$ such that for all $s' \in \varepsilon_0 B(s)$ for which $V(s') \geq V(s) - \varepsilon_0$, and $A^*(s') \cap \Lambda \neq \emptyset$.*

Theorem 2.1 *If the GMFCQ holds at any $a^*(s) \in A^*(s)$, and Clarke's hypothesis is satisfied, then V is Lipschitz and:*

$$\partial V(s) \subset cl \quad conv \left\{ \bigcup_{a^*(s) \in A^*(s)} \bigcup_{(\lambda,\mu) \in K^*(a_n^*(s),s)} \partial_s (f - \lambda g - \mu h)(a^*(s), s) \right\}.$$

Moreover, if the sequence $\{s_n\}$ converges to s, and if $a_n^* \in A^*(s_n) \cap \Lambda$, then there is a subsequence $\{a_n^*\}$ that converges to $a^* \in A^*(s)$.

Proof The Lipschitz property of V and the formula for the generalized gradient follow directly from Clarke [7] (Corollary 1, page 242). By continuity of V at s, $\exists \delta$ such that $s_n \in \delta B(s) \cap \varepsilon_0 B(s)$ implies $V(s_n) \geq V(s) - \varepsilon_0$; hence, by Clarke's hypothesis $A^*(s_n) \cap \Lambda \neq \emptyset$. Any sequence $\{a_n^*\}$ with $a_n^* \in A^*(s_n) \cap \Lambda$ is in the compact Λ and therefore contains a subsequence converging to some a^*. Now, D is closed at s; hence, $a^* \in D(s)$. By the continuity of V and f, $V(s_n) = f(a_n, s_n) \to f(a, s) = V(s)$; hence, $a^* \in A^*(s) \cap \Lambda$. □

Clarke's hypothesis is not expressed in terms of primitives and therefore cannot be immediately checked. Nevertheless, it is easily shown to be satisfied if any of the following three conditions is satisfied[3]:

(a) Clarke's growth condition (Clarke [7]): $\forall r \in \mathbb{R}$, $\{(a', s') \in A \times S,\ f(a', s') \geq r\}$ is compact;

[3] A fourth condition, in the form of a mild compactness restriction, is discussed in the next section of this paper.

(b) Inf-compactness condition (Bonnans and Shapiro [14]): There exist $r \in \mathbb{R}$ and a compact set $\Omega \subset A$ such that for every s' in a neighborhood of s the set $\{a' \in D(s'),\ f(a', s') \geq r\}$ is nonempty and contained in Ω;
(c) Uniform compactness condition (Gauvin and Dubeau [15]): There exists a neighborhood S' of s such that $cl\,[\cup_{s' \in S'} D(s)]$ is compact.

It is important to note that, when combined with the GMFCQ, any one of the above conditions implies a very powerful result analogous to Berge's maximum theorem: The value function is Lipschitz, and the set of optimal solutions $A^*(s)$ is upper hemicontinuous; the latter property is very important since it guarantees that as s_n converges to s, the maxima of $f(., s_n)$ become arbitrarily close to some of the maxima of $f(., s)$.

Proposition 2.1 *If the GMFCQ holds at any $a^*(s) \in A^*(s)$, and if any of the conditions* (a), (b), *or* (c) *above is satisfied, then the hypothesis of Theorem 2.1 holds and the optimal solution correspondence A^* is upper hemicontinuous at s.*

Proof We prove the upper hemicontinuity independently for each condition:

(a) Given any $s_n \to s$ and any sequence $\{a_n\}$ such that $a_n \in A^*(s_n) \subset D(s_n)$, by continuity of V, $V(s_n) = f(a_n, s_n)$ converges to $V(s)$. As a result, $\forall \varepsilon' > 0$ there exists N such that $\forall n \geq N$, $f(a_n, s_n) \geq V(s) - \varepsilon'$. The sequence $\{(a_n, s_n)\}_{n \geq N}$ belongs to the (compact, by hypothesis) set $\{(a', s') \in A \times S,\ f(a', s') \geq V(s) - \varepsilon'\}$ and therefore has a subsequence converging to (a, s) for some a. By continuity of f, $f(a_n, s_n) = V(s_n)$ converges to $f(a, s)$; hence, $f(a, s) = V(s)$. Finally, since D is closed at s, necessarily $a \in D(s)$. Thus, $a \in A^*(s)$.
(b) By the inf-compactness condition, $\exists r$ such that $A_s = \{a' \in D(s'),\ f(a', s') \geq r\}$ is nonempty for all $s' \in \delta B(s)$ and is included in a compact set Ω. Thus, there exists N such that $\forall n \geq N$, $a_n \in A^*(s_n) \subset A_s \subset \Omega$, and the sequence $\{(a_n, s_n)\}_{n \geq N}$ has a convergent subsequence to (a, s). By continuity of V and f $V(s_n) = f(a_n, s_n) \to f(a, s) = V(s)$ and since D is closed at s, the desired result follows.
(c) Since V is continuous at s, the map $L : s \to \{a,\ f(a, s) - V(s) \geq 0\}$ is closed at s. Under the uniform compactness condition, since $A^*(s) = L(s) \cap D(s)$, the correspondence $A^* : s \to A^*(s)$ is the intersection of the closed mapping L with the upper hemicontinuous (since closed and uniformly compact) mapping D. Consider then $s_n \to s$ and any $a_n \in A^*(s_n) = L(s_n) \cap D(s_n)$. Since D is upper hemicontinuous at s, there exists a subsequence of a_n converging to some $a \in D(s)$. Since L is closed at s, the limit a of the subsequence of a_n necessarily belongs to $L(s)$. Thus, $a \in A^*(s) = L(s) \cap D(s)$, which proves that A^* is upper hemicontinuous at s. □

Note again that the continuity of V cannot come directly from Berge's maximum theorem since the feasible correspondence D is not necessarily continuous, even though all constraints are continuous. For example, the correspondence D defined as:

$$D(s) = \{(x, y), x + y \leq s \ \text{ and } \ (s - 11)(10 - x) \leq 0\}$$

is not continuous at $s = 11$.

3 Generalized Envelope Theorems

The conditions for the preservation of the Lipschitz property under maximization, derived in the previous section, are shown below to be sufficient for the existence of specific bounds for the Dini derivatives of the value function. After proving this important result, we impose, in the rest of this section, additional restrictions on the primitives (such as concavity, differentiability, Clarke regularity, continuous differentiability) to derive sharper envelope theorems going beyond the simple existence of bounds all the way to precise C^1 envelopes.

3.1 A Central Result on Stability Bounds

Under the conditions of Theorem 2.1, the value function is Lipschitz and has Dini derivatives. Specific bounds for these Dini derivatives are obtained as a consequence of a result in Clarke [7], Corollary 4 (page 243) (see also Tarafdar [16] for an alternative proof independent of Clarke's results), and can be expressed in terms of the primitives.

Theorem 3.1 *If the GMFCQ holds at any $a^*(s) \in A^*(s)$, then under Criterion 2.1, for any $x \in \mathbb{R}^m$:*

$$D^+V(s;x) \leq \max_{a^*(s)\in A^*(s)} \left(\sup_{\lambda \in K(a^*(s),s)} \left(\max_{\theta \in \partial_s(f-\lambda g-\mu h)(a^*(s),s)} \{\theta \cdot x\} \right) \right)$$

and

$$\max_{a^*(s)\in A^*(s)} \inf_{\lambda\in K(a^*(s),s)} \left(\min_{\theta\in\partial_s(f-\lambda g-\mu h)(a^*(s),s)} \{\theta\cdot x\}\right) \leq D_+V(s;x).$$

Proof Omitting the equality constraints to simplify the notations (since equality constraints are associated with another multiplier), the Lipschitz program (1) becomes:

$$-V(s) = \min -f(a,s) \text{ s.t. } g(a,s) \leq 0$$

and is identical to the "modified program":

$$-V(s) = \min -f(a,a') \quad \text{s.t. } g(a,a') \leq 0 \text{ and } -a' + s = 0$$

with its associated Lagrangian[4]:

$$L_m((a,a'),s,\lambda,\theta) = -f(a,a') + \lambda g(a,a') + \theta(-a'+s).$$

The two programs have the same set of solutions, in the sense that $a^*(s) \in A^*(s)$ if and only if $(a^*(s),s) \in A^*_m(s)$, and the same set of multipliers, i.e., $\lambda \in K(a^*(s),s)$ if and only if $(\lambda,\theta) \in K_m(a^*(s),s)$.

[4] The subscript m is used to identify objects relevant to the "modified program."

By Theorem 6.1.1 in Clarke [7], there exist $\lambda \geq 0$ and θ such that:

$$\lambda g(a^*(s), s) = 0 \text{ and } 0 \in \partial_{(a,a')} L((a^*(s), s), s, \lambda, \theta).$$

The latter condition implies the existence of $(\sigma_a + \lambda\gamma_a) \in \partial_a(-f + \lambda g)(a^*(s), s)$ and of $(\sigma_{a'} + \lambda\gamma_{a'}) \in \partial_{a'}(-f + \lambda g)(a^*(s), s)$ such that, for all (u, v):

$$0 = (\sigma_a + \lambda\gamma_a)u + (\sigma_{a'} + \lambda\gamma_{a'})v - \theta v.$$

As a result, necessarily:

$$\sigma_a + \lambda\gamma_a = 0$$

and

$$\theta = \sigma_{a'} + \lambda\gamma_{a'} \in \partial_{a'}(-f + \lambda g)(a^*(s), s).$$

The assumptions of Corollary 4 in Clarke [7] are satisfied (in Clarke's notations, if the GMFCQ holds at each $a^*(s)$ then $M^0(\sum) = \{0\}$, where $\sum = A^*(s)$); hence,

$$D^+(-V)(s; x) \leq \sup_{(\lambda,\theta)\in M^1(a,s)} \{\theta \cdot x\} = \sup_{\lambda\in K(a^*(s),s)} \sup_{\theta\in\partial_{a'}(-f+\lambda g)(a^*(s),s)} \{\theta \cdot x\}$$

and

$$D_+(-V)(s; x) \geq \inf_{a^*(s)\in A^*(s)} \inf_{\lambda\in K(a^*(s),s)} \inf_{\theta\in\partial_s(-f+\lambda g)(a^*(s),s)} \{\theta \cdot x\}.$$

Since $D^+(-V)(s; x) = -D_+V(s; x)$ and that $D_+(-V)(s; x) = -D^+V(s; x)$, we obtain:

$$\begin{aligned} D^+V(s; x) &\leq - \inf_{a^*(s)\in A^*(s)} \inf_{\lambda\in K(a^*(s),s)} \inf_{\theta\in\partial_s(-f+\lambda g)(a^*(s),s)} \{\theta \cdot x\} \\ &= \sup_{a^*(s)\in A^*(s)} \sup_{\lambda\in K(a^*(s),s)} \sup_{\theta\in\partial_s(f-\lambda g)(a^*(s),s)} \{\theta \cdot x\} \end{aligned}$$

and

$$D_+V(s; x) \geq \inf_{\lambda\in K(a^*(s),s)} \inf_{\theta\in\partial_s(f-\lambda g)(a^*(s),s)} \{\theta \cdot x\}.$$

Since $a^*(s)$ in $A^*(s)$ is arbitrary in the last inequality, the theorem is proven, noting that both sets $\partial_s(f - \lambda g)(a^*(s), s)$ and $A^*(s)$ are compact sets so inf and sup, respectively, become min and max. □

Application Theorem 1 of Milgrom and Segal [4] is a direct consequence of this result when applied to the unconstrained version of Program (1):

$$V(s) = \sup_{a\in A} f(a, s).$$

In addition to giving precise information about the bounds for the rate of growth of the value function at any point, the above result also suggests that further differentiability properties of the value function may be derived by combining (i) a stronger CQ (to get uniqueness of the Lagrange multiplier), (ii) continuous differentiability hypothesis (for Clarke gradients to be singletons), and (iii) concavity assumptions (to get a unique optimal solution).

3.2 Clarke Regularity

While the optimal value function is generally not everywhere differentiable, sharp bounds or estimates of its rate of growth can be obtained by a directional analysis, providing that additional regularity conditions beyond the Lipschitz structure of the primitives are in place.

While we consider the traditional assumptions of concavity and differentiability below, we first discuss the role of Clarke regularity, a property which has received little attention in economic models. (Notable exceptions are Askri and LeVan [3] and Bonnisseau and LeVan [2].) The importance of Clarke regularity stems from the property that Clarke regular Lipschitz value functions have directional derivatives. In addition, intuitively, upper Clarke regularity at x gives "local" information, as it implies that the derivative in some direction d approximates the maximum rate of growth of the value function in the direction d in a whole neighborhood of x.

3.2.1 Locally Relaxing the Constraints

We begin with Lipschitz programs satisfying a mild compactness condition implying that constraints can be locally ignored in a neighborhood of a particular point s.

Criterion 3.1 *There exist a compact set $\Lambda \subset A$ and a compact neighborhood $N(s)$ of s such that $\forall s' \in N(s) \subset S$, $A^*(s') \subset \Lambda \subset D(s')$.*

We show this above condition is sufficient for the upper Clarke regularity[5] of the objective to be preserved to the value function under the maximization operation, noting in the process that Clarke's hypothesis (Criterion 2.1) is easily satisfied.

Proposition 3.1 *Under Criterion 3.1, if f is Lipschitz, then V is Lipschitz at s. Furthermore, if f is upper Clarke regular at s for each $a^*(s) \in D(s)$, then V is upper Clarke regular at s and*

$$V'(s; x) = \max_{a^*(s) \in A^*(s)} f_s(a^*(s), s; x).$$

Proof V is continuous at s by Berge's theorem of the maximum, since by hypothesis

$$V(s) = \max_{a \in \Lambda} f(a, s).$$

[5] Lower Clarke regularity for minimization programs.

Furthermore, Criterion 2.1 is satisfied by choosing any $\varepsilon_0 > 0$ such that $\varepsilon_o B(s) \subset N(s)$, so that $A^*(s') \subset \Lambda$ for any $s' \in \varepsilon_o B(s)$. V is therefore Lipschitz at s by Theorem 2.1.

The upper Clarke regularity of f permits squeezing together the Dini bounds of Theorem 3.1. Indeed, for all $a^*(s) \in A^*(s)$

$$\begin{aligned}\liminf_{t\downarrow 0} \frac{V(s+tx)-V(s)}{t} &= \liminf_{t\downarrow 0} \frac{f(a^*(s+tx), s+tx) - f(a^*(s), s)}{t} \\ &\geq \liminf_{t\downarrow 0} \frac{f(a^*(s), s+tx) - f(a^*(s), s)}{t} \\ &= f_s(a^*(s), s; x),\end{aligned}$$

where the last equality results from the Gâteaux differentiability of f in s for each $a^*(s)$. In addition,

$$\begin{aligned}\limsup_{t\downarrow 0} \frac{V(s+tx)-V(s)}{t} &\leq \max_{a^*(s)\in A^*(s)} \max_{\varsigma\in\partial f(a^*(s),s)} \{\varsigma \cdot x\} \\ &= \max_{a^*(s)\in A^*(s)} f^o(a^*(s), s; x) \\ &= \max_{a^*(s)\in A^*(s)} f_s(a^*(s), s; x),\end{aligned}$$

where the last equality follows from the upper Clarke regularity of f in s for each $a^*(s)$. Since upper and lower Dini's coincide, V is Gâteaux differentiable and

$$V'(s;x) = \max_{a^*(s)\in A^*(s)} f_s(a^*(s), s; x) = \max_{a^*(s)\in A^*(s)} f^o(a^*(s), s; x).$$

By definition of the upper Clarke derivative, for $a^*(s) \in A^*(s)$:

$$f^o(a^*(s), s; x) = \max_{\xi\in\partial_s f(a^*(s),s)} \{\xi \cdot x\}$$

and by Theorem 2.1:

$$\partial V(s) \subset \text{cl conv}\left\{\cup_{a^*(s)\in A^*(s)} \partial_s f(a^*(s), s)\right\};$$

thus,

$$\max_{\xi\in\partial V(s)} \{\xi\cdot x\} \leq \max_{\xi\in \text{cl conv}\{\cup_{a^*(s)\in A^*(s)}\partial_s f(a^*(s),s)\}} \{\xi\cdot x\} \leq \max_{a^*(s)\in A^*(s)} f^o(a^*(s), s; x).$$

As a result,

$$\begin{aligned}V'(s;x) \leq V^o(s;x) &= \max_{\xi\in\partial V(s)} \{\xi\cdot x\} \\ &\leq \max_{a^*(s)\in A^*(s)} f^o(a^*(s), s; x) = V'(s;x);\end{aligned}$$

hence, V is upper Clarke regular at x. □

There are, however, economic models in which the objective is not necessarily jointly Lipschitz in (a, s) (and possibly not even continuous in a for each s). Such is the case, in particular, in deterministic dynamic programs where the value function V is defined recursively through Bellman's equation:

$$V(s) = \max_{a \in \Lambda} \{U(a, s) + \beta V(a)\} \tag{2}$$

and for which Proposition 3.1 clearly does not apply. Nevertheless, it is possible to strengthen some of the hypotheses on the objective to ensure the preservation of upper Clarke regularity, as shown in the next result.

Proposition 3.2 *Under Criterion 3.1, if the correspondence $(a', s') \rightrightarrows \partial_s f(a', s')$ is closed at every $(a^*(s), s)$, and f is upper Clarke regular at s for each $a^*(s) \in D(s)$, then V is upper Clarke regular at s and*

$$V'(s; x) = \max_{a^*(s) \in A^*(s)} f_s(a^*(s), s; x).$$

Proof The Lipschitz condition of f at every $(a, s) \in \Lambda \times N(s)$ implies that f is continuous in a for each s. The compactness of Λ and $N(s)$ implies that f is globally Lipschitz on $\Lambda \times N(s)$. In particular, there exists some $K > 0$ such that for all $a \in \Lambda$ and all (s', s'') in $N(s)$:

$$\left| f(a, s'') - f(a, s') \right| \leq K \left\| s' - s'' \right\|.$$

The Proposition then follows from Theorem 2.8.2 in Clarke [7]. □

Application This result is used by Askri and LeVan [3] to prove the upper Clarke regularity of the value function V in Eq. (2) above for a large class of classical nonoptimal dynamic growth models. In these models, the mild compactness hypothesis follows from the hypothesis of interiority of solutions, together with the upper hemicontinuity of the optimal solution correspondence (see Askri and LeVan [3] Lemma 3.1), and V thus inherits the upper Clarke regularity of the utility function U.[6]

3.2.2 Constrained Programs

When constraints are active, multipliers (not necessarily unique) become nonzero and stronger conditions on the primitives are needed to prove anything beyond the Lipschitz property for the value function. We next prove the value function of a Lipschitz program with continuously differentiable objective and constraints is upper Clarke regular, provided the SMFCQ holds (and hence, the Clarke gradients are singletons and the multipliers are unique).

[6] Proposition 3.2 in Askri and LeVan is incorrectly stated since its proof requires the (upper) Clarke regularity of U (not the differentiability) so that their Theorem 2.1 can be used to prove the regularity of V.

Proposition 3.3 *Under Criterion 2.1, if the SMFCQ holds at every optimal solution $a^*(s) \in A^*(s)$, and if the primitives are continuously differentiable in s, then V is upper Clarke regular at s and*

$$V^o(s;x) = V'(s;x) = \max_{a^*(s)\in A(s)} \left\{ L_s(a^*(s), s, \lambda, \mu) \cdot x \right\}.$$

Proof Given the uniqueness of the multiplier under the SMFCQ, the Dini bounds in Theorem 3.1 coincide; hence, V is Gâteaux differentiable with:

$$V'(s;x) = \max_{a^*(s)\in A^*(s)} \{L_s(a^*(s), s, \lambda, \mu) \cdot x\}.$$

Clearly, for all $a^*(s) \in A^*(s)$:

$$L_s(a^*(s), s, \lambda, \mu) \cdot x \le \max_{a^*(s)\in A^*(s)} \{L_s(a^*(s), s, \lambda, \mu) \cdot x\};$$

hence,

$$\forall \theta \in \operatorname{clconv} \left\{ \bigcup_{a^*(s)\in A^*(s)} L_s(a^*(s), s, \lambda, \mu) \right\}, \theta \cdot x \\ \le \max_{a^*(s)\in A^*(s)} \{L_s(a^*(s), s, \lambda, \mu) \cdot x\}.$$

Recall that by Theorem 2.1:

$$\partial V(s) \subset \operatorname{cl\,conv} \left\{ \bigcup_{a^*(s)\in A^*(s)} L_s(a^*(s), s, \lambda, \mu) \right\};$$

therefore,

$$V^o(s;x) = \max_{\xi\in\partial V(s)} \{\xi . x\} \le \max_{a^*(s)\in A^*(s)} \{L_s(a^*(s), s, \lambda, \mu) \cdot x\} = V'(s;x),$$

which implies $V'(s;x) = V^o(s;x)$. Hence, V is upper Clarke regular. □

3.3 Concavity and Differentiability

Although lower Clarke regularity is not typically preserved under maximization, the results concerning the preservation of concavity are well known. As any concave function is lower Clarke regular, we combine concavity with upper Clarke regularity of the objective to generate a large class of objectives for which the value function will be *continuously differentiable* under Criterion 3.1. Our result is stated in the following result:

Proposition 3.4 *Under Criterion 3.1, if f is concave in (a, s) and differentiable at s for each a and if $graph D$ is convex, then V is concave and continuously differentiable.*

Proof f is concave in s, and thus lower Clarke regular, in addition to be differentiable in s. It is therefore continuously differentiable in s (see Appendix) and therefore upper Clarke regular. By Proposition 3.1 V inherits the upper Clarke regularity of f, but also the concavity of f since $graph D$ is convex. V is thus continuously differentiable. □

Application This is the well-known result on the differentiability of the value function under interiority of solutions, generally obtained as a consequence of a theorem of Benveniste and Scheinkman (see, for instance, Lucas and Prescott [18], Theorem 4.11).

In the presence of multiple Lagrange multipliers, an alternative to SMFCQ is to postulate enough concavity in the primitives to "squeeze" the lower and upper Dini bounds of Theorem 3.1 to obtain Gâteaux differentiability.

Proposition 3.5 *Under Clarke's hypothesis, if the GMFCQ holds at every $a^*(s) \in A^*(s)$, and if the primitives are continuously differentiable in s, f and $-g$ are concave, and h is affine in a, then V is Gâteaux differentiable and*

$$V'(s; x) = \max_{a^*(s) \in A^*(s)} \min_{(\lambda,\mu) \in K(a^*(s),s)} \left\{ L_s(a^*(s), s, \lambda, \mu) \cdot x \right\}.$$

Proof By Theorem 3.1:

$$\max_{a^*(s) \in A^*(s)} \min_{\lambda \in K(a^*(s),s)} L_s(a^*(s), s, \lambda, \mu) \cdot x \leq D_+ V(s; x).$$

Imposing additional conditions on the primitive helps tighten the upper bound as follows.

First, choose a sequence $\{t_n\} \downarrow 0$ such that:

$$DV^+(s; x) = \limsup = \lim_{n \to \infty} \frac{V(s + t_n x) - V(s)}{t_n}$$

and consider the associated sequence $\{a^*(s + t_n x)\}$ with $a^*(s + t_n x) \in A^*(s + t_n x)$ for all $n \in \mathbb{N}$. Any $a^* \in A^*(s)$ is a global maximum given the concavity assumptions.

By definition of the value function, for any $a^* \in A^*(s)$ and for any $(\lambda, \mu) \in K(a^*, s)$ and $(\lambda_n, \mu_n) \in K(a^*(s + t_n x), s + t_n x)$:

$$\frac{V(s + t_n x) - V(s)}{t_n} = \frac{L(a^*(s + t_n x), s + t_n x, \lambda_n, \mu_n) - L(a^*, s, \lambda, \mu)}{t_n},$$

where $(\lambda_n, \mu_n) \in K(a^*(s + t_n x), s + t_n x)$.

By strong duality, $(a_n^*(s + t_n x), s + t_n x, \lambda_n, \mu_n)$ is a global saddle point of L; therefore,

$$L(a^*(s + t_n x), s + t_n x, \lambda_n, \mu_n) \leq L(a^*(s + t_n x), s + t_n x, \lambda, \mu)$$

and

$$L(a^*(s+t_nx), s, \lambda, \mu) \leq L(a^*, s, \lambda, \mu).$$

Consequently, for any $a^* \in A^*(s)$ and for any $(\lambda, \mu) \in K(a^*, s)$:

$$\begin{aligned}
&\frac{V(s+t_nx)-V(s)}{t_n} \\
&\quad \leq \frac{L(a^*(s+t_nx), s+t_nx, \lambda, \mu) - L(a^*(s+t_nx), s, \lambda, \mu)}{t_n} \\
&\quad \leq L_s(a^*(s+t_nx), s', \lambda, \mu) \cdot x
\end{aligned}$$

for some $s'_n \in [s, s+t_nx]$ by the mean value theorem. By assumption, L_s is upper semicontinuous in (a, s). Hence, for any $a^* \in A^*(s)$ and $(\lambda, \mu) \in K(a^*, s)$:

$$\begin{aligned}
&\lim_{n\to\infty} \frac{V(s+t_nx)-V(s)}{t} \\
&\quad \leq L_s(a^*, s, \lambda, \mu) \cdot x;
\end{aligned}$$

therefore,

$$\begin{aligned}
&D_+V(s; x) \\
&\quad \leq \max_{a^*(s)\in A^*(s)} \min_{(\lambda,\mu)\in K(a^*(s),s)} L_s(a^*(s), s, \lambda, \mu) \cdot x;
\end{aligned}$$

thus, $D_+V = D^+V$, and the result follows. □

Application The above result can be seen as a generalization of some of the findings of Milgrom and Segal [4]. In particular, Proposition 3.3 extends their Theorem 5 to constrained programs satisfying the SMFCQ, while Proposition 3.5 mirrors their Corollary 5, but holds under a weaker CQ.

Finally, in the absence of equality constraints, we note that the hypothesis of joint concavity, together with the continuous differentiability in s, implies if the SMFCQ holds at every optimal solution, then the value function must be at least continuously differentiable. Our result is thus related to the analysis of concave programs of Rincon-Zapatero and Santos [5], although our constraint qualification is weaker than the linear CQ used by these authors.

Proposition 3.6 *Under Criterion 2.1, if the primitives are continuously differentiable in (a, s) and the SMFCQ holds at all $a^*(s)$ in $A^*(s)$, and if f and $-g$ are jointly concave in (a, s) and $h = 0$, then V is continuously differentiable and*

$$V'(s) = L_s(a^*(s), s, \lambda)$$

for any $a^*(s) \in A^*(s)$.

Proof Under the SMFCQ, and with continuously differentiable primitives, the multiplier is unique, and by Proposition 3.5, V is Gâteaux differentiable with:

$$V'(s,x) = \max_{a^*(s)\in A^*(s)} \left\{\left(f_s(a^*(s),s) - \lambda g_s(a^*(s),s)\right)\cdot x\right\}$$

in which $\{\lambda\} = K(a^*(s),s)$. In addition,

$$-V'(s;-x) \leq V'(s;x).$$

The concavity of V (inherited from that of f and $-g$), together with its Gâteaux differentiability, implies it is continuously differentiable, hence lower Clarke regular; therefore,

$$V'(s;x) = V^{-o}(s;x) \leq V^{o}(s;x) = -V^{-o}(s;-x) = -V'(s;x).$$

As a result,

$$V^{o}(s;x) = V^{-o}(s;x);$$

that is,

$$\begin{aligned}&\max_{a^*(s)\in A^*(s)} \left\{\left(f_s(a^*(s),s) - \lambda g_s(a^*(s),s)\right)\cdot x\right\}\\ &\quad= \min_{a^*(s)\in A^*(s)} \left\{\left(f_s(a^*(s),s) - \lambda g_s(a^*(s),s)\right)\cdot x\right\};\end{aligned}$$

hence, V is continuously differentiable at s and

$$V'(s) = f_s(a^*(s),s) - \lambda g_s(a^*(s),s)$$

for any $a^*(s)\in A^*(s)$ and $\{\lambda\} = K(a^*(s),s)$. □

We note that Proposition 3.6 does not require the set of optimal solutions to be a singleton. Further, any optimal solution along with its associated unique multiplier can be used to calculate the gradient of the value function.

4 Applications and Extensions

In this section, we show how the general results derived above can be applied to various economic models characterized by nonconvexities and/or nonsmooth objectives and constraints.

4.1 Optimization Problems with Discrete Choice Variables

We first illustrate the use of our results in a simple model with nonsmooth primitives while also allowing for discrete choices. To apply our results, both objective and constraints are only Lipschitz and discrete choice constraints are rewritten as equality constraints. The upside of this strategy is that equality constraints thus become continuously differentiable; the downside is that they restrict the choice of directions satisfying the CQ (since such direction must be in the set of tangent directions of equality constraints).

Our results can be applied to a large class of models, but for clarity we focus on a simple utility maximization problem in which a consumer, endowed with one unit of time and $e > 0$ units of the consumption good, chooses jointly a level of consumption c and whether to work or not ($l = 1$ or $l = 0$) and so as to maximize utility $U(c, l)$. Function $U : \mathbb{R}_+ \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ is assumed to be continuous, increasing, and Lipschitz at every point in the interior of its domain, and $U(c, 1) > U(c, 0)$ for all $c \geq 0$. The consumption good is produced by a firm with production function $F : \mathbb{R} \longrightarrow \mathbb{R}$ assumed to be increasing, continuous, and Lipschitz at every point.

Associated with the consumer's problem is the program:

$$V(e) = \max U(c, l)$$

subject to:

$$\begin{aligned} h(c, l) &= l(1 - l) = 0 \\ g_1(e, (c, l)) &= c - F(1 - l) - e \leq 0 \\ g_2(e, (c, l) &= -c \leq 0. \end{aligned}$$

The multiplier rule characterizing KKT points is:

$$0 \in \partial_{(c,l)}\{U - \lambda_1 g_1 - \lambda_2 g_2 - \mu h\}(e, (c^*, l^*))$$

so that:

$$\begin{aligned} &\lambda_1 - \lambda_2 \in \partial_c U(c^*, l^*) \\ &\quad \mu\left(1 - 2l^*\right) \in \partial_l U(c^*, l^*) - \lambda_1 \partial_l F(1 - l^*), \end{aligned}$$

together with the complementary slackness conditions:

$$\begin{aligned} \lambda_1(c^* - F(1 - l^*) - e) &= 0 \\ \lambda_2 c^* &= 0. \end{aligned}$$

First, we note that Clarke's hypothesis is satisfied because the choice domain $D(e)$ (a closed set) is included in the compact set $[0, F(1) + e] \times [0, 1]$; hence, the uniform compactness condition is trivially satisfied. Second, it is easy to verify that the GMFCQ is satisfied at any solution (c^*, l^*). Indeed as:

$$\begin{aligned} \partial_{(c,l)} h &= \{(0, 1 - 2l^*)\} \\ \partial_{(c,l)} g_1 &= \{1\} \times -\partial_l F(1 - l^*) \\ \partial_{(c,l)} g_2 &= \{(-1, 0)\} \end{aligned}$$

if $c^* = 0$ (and thus $\lambda_1 = 0$ and $l^* = 1$) then any $(y_1, 0)$ with $y_1 > 0$ satisfies the GMFCQ, while if $c^* > 0$ (hence $\lambda_2 = 0$) any $(y_1, 0)$ with $y_1 < 0$ satisfies the GMFCQ. The following result is then a direct consequence of Theorem 3.1.

Proposition 4.1 *V is Lipschitz at any $e > 0$ and*

$$D^+ V(e; x) \le \max_{(c^*, l^*) \in A^*(e)} \sup_{\lambda_1 \in K((c^*, l^*), e)} \lambda_1$$

and

$$D_+ V(e; x) \ge \max_{(c^*, l^*) \in A^*(e)} \inf_{\lambda_1 \in K((c^*, l^*), e)} \lambda_1 \ge 0,$$

where $\lambda_1 = 0$ if $c^* = 0$ and $\lambda_1 \in \partial_c U(c^*, l^*)$ if $c^* > 0$.

We note that the SMFCQ can only be satisfied if none of the inequality constraints are active, i.e., if $0 < c^* < F(1 - l^*) + e$. When that is the case, Proposition 3.3 implies that V is Gâteaux differentiable at any $e > 0$ since the primitives are trivially continuously differentiable in e.

Alternatively, if U is continuously differentiable in its first argument, then the multiplier is unique since $\lambda_1 = U_c(c^*, l^*) \ge 0$ if $c^* > 0$, and V is Gâteaux differentiable with:

$$V'(e; x) = \max_{(c^*, l^*) \in A^*(e)} U_c(c^*, l^*) \text{ if } c^* > 0.$$

4.2 Lipschitz Dynamic Programming

The quantitative property of Lipschitz continuity of the value function in optimization programs is very important, at minimum for computational reasons, so a number of researchers have produced sufficient condition for the preservation of the Lipschitz property in maximization programs (see, for instance, Laraki and Sudderth [19] and Hinderer [20]) expressed generally in terms of global Lipschitz conditions. We seek here to weaken these conditions to just (local) Lipschitzness for recursive dynamic programs common in the economic literature.

Consider the following generic recursive dynamic program (see, for instance, [18]):

$$V_{k+1}(s) = T(V_k)(s) = \max_{a \in D(s)} \{f(a, s) + \beta V_k(a)\}$$

in which $D(s) = \{a \in A \subset \mathbb{R}^n,\ g(a,s) \leq 0\}$, $s \in S \subset \mathbb{R}^m$ and $V_0 = 0$, with its corresponding Lagrangian:

$$L_{k+1}(a,s) = f(a,s) + \beta V_k(a) - \lambda g(a,s)$$

and solution set $A^*_{k+1}(s) = \arg\max_{a\in D(s)} \{f(a,s) + \beta V_k(a)\}$. Both functions f and g are only assumed to be Lipschitz in (a,s), and $0 < \beta < 1$.

Given $V_0 = 0$, the following result is a direct consequence of a repeated application of Theorem 2.1.

Proposition 4.2 *If the GMFCQ is satisfied for all $a^*_{k+1}(s) \in A^*_{k+1}(s)$ for each n, then if Clarke's hypothesis is satisfied (or under any of the conditions in Proposition 2.1) the sequence $\{V_k\}$ is a sequence of Lipschitz functions, with Clarke gradients satisfying:*

$$\partial V_{k+1}(s) \subset cl\ conv \left\{ \bigcup_{a^*_{k+1}(s)\in A^*_{k+1}(s)} \ \bigcup_{\lambda\in K^*(a^*(s),s)} \partial_s (f - \lambda g)(a^*_{k+1}(s), s) \right\}.$$

Of course, whether or not the GMFCQ and Clarke's hypothesis are satisfied will depend on the specific problem considered. The GMFCQ is easier to satisfy, the fewer the number of active constraints (it is trivially satisfied when all constraints are inactive), while Clarke's hypothesis is automatically satisfied if the choice domain is uniformly bounded (as in bounded growth models).

It is well known that the sequence $\{V_k\}$ of Lipschitz functions converges uniformly to the unique continuous function V satisfying $V = T(V)$. Unfortunately, uniform limits of sequences of locally Lipschitz functions are not necessarily locally Lipschitz,[7] but it is nevertheless possible to prove that V is (at least) Lipschitz under variety of cases.

One can, for instance, impose a global or uniform Lipschitz condition, as in Laraki and Sudderth [19] or Hinderer [20] , in which case an upper bound for the Lipschitz modulus of the value function can be derived. Alternatively, we illustrate the use of our results in the growth model with nonconvex technology of Askri and LeVan [3] in which the objective is only assumed to be upper Clarke regular and solutions are interior.

4.2.1 Clarke Regular Dynamic Programming

Consider the growth model of Askri and LeVan [3] and the associated Bellman equation:

$$V(k_0) = \max_{k_1\in G(k_0)} \{U(k_0, k_1) + \beta V(k_1)\}.$$

[7] The Weistrass approximation theorem asserts that any continuous functions, Lipschitz or not, may be uniformly approximated by polynomials (which are Lipschitz).

While the assumptions made in Askri and LeVan are sufficient to ensure that Criterion 3.1 is fulfilled (as shown in their Lemma 3.1), we amend their results (Propositions 3.2, 3.3) to emphasize the importance of upper Clarke regularity.

First, we correct their Proposition 3.2 by noting that the critical assumption on U is that of upper Clarke regularity in its first argument. Indeed, a careful reading of their proof reveals that the upper Clarke regularity of V relies primarily on the upper Clarke regularity of U (and not, simply, its differentiability[8]), which is obtained through their Theorem 2.1.

Second, we note that the addition of upper Clarke regularity of U in its second argument is sufficient to prove their Proposition 3.3, namely that V is not just differentiable, but *once continuously differentiable* at every point on the optimal path. In this case, even though the primitives are not smooth, V has that property on the optimal path. A crucial feature in this problem is the interiority of all optimal solutions, a property guaranteed by assuming Criterion 3.1 just like in Askri and LeVan [3].

Proposition 4.3 *If Criterion 3.1 holds and if U is upper Clarke regular in its first argument, then V is upper Clarke regular at $k_0 > 0$. Furthermore, if U is also upper Clarke regular in its second argument, then V is continuously differentiable at every point along the optimal path.*

Proof The upper Clarke regularity of V follows from Proposition 3.2 .

Since all solutions are interior, and given that the directional derivatives of the sum of two Clarke regular functions are additive by Clarke [7] (Theorem 2.9.8), the first-order conditions for any directions d and $-d$ are:

$$U_2'(k_0, k_1; d) + \beta V'(k_1; d) \leq 0$$
$$U_2'(k_0, k_1; -d) + \beta V'(k_1; -d) \leq 0.$$

Upper Clarke regularity implies that:

$$U_2^0(k_0, k_1; d) + \beta V^0(k_1; d) \leq 0$$
$$U_2^0(k_0, k_1; -d) + \beta V^0(k_1; -d) \leq 0;$$

hence, for $\nu \in \partial U_2(k_0, k_1), \xi \in \partial V(k_1)$,

$$\max\{\nu \cdot d\} + \beta \max\{\xi \cdot d\} \leq 0$$
$$\max\{\nu \cdot (-d)\} + \beta \max\{\xi \cdot (-d)\} \leq 0.$$

As a result, the last inequality simplifies to

$$\min\{\nu \cdot d\} + \beta \min\{\xi \cdot d\} \geq 0;$$

thus, necessarily, for all $d \in \mathbb{R}^n$

$$\min\{\nu \cdot d\} = \max\{\nu \cdot d\}, \quad \min\{\xi \cdot d\} = \max\{\xi \cdot d\}.$$

[8] Differentiable functions need not be upper Clarke regular.

This implies that the Clarke gradients $\partial_2 U(k_0, k_1)$ and $\partial V(k_1)$ are both singleton for any optimal k_1, given k_0 ; hence, V is once continuously differentiable at any optimal point k_1. □

4.2.2 Concave Dynamic Programming

Much more (than just the Lipschitz property) can be established concerning the differentiability properties of the value function in the presence of concavity, providing the primitives are also assumed to be differentiable with respect to s. Concavity itself, a feature of many economic models, is an important property since it is preserved under pointwise limits and since concave functions are Lipschitz on the interior of their domain.

The presence of multiple multipliers is of course a hindrance, but that too can be set aside if one assumes that the SMFCQ holds.

Proposition 4.4 *Assuming (i) f and g are Lipschitz and concave in (a, s) as well as continuously differentiable in s, (ii) the derivatives f_s and g_s are upper semicontinuous in a, (iii) the GMFCQ is satisfied at every optimal solution, and (iv) Criterion 2.1 is satisfied, then the Lipschitz value function V is concave and Gâteaux differentiable with:*

$$V'(s, x) = \max_{a^*(s)\in A^*(s)} \min_{\lambda\in K(a^*(s),s)} \big(F_s(a^*(s), s) - \lambda g_s(a^*(s), s)\big) \cdot x.$$

If, in addition, the SMFCQ is satisfied at every optimal solution, then V is continuously differentiable and

$$V'(s) = f_s(a^*(s), s) - \lambda_{a^*(s)} g_s(a^*(s), s)$$

for any $a^*(s) \in A^*(s)$ and $\lambda_{a^*(s)} = K(a^*(s), s)$.

Proof V is concave, hence Lipschitz, and satisfies the Lipschitz program:

$$V(s) = \max_{a\in D(s)} \{f(a, s) + \beta V(a)\}.$$

Under the MFCQ a direct application of Proposition 3.5 to this program implies that V is Gâteaux differentiable with:

$$V'(s, x) = \max_{a^*(s)\in A^*(s)} \min_{\lambda\in K(a^*(s),s)} \big(f_s(a^*(s), s) - \lambda g_s(a^*(s), s)\big) \cdot x,$$

but the existence of multiple Lagrange multipliers generically prevents V from being continuously differentiable.

However, assuming next that the SMFCQ is satisfied at each optimal solution, the multiplier set is a singleton and by Proposition 3.6:

$$V'(s) = f_s(a^*(s), s) - \lambda g_s(a^*(s), s)$$

for any $a^*(s) \in A^*(s)$ and $\lambda \in K(a^*(s), s)$. □

Our result on the continuous differentiability generalizes Benveniste and Scheinkman [17] by allowing the inequality constraints to be active at the optimal solution. The cost is a stronger constraint qualification (SMFCQ), although weaker than the LICQ in Rincon-Zapatero and Santos [5].

4.3 Supermodularity of the Value Function in Recursive Dynamic Programs

The global convergence of monotone Markov processes is at the core of many results concerning the long-run behavior of many recursive economic models or dynamical systems. In stochastic dynamic optimization, the existence of monotone controls pretty much ensures that the state follows a monotone Markov process, as clearly illustrated in standard capital accumulation models (see, for instance, Hopenhayn and Prescott [21] or Stokey et al. [18]).

In these models, as in many recursive dynamic economic models, monotone controls are a consequence of the supermodularity of the value function. In Hopenhayn and Prescott [21], this supermodularity follows from the application of a powerful theorem of lattice theory on the preservation of supermodularity under maximization (Theorem 2.7.6 in Topkis [22]), but is restricted to programs with sublattice-valued choice domains. In the context of the standard classical growth model, this means that practically all but Leontieff-type production functions are excluded. Alternatively, in Mirman, Morand and Reffett [23], the proof of supermodularity relies on the differentiability properties of the value function and requires the primitives to be smooth.

We propose to use a generalized envelope result to prove supermodularity without relying on Topkis' theorem and without assuming smooth primitives. To illustrate our argument, we work with a deterministic dynamic growth model with nonconvexities and prove the supermodularity of the value function under fairly general conditions.

4.3.1 Model and Definition of Recursive Equilibrium

We consider a class of models with a continuum of identical infinitely lived households/firms. Each household enters period $t = \{0, 1, 2, \dots\}$ with an individual stock of capital k_t, and supplies inelastically one unit of time to firms. Common in the literature (e.g., Coleman [25], Greenwood and Huffman [26]), and consistent with more recent work (e.g., Tanaka [24] Kamihigashi and Roy [27,28]), we adopt a "reduced-form" production function $F(k, n, K, N)$, where k and n are, respectively, the firm's capital and labor inputs, while K and N are the average or per capita corresponding quantities. Since $n = N = 1$ we use the notation $f(k, K) = F(k, 1, K, 1)$ and make the following standard assumptions.

Assumption (i) *There exists* $\hat{k} > 0$ *such that* $F(\hat{k}, 1, \hat{k}, 1) = \hat{k}$ *and* $F(k, 1, k, 1) < k$ *, for all* $k > \hat{k}$, so we denote by $\mathbb{K}$ *the interval* $[0, \widehat{k}]$. Function $F : \mathbb{K} \times [0, 1] \times \mathbb{K} \times [0, 1] \to \mathbb{R}$ *is continuous, increasing, concave in its first two arguments and exhibits constant returns to scale in* (k, n).

Assumption (*ii*) $u : \mathbb{K} \mapsto \mathbb{R}$ *is increasing, continuous, and concave and satisfies* $u(0) = 0$.

We replace the usual Inada condition $\lim_{c\to 0} u'(c) \to \infty$ by the following assumption:

Assumption (*iii*) *For all* $M > 0$, *there exists* $x_0 \in \mathbb{K}, x_0 > 0$ *such that* $\xi > M$ *for all* $\xi \in \partial u(x_0)$.

As in Hopenhayn and Prescott, we also need a curvature condition requiring that the degree of complementarity between private and aggregate per capita capital stocks be high relative to the curvature of the utility function.

Assumption (*iv*) $\forall k' \geq k, \forall K' \geq K$ and $\forall y \in [0, f(k, K)]$:

$$u(f(k', K') - y) - u(f(k, K') - y) \geq u(f(k', K) - y) - u(f(k, K) - y).$$

Note that the function $u(f(k, .) - y)$ is Lipschitz at any $K > 0$ satisfying $f(k, K) - y > 0$; hence, Assumption (iv) will be satisfied if, at points where $u(f(k, .) - y)$ is differentiable, $u'(f(k, K) - y) f_1(k, K)$ is increasing in K.

Given $k_0 = K_0 > 0$, a consumer seeks to maximize:

$$E_0 \left\{ \sum_{i=0}^{\infty} \beta^i u(c_i) \right\}$$

subject to:

$$c_t + k_{t+1} \leq f(k_t, K_t) \quad t = 0, 1, \ldots.$$

In addition, the consumer uses the law of motion h to recursively compute the sequence $\{K_t\}$ of per capita capital stocks as $K_{t+1} = h(K_t)$.

A recursive equilibrium is a particular law of motion $h^* \in \boldsymbol{B}$ such that for all $k \in \mathbb{K}$:

$$h^*(k) \in Y^*(k, K; h) = \arg \sup_{y \in \Gamma(k,K)} \{u(f(k, K) - y) + V(y, h(K))\}$$

and where

$$\boldsymbol{B} = \{h : \mathbb{K} \to \mathbb{K}, 0 \leq h(k) \leq f(k, k), h \text{ usc and increasing}\}.$$

Exploiting the lattice structure generated by the pointwise partial order, we denote by $\vee Y^*$ and $\wedge Y^*$ the greatest and least selections of Y^*. Function V is the unique value function satisfying Bellman's equation:

$$V(k, K) = \sup_{y \in \Gamma(k,K)} \{u(f(k, K) - y) + V(y, h(K))\},$$

where $\Gamma(k, K) = \{y \in \mathbb{K}, 0 \leq y \leq f(k, K)\}$.

The proof of existence of equilibrium follows the argument in Mirman, Morand and Reffett [32] (Theorems 6 and 9) which relies on the order-preserving properties of a self-mapping on $\boldsymbol{B}$ defined by:

$$h \to \vee Y^*(k, k; h).$$

It requires V to be supermodular, a property which we establish next.

4.3.2 Proof of Supermodularity of the Value Function

Given $V_0 = 0$ we prove by induction that each element of the sequence $\{V_n = T^{(n)} V_0\}$ is supermodular; V naturally inherits that property as the pointwise limit of that sequence.

Fixing $h \in \boldsymbol{B}$, and assuming that V_n is Lipschitz and supermodular, we consider the Lipschitz program:

$$V_{n+1}(k, K) = \max_{0 \le y \le f(k,K)} \{u(f(k, K) - y) + V_n(y, h(K))\}.$$

Given that h and f are increasing and that u is concave, the objective above has increasing differences in $(y; (k, K))$.[9] Since the choice correspondence $[0, f(k, K)]$ is strong set order ascending, the greatest and least selections $\vee Y^*_{n+1}$ and $\wedge Y^*_{n+1}$ are both increasing in (k, K) by Theorems 2.8.1 and 2.8.3 in Topkis [22]. Note that by the same argument both $f - \vee Y^*_{n+1}$ and $f - \wedge Y^*_{n+1}$ are also increasing in (k, K).

Inada conditions imply interiority of solutions (all multipliers are thus 0); hence, by Theorem 3.1, V_{n+1} is Lipschitz with:

$$\max_{y^*(k,K)\in Y^*_{n+1}(k,K)} \left(\min_{\theta \in \partial(u(f(k,K)-y^*(k,K)))} \theta \cdot x \right) \le D_+ V_{n+1}(k, K; x)$$

and

$$D^+ V_{n+1}(k, K; x) \le \max_{y^*(k,K)\in Y^*_{n+1}(k,K)} \left(\max_{\theta \in \partial(u(f(k,K)-y^*(k,K))} \theta \cdot x \right)$$

in which the Dini derivatives are with respect to the first variable.

The concavity of u implies that if $c' > c$ then $\forall (\theta, \theta') \in \partial u(c) \times \partial u(c')$ necessarily $0 \le \theta' \le \theta$. As a result, for any $x > 0$:

$$\max_{y^*(k,K)\in Y^*_{n+1}(k,K)} \left(\max_{\theta \in \partial(u(f(k,K)-y^*(k,K))} \theta \cdot x \right) \le \max_{\theta \in \partial(u(f(k,K)-\wedge Y^*_{n+1}(k,K))} \theta \cdot x$$

[9] Supermodularity and increasing differences are equivalent properties on $\mathbb{R}^2$.

and also, given any $\overline{k} \in \mathbb{K}$ with $\overline{k} > 0$, $\forall k \in [\widehat{k}, \overline{k}]$:

$$\max_{\theta \in \partial(u(f(k,K) - \wedge Y^*_{n+1}(k,K))} \theta \cdot x \leq \max_{\theta \in \partial(u(f(\overline{k},K) - \wedge Y^*_{n+1}(\overline{k},K))} \theta \cdot x.$$

Thus, $\forall k \in [\widehat{k}, \overline{k}]$ and $\forall x > 0$:

$$\begin{aligned} 0 &\leq D_+ V_{n+1}(k, K; x) \\ &\leq D^+ V_{n+1}(k, K; x) \leq \max_{\theta \in \partial(u(f(\overline{k},K) - \wedge Y^*_{n+1}(\overline{k},K))} \theta \cdot x, \end{aligned}$$

which proves that the Dini derivatives of V_{n+1} are uniformly bounded on any interval $[\overline{k}, \widehat{k}]$ (a symmetric argument holds for the direction $x < 0$). As a result, the function $k \to V_{n+1}(k, K)$ is absolutely continuous on $[\overline{k}, \widehat{k}]$ for any $0 < \overline{k} < \widehat{k}$.

Such absolute continuity, together with the properties that V_{n+1} is increasing and continuous in its first argument, implies that V_{n+1} is absolutely continuous on $\mathbb{K} = [0, \widehat{k}]$ (see Problem 37 in Royden [29]). By the fundamental theorem of integral calculus (Royden [29]), $k \to V_{n+1}(k, K)$ is therefore almost everywhere differentiable and, $\forall k \in \mathbb{K}$:

$$V_{n+1}(k, K) = \int_0^k V'_{n+1}(s, K) ds. \tag{3}$$

At the points where $k \to V_{n+1}(k, K)$ is differentiable, by definition both Dini's must coincide; hence, for all s:

$$V'_{n+1}(s, K) = u'(f(s, K) - \wedge Y^*_{n+1}(s, K)) f_1(s, K).$$

Finally, we note that for any $K' > K$, wherever the derivative exists:

$$\begin{aligned} V'_{n+1}(s, K) &= u'(f(s, K) - \wedge Y^*_{n+1}(s, K)) f_1(s, K) \\ &\leq u'(f(s, K') - \wedge Y^*_{n+1}(s, K')) f_1(s, K') = V'_{n+1}(s, K') \end{aligned}$$

the inequality following from Assumption (iv) and the monotonicity of $\wedge Y^*_{n+1}$.

The above inequality, together with (3) prove the supermodularity of V_{n+1}. By induction, the sequence $\{V_n\}_{n=0}^{\infty}$ is a collection of supermodular functions in (k, K) and its pointwise limit V inherits that property.

We note that the strategy of relying on a nonsmooth envelope can also be used to derive sufficient conditions for the value function of a stochastic one sector growth model with Lipschitz primitives to be supermodular and therefore to prove the existence of monotone controls. Based on this monotonicity, strong results about the convergence of the state to an invariant distribution can then be derived.

4.4 Normal Goods in the Consumer's Problem via Minimax Dynamic Lattice Programming

In our final application, we construct monotone comparative statics in a consumer's problem via aggregation. The construct is essentially a dynamic programming argument. That is, our strategy corresponds to solving the consumer's problem in two stages. In the first stage, given a (feasible) level of consumption for some good i, the consumer maximizes utility over choices of all other goods. Then, in the second stage, the consumer then chooses a level of consumption for good i.

We study the normality/comparative statics of the second-stage decision relative to choosing good i. This aggregation procedure enables us apply Proposition 3.5 under criterion 2.1 to show how envelope theorems can be used to prove the existence of monotone controls in an important example of a constrained lattice programming problem (namely, provide sufficient conditions for the normality of consumption demand for a single good in a vector of consumption decisions).

It is well known that standard lattice programming techniques in Euclidean orders cannot be applied to the consumer's problem to derive conditions for the normality of demand for a single good. (e.g., Antoniadou[30,31], Mirman and Ruble [32], Quah [33]). For example, ordered changes in income do not induce strong set ordered changes in the budget correspondence in the standard Euclidean lattices.[10] To deal with this situation, researchers have changed the order structures on commodity spaces. In some cases, these new partial orders induce lattice structure on commodity spaces (e.g., Antoniadou [30,31] and Mirman and Ruble [32]). In other cases, the new orders are not even generally partial orders, let alone lattices, but they are sufficient to infer the existence of monotone comparative statics (e.g., the flexible set orders in Quah [33]). Unfortunately, in all this work, the characterizations of monotone selections are "weak" when compared to the strong set order comparative statics in standard lattice programming problems, where objectives are supermodular in consumption and have increasing differences with parameters such as income. In this section, we combine our nonsmooth envelope theorems with duality to construct *strong set order comparative statics* for the consumer's problem in income. We refer to the resulting lattice programming approach as *minimax lattice programming*.

To formalize the consumer's problem, let $A = C \subset \mathbb{E}^n_+$ be the (finite-dimensional) commodity space,[11] C a nonempty sublattice, and $s \in \mathbb{S} = [0, \infty[$ the level of income. Fix the vector of prices to be $p \in \mathbb{S}_{n-1}$, with $p >> 0$, and denote by $\mathbb{S}_{n-1}$ the $n-1$-dimensional simplex. Define $C_i = \{c_i \in \mathbb{E}_+, c = (c_1, c_2, \ldots, c_{i-1}, c_i, c_{i+1}, \ldots, c_n) \in C\}$ and $C_{-i} = \{c_{-i} = (c_1, c_2, \ldots, c_{i-1}, c_{i+1}, \ldots, c_n) \in \mathbb{E}^{n-1}, c \in (c_1, c_2, \ldots, c_{i-1}, c_i, c_{i+1}, \ldots, c_n)\}$. As we seek strong set order monotone comparative statics, we will consider comparative statics in the following Veinott sublattice power domains: $\mathcal{L}(E_i) = \{L_i \subset \mathbb{E}_+ | L_i$ a

[10] A finite-dimensional Euclidean lattice $\mathbb{E}^n$ is the pair $(\mathbb{R}^n, \geq_e)$, where $\geq_e$ is the standard componentwise (product) Euclidean order.

[11] The extension to infinite-dimensional commodity spaces is possible using similar duality arguments in this section.

nonempty sublattice of $\mathbb{E}_+$), and $\mathcal{L}(\mathbb{E}_{-i}) = \{L_{-i} | L_{-i}$ a nonempty sublattice of $\mathbb{E}_+^{n-1}\}$, each set endowed with Veinott's strong set order.

We assume that the consumer's preferences can be represented by a real-valued utility function $u : C \to \mathbb{R}$, where $u(c)$ satisfies the cardinal complementarity conditions in Quah [33]:

Assumption Q: ($i-$concave supermodularity). Preferences are represented by a utility function $u(c)$ that is (a) supermodular and bounded below on $C \subset \mathbb{E}_+^n$,[12] (b) concave in c_{-i} for each c_i, and (c) increasing and locally Lipschitz in c_i.

Notice, in Assumption Q, we do *not* require good i to be an element of a *convex* set; we only require the objective function in good i to be increasing and locally Lipschitz in its ith argument. So our results are *not* contained in Quah [33].

We ask the following question: For each fixed price vector $p >> 0$, what are sufficient conditions for the consumer's demand correspondence for good c_i to be strong set order isotone in income $s \in \mathbb{S}$ to $\mathcal{L}(C_i)$? To answer this question about monotone comparative statics, we first build "cardinal" conditions based on Assumption Q and then take increasing transformations to obtain ordinal versions (based on quasi-concavity and quasi-supermodularity).

The consumer's problem can be described as follows. For a fixed price vector $p = (p_1, p_2, \ldots, p_{i-1}, 1, p_{i+1}, \ldots, p_n) >> 0$, $p \in \mathbb{S}_{n-1}$, and an income level $s \in \mathbb{S}$, the consumer solves [13]:

$$V(s) = \max_{c \in D(s)} u(c), \tag{4}$$

where $V(s)$ is the value function at s. The budget correspondence is:

$$D(s) = \{c : p \cdot c \leq s, c \in C\}$$

and the set of optimal solutions is:

$$C^*(s) = \arg \max_{c \in D(s)} u(c). \tag{5}$$

We shall refer to Program (4) as the *primal problem,* in which Assumption Q is well defined and attains its maximum.

For any $s > 0$ and $c_i \in in[0, s[$, define the "*aggregated primal*" as:

$$V^{ap}(c_i, s) = \max_{c_{-i} \in D(c_i, s)} u(c_{-i}; c_i) \tag{6}$$

[12] If the utility function is not bounded below, the argument in this section can be modified to accommodate this case

[13] As our arguments must hold for each fixed p, to economize on notation, we suppress the p in the notation for parameters except where emphasis is needed, and/or the context is not clear.

with optimal solutions given by:

$$C_{-i}^{ap*}(c_i, s) = \arg \max_{c_{-i} \in D(c_i, s)} u(c_{-i}, c_i).$$

We can then recast our original monotone comparative statics problem as the existence of a single-crossing property (or increasing differences) in the following problem:

$$V(s) = \max_{c_i \in [0,s]} V^{ap}(c_i, s), \tag{7}$$

where $V(s)$ is the value function of the primal problem (4). Notice Program (7) is a *standard* (supermodular) lattice programming problem if for each $p \in \mathbb{S}_{n-1}$ with $p >> 0$, the value function $V^{ap}(c_i, s)$ is supermodular in (c_i, s), as the feasible correspondence in program (7) is Veinott strong set order isotone in s.

To prove the supermodularity of the value function V^{ap} in (6), we first note the feasible correspondence is generated by the single constraint $g(c; p) = p \cdot c$ intersected with the condition that $c \in C$, where the constraint is a smooth, convex, increasing valuation (i.e., both supermodular and submodular). So the feasible correspondence $D(c_i, s) = \{c_{-i} \in C_{-i} : c_i \in [0, s[. \ p \cdot c \leq s\} \subset \mathbb{R}_+^{n-1}$ has a convex graph in s, each $c_i \in [0, s[$.[14] Since the objective function in the aggregated primal is concave (although not necessarily smooth) in the choice variables c_{-i} for each $c_i \in [0, s]$, the value function $V^{ap}(c_i, s)$ is a concave function of s for each $c_i \in [0, s]$. Therefore, for each $c_i \in]0, s[$, $s > 0$, it has a nonempty partial subdifferential in s. Further, by Proposition 3.5, this value function is Gateaux differentiable. This implies that our monotone comparative statics problem has been reduced to characterizing a (nonsmooth) envelope using Proposition 3.5 for the aggregated Problem (6).

To characterize this nonsmooth envelope, we first conjugate the aggregated primal with the classical Lagrangian dual. Under Assumption Q, as the aggregated primal in (6) is a standard concave programming problem in c_{-i} for each $c_i \in [0, s[$ and $s > 0$, we can define:

$$\begin{aligned} L(c_{-i}, \lambda; c_i, s) &= u(c_{-i}; c_i) - \lambda(p \cdot c - s) \text{ if } c_{-i} \in C_{-i}, \phi \in \Omega \\ &= +\infty \text{ if } c_{-i} \in C_{-i}, \phi \notin \Omega \\ &= -\infty \text{ if } c_{-i} \notin C_{-i} \qquad (8) \\ &= \text{ any number in } \mathbb{R}^* \text{ else.} \qquad (9) \end{aligned}$$

Using this Lagrangian, we define the "aggregated dual" problem as follows for $c_i \in [0, s)$, $s > 0$:

$$V^{ad}(c_i, s) = \inf_{\lambda} \sup_{c_{-i}} L(c_{-i}, \lambda; c_i, s), \tag{10}$$

[14] Notice the interesting cases here of $D(c_i, s)$ occur when $c_i \in [0, s[$, with $s > 0$. When $c_i = s$, as utility is bounded below, the value function can be trivially defined at $c_i^* = s$, and $c_{-i}^* = 0$.

where $\lambda \in \Lambda \subset \mathbb{R}_+$, $c_{-i} \in C_{-i}$, so V^{ad} is the value function for the aggregated dual. By a standard strong duality argument, we know this problem has zero duality gap (i.e., $V^{ap}(c_i, s) = V^{ad}(c_i, s)$), necessary and sufficient standard Lagrange multiplier rules apply, and we have saddle point stability for optimal solutions, among other properties (see Rockafellar [34]).

The objective function in the Lagrangian dual is:

$$V^L(\lambda, c_i, s) = \sup_{c_{-i}} L(c_{-i}; \lambda, c_i, s), \tag{11}$$

and the corresponding set of optimal solution is:

$$C_{-i}^{L*}(\lambda, c_i, s) = \arg \sup_{c_{-i} \in C_{-i}} L(c_{-i}; \lambda, c_i, s).$$

Using standard lattice programming techniques, we can establish the following properties of the correspondence $C_{-i}^{L*}(\lambda, c_i, s)$ for any $p >> 0$:

Lemma 4.1 *Under Assumption Q,* $C_{-i}^{L*} : \Omega \times C_i \times \mathbb{S} \times \mathbb{S}_{n-1} \to \mathcal{L}(\mathcal{C}_i)$ *is* (i) *descending in* $\lambda \in \Omega$ *(with all selections antitone) for each* $(c_i, s) \in C_i \times \mathbb{S}$, *and* (ii) *ascending in* c_i *(with all selections isotone) for each* $(\lambda, s) \in \Omega \times \mathbb{S}$.

Proof Since $C_{-i} \subset \mathbb{E}^{n-1}$ is a section of C, C_{-i} is a sublattice (e.g., Topkis ([22], p. 16). Further, for each $(\lambda, c_i, s) \in \Omega \times C_i \times \mathbb{S}$, noting the presence of an indicator function for the sublattice C_{-i} in the definition of the Lagrangian in (11), we first consider the parameterized Lagrangian $L_{(\lambda,c_i,s)}$: $\mathbb{E}^{n-1} \mapsto \mathbb{R}_-^* = \mathbb{R} \cup -\infty$. As (i) $(\mathbb{R}_-^*, *)$ is a posemigroup (with identity element) for $* = +$, and (ii) the indicator function for any sublattice $C_i^{'} \subset C_i$ super$*$ (e.g., Veinott [35], Chapter 6, p. 18) with $*$=+ properly increasing in this posemigroup, under assumption Q, the Lagrangian is super$*$ in c_i for $* = +$ (and, hence, $L_{(\lambda,c_i,s)}(c_{-i})$ supermodular in c_{-i}, for each (λ, c_i, s)). Further, $L_{(\lambda,c_i,s)}(c_{-i})$ has strict decreasing differences between $(c_{-i}; \lambda)$ for each (c_i, s), and increasing differences between $(c_{-i}; c_i)$ each (λ, s) (and is a valuation in $(c_i; s)$ for each (λ, c_i)). Therefore, by Topkis' theorem (Topkis [22], Theorem 2.8.3), the correspondence $C_{-i}^{L*}(\lambda, c_i, s) = C_{-i}^{L*}(\lambda, c_i)$ is a Veinott strong set order descending in λ, each c_i, and ascending in c_i, each λ. We note that the strict decreasing (respectively, increasing) differences imply that every selection of $C_{-i}^{L*}(\lambda, c_i)$ is antitone in λ, for each c_i (respectively, isotone in c_i, for each λ) and independent of s. This proves the result. □

By a standard argument, the value function in $V^L(\lambda, c_i, s)$ is convex in λ for each (c_i, s). It is also submodular (noting that for the posemigroup $(\mathbb{R}_+^*, *)$ with $\mathbb{R}_+^* = \mathbb{R} \cup \infty$ and $* = +$, indicator functions of sublattices in the space are sub$*$ with $*$ properly increasing). Also, the effective domain of V^L is a convex sublattice. By an application of a standard version of Danskin's theorem (e.g., Grinold [36], Lemma, p. 186), the right directional envelope $V^L(\lambda, c_i, s)$ in λ in direction $e^+ > 0$ is given by:

$$V_\lambda^L(\lambda, c_i; e^+) = \min_{c_{-i} \in C_{-i}^{L*}} (s - p_{-i} c_{-i}(\lambda, c_i, s) - c_i) \cdot e^+,$$

while for $e^- < 0$ the left directional is:

$$V_\lambda^L(\lambda, c_i; e^-) = \max_{c_{-i} \in C_{-i}^{L*}} (s - p_{-i}c_{-i}(\lambda, c_i, s) - c_i) \cdot e^-.$$

Let $\Lambda^*(c_i, s)$ be the set of KKT multipliers in the aggregated dual problem. We now state the following lemma:

Lemma 4.2 *Under Assumption Q, in the dual problem* (11), *the set* $\Lambda^*(c_i, s)$ *of multipliers is* (1.a) *ascending in Veinott's strong set order in* c_i *for each* $s \in \mathbb{S}$, *with* $\wedge\Lambda^*(c_i, s)$ *and* $\vee\Lambda^*(c_i, s)$ *isotone selections in* c_i, (1.b) *descending in Veinott's strong set order in s for each* $c_i \in [0, s]$, *with* $\wedge\Lambda^*(c_i, s)$ *and* $\vee\Lambda^*(c_i, s)$ *antitone selections in s. Moreover,* (1.c) *the aggregated primal* $V^{ad}(c_i, s)$ *has increasing differences (and, hence, is supermodular) in* (c_i, s) *and* (1.d) *if* $H : \mathbb{R} \to \mathbb{R}$ *is a strictly increasing transformation, then* $V^{*ad}(c_i; s) = H(V^{ad}(c_i; s))$ *has the single-crossing property in* $(c_i; s)$.

Proof Given the partial directional derivative $V_\lambda^L(\lambda, c_i, s)$ in any direction e, just above the lemma, and since C_{-i}^{L*} is ascending in c_i for each s, and independent of s, it is easy to see that $V^L(\lambda, c_i, s)$ has (i) decreasing differences between (λ, c_i) for each s and (ii) increasing differences between (λ, s) for each c_i. Thus, (1.a) and (1.b) follow directly from Topkis' theorem (Topkis [22], Theorem 2.8.3).

To see (1.c), appealing to the partial concavity of $V^{ad}(c_i, s)$ in s, for each $c_i \in [0, s]$, we can simplify the calculation of the Clarke gradient of $V^{ap}(c_i, s)$ in Program (4) by exploiting the zero duality gap noted before (i.e., $V^{ap}(c_i, s) = V^{ad}(c_i, s)$) and Proposition 3.5 so that for $e^+ > 0$:

$$\begin{aligned} V_s^{ap}(c_i, s; e^+) &= \min_{\lambda \in \Lambda^*} \max_{c_{-i} \in C_{-i}^{d*}} \partial_s L(c_{-i}^d(c_i, s), \lambda^*(c_i, s), c_i, s) \cdot e^+ \\ &= \min_{\lambda \in \Lambda^*} \Lambda^*(c_i, s) \cdot e^+ \\ &= \wedge\Lambda^*(c_i, s) \cdot e^+ \end{aligned}$$

which is well defined as $\Lambda^*(c_i, s)$ is subchained. Since the selection $\wedge\Lambda^*(c_i, s)$ is isotone in c_i for each s, $V^{ap}(c_i, s)$ has increasing differences for any right perturbation. (A similar proof holds for left perturbations, i.e., when $e^- < 0$.) This proves that $V^{ap}(c_i, s)$ has increasing differences in (c_i, s) and is therefore supermodular. (1.d) follows as a strictly increasing transformation of a supermodular function is quasi-supermodular (hence, has the requisite single-crossing property). □

We now have our main comparative statics result:

Theorem 4.1 *Under Assumption Q,* (1.a) $C_i^* : \mathbb{S} \to \mathcal{L}(C_i)$ *is ascending in Veinott strong set order,* (1.b) $\vee C_i^*(s)$ *and* $\wedge C_i^*(s)$ *are isotone selections. Further,* (1.b) *if* $H{:}\mathbb{R} \to \mathbb{R}$ *is a strictly increasing transformation, the monotone comparative statics in* (1.a) *is also obtained. Finally,* (1.c) *if Assumption Q holds for all* $i \in \{1, 2, \ldots, n\}$, *then statements* (1.a) *and* (1.b) *are true for all i, and* $\vee C_i^*(s)$ *and* $\Lambda C_i^*(s)$ *each are Lipschitz selections of modulus* 1.

Proof Claims (1.a) and (1.b) follow from Topkis' theorem (e.g., Topkis ([22], Theorem 2.8.2) noting that the feasible set $[\wedge \sum_{j \neq i} c_j^*(\lambda^*(s, c_i), c_i), s]$ is strong set order increasing in s for all selection $c_j^* \in C_j^*(\lambda^*, c_i)$, and $\lambda^* \in \Lambda^*(s, c_i)$. The last claim in the theorem follows from Curtat ([37], Theorem 2.3) noting that $\vee C_i^*(s) = s - \sum_{-i} p_{-i} \cdot \vee C_{-i}^*(s)$ and $\sum_{-i} p_{-i} \cdot \vee C_{-i}^*(s)$ (resp., $\wedge C_i^*(s) = s - \sum_{-i} p_{-i} \cdot \wedge C_{-i}^*(s)$ and $\sum_{-i} p_{-i} \cdot \wedge C_{-i}^*(s)$) are both increasing in s, for all i, at $\vee C_i^*(s)$ (resp., $\wedge C_i^*(s)$) and s is Lipschitz of modulus 1. □

A few remarks about Theorem 4.1, which is basically a nonsmooth version of Chipman's result (e.g., Chipman ([38]). We note first that a similar approach can be taken to establish monotone comparative statics per price effects, since issues of (gross) substitutability or complementarity in demand are reduced to sufficient condition in the dual program (11) for $\lambda^*(c_i, p)$ to be falling or rising in p_i (for own substitution effects or cross-substitution effects). This comparative statics is rather easily incorporated into our dual formulation, but appears somewhat difficult to accommodate in Quah [33]. Second, if we only need good i strong set order isotone in s, we do not need the choice space $C_i \in \mathbb{E}_+$ to be convex, so the section C_i can be a discrete lattice. Quah's flexible set ordering approach cannot handle this case. Also, conditions for normality of all the selections for good i simply require strict increasing differences between $(c_{-i}; c_i)$, a result which follows from the arguments above. Further, in principle, our methods can be extended to integer programming problems (at least in static problems) using conjugation schemes developed in the literature on Lagrangian relaxation methods.

Finally, to study the ordinal case in principle using dual methods involves conjugating quasi-concave programs. For this, we need to pick a different duality scheme with an "ordinal" Lagrangian (for discussion, see, for example, Crouzeix [39] or Trach [40], and references therewithin).

5 Conclusions

In this paper we have shown that the Lipschitz property is preserved under maximization under very general conditions and therefore argued that the set of Lipschitz function is appropriate for a large class of economic models. Although not differentiable, Lipschitz functions have enough properties (they have Dini derivatives and Clarke derivatives, among other properties) for the formulation of nonsmooth envelope theorems for Lipschitz programs, and we derived a number of such envelopes. Our initial applications of some of our results to various economic models, in particular models associated with recursive dynamic programs, show that powerful methods can be developed when nonsmooth envelope theorems are combined with lattice programming techniques. We intend to further explore this promising combination.

We do not address the issue of computing optimal solutions for Lipschitz programming problems with nonsmooth constraints, but there is actually a rather large literature on this question. For theoretical issues, the interested reader can consult the excellent early survey by Sun and Han [41] as well as the classic monograph of Klatte and Kummer [42], especially chapters 9 and 10 which apply directly to the class of nonsmooth optimization problems we study in this paper.

When applied to nonsmooth programs typically studied in this paper, smooth gradient-based algorithms may fail to converge and/or traditional gradients of Lipschitz functions may not exist. It is important to notice, however, that the literature on nonsmooth optimization methods has developed numerous special tools for solving nonsmooth programs, such as replacing gradients with arbitrary subgradients (the "subgradient methods") or a combination of such (the "bundle methods"), or approximating subdifferentials by random sampling of gradients (the "gradient sampling methods"). These approaches represent a powerful set of numerical algorithms applicable to the computation of optimal solutions for the class of problem studied in this paper. For both an extensive discussion of the theoretical issues involved, as well as for a concise summary of the collection of numerical implementations for the actual computation of solutions that can be applied to approximate optimal solutions, we refer the interested reader to this literature, in particular the recent monograph of Bagirov et al. [43].

Acknowledgements We thank Rabah Amir, Bob Becker, Hector Chade, Bernard Cornet, Amanda Friedenberg, Martin Jensen, Cuong LeVan, Len Mirman, Ed Prescott, Juan-Pablo Rincon-Zapatero, and Carlo Strub for helpful conversations. The usual caveats apply.

Appendices

Appendix A discusses some of the features of Lipschitz functions, Appendix B presents some properties of correspondences, and Appendix C discusses some features of lattices and lattice programming.

Appendix A. Properties of Lipschitz Functions

Lipschitz Property. Given an open set $\Omega \subset \mathbb{R}^n$, a function $f : \Omega \to \mathbb{R}^m$ is said to be Lipschitz continuous (or simply "Lipschitz") at $x \in \Omega$ of modulus $k \geq 0$ if $\exists\ \delta > 0$ such that:

$$\forall x', x'' \in \delta B(x), \quad \left| f(x'') - f(x') \right| \leq k \left| x'' - x' \right|,$$

where $B(x)$ is the open ball of radius 1 centered on x. If the modulus k can be chosen independently of x on an entire subset of Ω, f is said to be globally Lipschitz on that subset. Note, for instance, that the function $x \longmapsto Log(x)$ is globally Lipschitz on any $[a, b]$ for $0 < a < b$, but only Lipschitz at every $x \in \Omega =]0, +\infty[$.

Dini Derivatives. Intuitively, to be Lipschitz at x means that the rate of change of f around x, no matter how it is calculated, cannot exceed the modulus k. In particular, it implies that the Dini derivatives, upper and lower bounds for the rate of growth of f at x in the direction d, respectively, defined as the functions:

$$d \longmapsto D^+ f(x; d) = \limsup_{t \downarrow 0} \frac{f(x+td) - f(x)}{t} \text{ and}$$
$$d \longmapsto D_+ f(x; d) = \liminf_{t \downarrow 0} \frac{f(x+td) - f(x)}{t},$$

always exist.

In the event the two Dini bounds coincide, f is said to be Gâ teaux (or directionally) differentiable at x, with Gâteaux derivative given by the common bound:

$$f'(x; d) = \lim_{t \downarrow 0} \frac{f(x+td) - f(x)}{t}.$$

If this Gâteaux derivative is a linear function of d, i.e., if $f'(x; d) = \nabla f(x) \cdot d$, then f is said to be differentiable at x. Note, for instance, that the Lipschitz function $x \longmapsto |x|$ is directionally (i.e., Gâteaux) differentiable, but not differentiable at $x = 0$. Failure to be differentiable "rarely happens," since Rademacher's theorem guarantees that if f is Lipschitz at all point of an open set $\Theta \subset \Omega$, then it is almost everywhere differentiable on Θ. Finally, if the function $x \longrightarrow \nabla f(.)$ is continuous at x, then f is said to be continuously differentiable at x.

Clarke Derivatives and Clarke Gradients. Being Lipschitz at x is a local condition, as it requires that the rate of change around x be bounded by k. Consequently, the upper and lower Clarke derivatives, respectively, defined as:

$$d \longmapsto f^o(x; d) = \lim \sup_{\substack{y \to x \\ t \downarrow 0}} \frac{f(y+td) - f(y)}{t} \text{ and:}$$
$$d \longmapsto f^{-o}(x; d) = \lim \inf_{\substack{y \to x \\ t \downarrow 0}} \frac{f(y+td) - f(y)}{t}$$

also always exist when f is Lipschitz at x.

The upper Clarke derivative is upper semicontinuous in x (hence the lower Clarke derivative is lower semicontinuous) as established by Clarke [7] (Proposition 2.1.1), and the Clarke derivatives define wider bounds than the Dini derivatives since:

$$f^{-o}(x; d) \leq Df_+(x; d) \leq Df^+(x; d) \leq f^o(x; d) = -f^{-o}(x; -d).$$

The Clarke gradient of a Lipschitz function f at x is the nonempty compact convex set:

$$\partial f(x) = cl \quad conv \left\{ \lim \nabla f(x_i) : x_i \to x, x_i \notin \Theta, x_i \notin \Omega_f \right\},$$

where *cl conv* denotes the closure of the convex hull, Θ is any set of Lebesgue measure zero in the domain, and Ω_f is a set of points at which f fails to be differentiable.

Clarke [7] (Proposition 2.1.5) shows that the correspondence $x \rightrightarrows \partial f(x)$ is upper hemicontinuous, and Clarke [7] (Proposition 2.1.2) shows that:

$$f^o(x; d) = \max_{\zeta \in \partial f(x)} \{\zeta . d\};$$

hence, $f^o(x; d)$ is a convex function of d.

Clarke Regular Functions. An important class of Lipschitz functions is the upper Clarke regular ("upper Clarke regular") functions, Lipschitz functions that are Gâteaux differentiable and for which the Gâteaux derivative coincides with the upper Clarke derivative, that is, $f^o(x; d) = f'(x; d)$. Lower Clarke regular functions are similarly defined, i.e., as Gâteaux differentiable Lipschitz functions such that $f^{-o}(x; d) = f'(x; d)$.

Several features combine to make the set of upper Clarke regular functions a natural extension of convex functions well suited to the study of economies with nonconvexities. First, except for some pathological cases, convex functions are upper Clarke regular. (Concave functions are lower Clarke regular, since f is upper Clarke regular iff $-f$ is lower Clarke regular.) Second, the theory, calculus, and properties of Clarke gradients have its precise counterpart in smooth and convex analysis. In particular, the Clarke gradient of a convex function coincides with the subgradient of convex analysis, i.e., the set of $p \in M_{m\times n}$ satisfying $\forall d,\ p \cdot d \leq f(x_0 + d) - f(x_0)$. Third, upper Clarke regularity grants more power to a Gâteaux derivative at a specific point and in a particular direction, since it then becomes an approximation for the maximum rate of growth of f in a *whole neighborhood* of that point in that direction. Such local behavior is one step short of continuous differentiability, as shown in the following result.

Lemma A. 1 *If $f : \Omega \to \mathbb{R}^m$ is upper Clarke regular and differentiable at x, then f is continuously differentiable at x.*

Proof Differentiability and upper Clarke regular together imply:

$$f^o(x; d) = f'(x, d) = \nabla f(x).d;$$

hence,

$$\begin{aligned} f^{-o}(x; d) &= -f^o(x; -d) = -\nabla f(x).(-d) \\ &= \nabla f(x).d = f^o(x; d). \end{aligned}$$

Function $x \to \nabla f(x)$ is thus both upper and lower semicontinuous at x; hence, f is continuously differentiable at x. □

Similarly, lower Clarke regular differentiable functions are continuously differentiable. Consequently, we note that if f is both upper and lower Clarke regular at x then f is continuously differentiable at x since:

$$f^o(x; d) = f'(x; d) = f^{-o}(x; d)$$

implies that $f'(x; d)$ is both convex and concave in d, hence linear in d. Thus, f is differentiable at x and therefore continuously differentiable by the lemma above.

Appendix B. Properties of Correspondences

A significant advantage to working in metric spaces is that the topological properties of correspondences can be stated exclusively in terms of sequences.

Definition B. 1 Given $A \subset \mathbb{R}^n$ and $S \subset \mathbb{R}^m$, a nonempty-valued correspondence $D : S \twoheadrightarrow A$ is:

(i) lower hemicontinuous at s if for every $a \in D(s)$ and every sequence $s_n \to s$ there exists a sequence $\{a_n\}$ such that $a_n \to a$ and $a_n \in D(s_n)$.
(ii) upper hemicontinuous at s if for every sequence $s_n \to s$ and every sequence $\{a_n\}$ such that $a_n \in D(s_n)$ there exists a convergent subsequence of $\{a_n\}$ whose limit point a is in $D(s)$.
(iii) closed at s if $s_n \to s$, $a_n \in D(s_n)$ and $a_n \to a$ implies that $a \in D(s)$. (In particular, this implies that $D(s)$ is a closed set.)
(iv) open at s if for any sequence $s_n \to s$ and any $a \in D(s)$, there exists a sequence $\{a_n\}$ and a number N such that $a_n \to a$ and $a_n \in D(s_n)$ for all $n \geq N$.

Note that $D(s) = \{a \in A,\ g_i(a, s) \leq 0,\ i = 1, \dots, p\}$, where the g_i are Lipschitz (thus continuous), is necessarily closed at s. The same property holds true in the presence of Lipschitz equality constraints.

Another property of correspondences which is critical in our analysis is that of uniform compactness.

Definition B. 2 A nonempty-valued correspondence D is said to be uniformly compact near s if there exists a neighborhood S' of s such that $cl\ [\cup_{s' \in S'} D(s)]$ is compact.

We note the result in Hogan [44] that if D is uniformly compact near s, then D is closed at s if and only if $D(s)$ is a compact set and D is upper hemicontinuous at s. When D is defined by a system of continuous equality and inequality constraints, uniform compactness near s thus implies compactness and upper hemicontinuity at s. In fact, for any s' sufficiently close to s, since $D(s')$ is a closed subset of $cl\ [\cup_{s' \in S'} D(s)]$ it is therefore compact.

Finally, we will need the following property of hemicontinuous correspondences (and thus of Clarke gradients).

Proposition B. 1 *If D is an upper hemicontinuous correspondence, then for every compact neighborhood K of x, the set:*

$$\bigcup_{z \in K} D(z)$$

is compact.

Proof Consider a sequence $\{y_n\}$ in $\bigcup_{z \in K} D(z)$ so that $y_n \in D(z_n)$ for some z_n in K. The sequence $\{z_n\}$ is the compact K, so there exists a subsequence of $\{z_{\varphi(n)}\}$ of $\{z_n\}$ converging to some $z' \in K$. By upper hemicontinuity of D at z', there exists a subsequence of $\{y_{\varphi(n)}\}$ converging to some $y \in D(z')$. This proves that the initial sequence $\{y_n\}$ has a convergent subsequence and therefore that the set $\bigcup_{x \in K} D(x)$ is compact. □

Appendix C. Posets, Lattices, Supermodularity, and Lattice Programming

A *partially ordered set* (or poset) is a set X ordered with a reflexive, transitive, and antisymmetric relation. If any two elements of X are comparable, X is referred to as a complete partially ordered set, or chain. An upper (resp., lower) bound of $B \subset X$ is an element x^u (resp., x^l) in B such that $\forall x \in B$, $x \leq x^u$ (resp., $x^l \leq x$). A *lattice* is a set X ordered with a reflexive, transitive, and antisymmetric relation $\geq$ such that any two elements x and x' in X have a least upper bound in X, denoted $x \wedge x'$, and a greatest lower bound in X, denoted $x \vee x'$. The product of an arbitrary collection of lattices equipped with the product (coordinatewise) order is a lattice. $B \subset X$ is a *sublattice* of X if it contains the sup and the inf (with respect to X) of any pair of points in B.

Let $(X, \geq_X)$ and $(Y, \geq_Y)$ be posets. A mapping $f : X \to Y$ is *isotone* (or *increasing*) on X if $f(x') \geq_Y f(x)$, when $x' \geq_X x$, for $x, x' \in X$. A correspondence (or multifunction) $F : X \to 2^Y$ is *ascending* in the set relation on 2^Y denoted by $\geq_S$ if $F(x') \geq_S F(x)$, when $x' \geq_X x$. A particular set relation of interest is Veinott's strong set order (see Veinott [35], Chapter 4). Let $L(Y) = \{A | A \subset Y,\ A \text{ a nonempty sublattice}\}$ be ordered with the *strong set order* $\geq_a$: if $A_1, A_2 \in L(Y)$, we say $A_1 \geq_a A_2$ if $\forall (a, b) \in A_1 \times A_2$, $a \wedge b \in A_2$ and $a \vee b \in A_1$.

Let X be a lattice. A function $f : X \to R$ is *supermodular (resp., strictly supermodular)* in x if $\forall (x, y) \in X^2$, $f(x \vee y) + f(x \wedge y) \geq$(resp., $>$) $f(x) + f(y)$. The class of supermodular functions is closed under pointwise limits (see Topkis [22], Lemma 2.6.1). Consider a partially ordered set $\Psi = X_1 \times P$ (with order $\geq$), and $B \subset X_1 \times P$. The function $f : B \longrightarrow R$ has *increasing differences* in (x_1, p) if for all $p_1, p_2 \in P$, $p_1 \leq p_2 \Longrightarrow f(x, p_2) - f(x, p_1)$ is nondecreasing in $x \in B_{p_1}$, where B_p is the p section of B. If this difference is strictly increasing in x then f has *strictly increasing differences* on B.

References

1. Amir, R., Mirman, L., Perkins, W.: One-sector nonclassical optimal growth: optimality conditions and comparative dynamics. Int. Econ. Rev. **32**(3), 625–644 (1991)
2. Bonnisseau, J.M., LeVan, C.: On the subdifferential of the value function in economic optimization problems. J. Math. Econ. **25**(1), 55–73 (1996)
3. Askri, K., LeVan, C.: Differentiability of the value function of nonclassical optimal growth models. J. Optim. Theory Appl. **97**(3), 591–604 (1998)
4. Milgrom, P., Segal, I.: Envelope theorems for arbitrary choices. Econometrica **70**(2), 583–601 (2002)
5. Rincon-Zapatero, J., Santos, M.: Differentiability of the value function without interiority assumptions. J. Econ. Theory **144**(5), 1948–1964 (2009)
6. Morand, O., Reffett, K., Tarafdar, S.: A nonsmooth approach to envelope theorems. J. Math. Econ. **61**, 157–165 (2015)
7. Clarke, F.: Optimization and Nonsmooth Analysis. SIAM, Philadelphia (1983)
8. Chatterjee, K., Krishna, R.V.: A nonsmooth approach to nonexpected utility theory under risk. Math. Soc. Sci. **62**(3), 166–175 (2011)
9. Maroto, J., Moran, M.: Lipschitz continuous dynamic programming with discounting. Nonlinear Anal. **62**, 877–894 (2005)
10. Hiriart Urruty, J.B.: Refinement of necessary optimality conditions in nondifferentiable programming. Appl. Math. Optim. **5**(1), 62–82 (1979)

11. Auslender, A.: Differentiable stability in non convex and non differentiable programming. Math. Program. Stud. **10**, 29–41 (1979)
12. Gauvin, J.: A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming. Math. Program. **12**(1), 136–138 (1979)
13. Kyparisis, J.: On the uniqueness of Kuhn–Tucker multiplier in nonlinear programming. Math. Program. **32**(2), 242–246 (1985)
14. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, New York (2000)
15. Gauvin, J., Dubeau, F.: Differentiable properties of the marginal function in mathematical programming. Math. Program. Stud. **19**, 101–119 (1982)
16. Tarafdar, S.: Optimization in Economies with Nonconvexities. Ph.D. Dissertation, Arizona State University (2010)
17. Benveniste, L., Scheinkman, J.: On the differentiability of the value function in dynamic models of economics. Econometrica **47**(3), 727–32 (1979)
18. Stockey, N., Lucas, R., Prescott, E.: Recursive Methods in Economics Dynamics. Harvard University Press, Cambridge (1989)
19. Laraki, R., Sudderth, W.: The preservation of continuity and Lipschitz continuity of optimal reward operators. Math. Oper. Res. **29**(3), 672–685 (2004)
20. Hinderer, K.: Lipschitz Continuity of value functions in Markovian decision processes. Math. Methods Oper. Res. **62**(1), 3–22 (2005)
21. Hopenhayn, H., Prescott, E.: Stochastic monotonicity and stationary distributions in dynamic economies. Econometrica **60**(6), 1387–1406 (1992)
22. Topkis, D.: Supermodularity and Complementarity. Princeton University Press, Chichester (1998)
23. Mirman, L., Morand, O., Reffett, K.: A qualitative approach to Markovian equilibrium in infinite horizon economies with capital. J. Econ. Theory **139**(1), 75–98 (2008)
24. Tanaka, Y.: Nonsmooth Optimization for Production Theory. Hokkaido University, Sapporo (2008)
25. Coleman, W.J.: Uniqueness of an equilibrium in infinite-horizon economies subject to taxes and externalities. J. Econ. Theory **95**(1), 71–78 (2000)
26. Greenwood, J., Huffman, G.: On the existence of nonoptimal equilibria in dynamic stochastic economies. J. Econ. Theory **65**(2), 611–623 (1995)
27. Kamihigashi, T., Roy, S.: Dynamic optimization with a nonsmooth, nonconvex technology: the case of a linear objective function. Econ. Theory **29**(2), 325–340 (2006)
28. Kamihigashi, T., Roy, S.: A nonsmooth, nonconvex model of optimal growth. J. Econ. Theory **132**(1), 435–460 (2007)
29. Royden, H.: Real Analysis. MacMillan Press, Basingstoke (1968)
30. Antoniadou, E.: Lattice Programming and Economic Optimization. Ph.D. Dissertation, Stanford University (1996)
31. Antoniadou, E.: Comparative statics for the consumers problem. Econ. Theory **31**(1), 189–203 (2007)
32. Mirman, L., Ruble, R.: Lattice theory and the consumer's problem. Math. Oper. Res. **33**(2), 301–314 (2008)
33. Quah, J.: The comparative statics of constrained optimization problems. Econometrica **75**(2), 401–431 (2007)
34. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
35. Veinott, A.: Lattice Programming: Qualitative Optimization and Equilibria. Stanford University, Stanford (1992)
36. Grinold, R.: Lagrangian subgradient. Manag. Sci. **17**(3), 185–188 (1970)
37. Curtat, L.: Markov equilibria of stochastic games with complementarities. Games Econ. Behav. **17**(2), 177–199 (1996)
38. Chipman, J.: An empirical implication of Auspitz–Lieben–Edgeworth–Pareto complementarity. J. Econ. Theory **14**(1), 228–231 (1977)
39. Crouzeix, J.P.: Contribution á l'é tude des fonctions quisiconvexes" Thése, Université de Clemont (1977)
40. Trach, P.T.: Quasiconjugates of functions, duality relationship between quasiconvex minimization under reverse convex constraint and quasiconcave maximization under a convex constraint, and applications. J. Math. Anal. Appl. **159**(2), 299–322 (1991)
41. Sun, D., Han, J.: Newton and quasi-Newton methods for a class of nonsmooth equations and related problems. SIAM J. Optim. **7**(2), 463–480 (1997)

42. Klatte, D., Kummer, B.: Nonsmooth Equations in Optimization: Regularity, Calculus, Methods, and Applications. Kluwer, Dordrecht (2002)
43. Bagirov, A., Karmitsa, N., Makela, M.: Introduction to Nonsmooth Optimization: Theory, Practice, and Software. Springer, Berlin (2014)
44. Hogan, W.: Point-to-set maps in mathematical programming. SIAM Rev. **15**(3), 591–603 (1973)