Generalized Envelope Theorems with Applications to Dynamic Programming*

Olivier Morand† Kevin Reffett‡ Suchismita Tarafdar§

Preliminary Draft
Please do not Circulate

Abstract

We derive a collection of generalized envelope theorems for a broad class of parameterized Lipschitizian optimization problems with both nonsmooth objectives and constraints applicable to many economic environments where nonconvexities play a key role.

We first provide sufficient conditions for the value function to be Lipschitz and obtain bounds for its upper and lower Directional Dini derivatives of this value function. Next we establish sufficient conditions for the directional differentiabiliy and/or differentiability of the value function, and show how standard smooth envelope theorems are special cases of our results.

We then apply our findings to derive new results on the existence and characterization of Markov equilibrium in dynamic economies with nonconvexities, dynamic programming with discrete choices, incentive constrained dynamic programming, and monotone comparative statics in constrained optimization problems.

1 Introduction

The use of envelope theorems to characterize optimal solutions of constrained optimization problems is widespread in microeconomic and macroeconomic theory. An envelope is basically an equality between the derivative of the value function and the derivative of the objective evaluated at the optimum along a fixed direction ignoring the "indirect" effects due to changes in the optimal solution. While envelopes of convex unconstrained programs are derived directly from the objective of the optimization program, in convex constrained programs they are generally obtained from the Lagrangian, both cases involving the use of super/subdifferentials.

*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.

†Department of Economics, University of Connecticut

‡Department of Economics, WP Carey School of Business, Arizona State University

§Department of Economics, Shiv Nadar University

In smooth (meaning continuously differentiable) convex programs, envelopes are typically standard derivatives giving precise information concerning the rate of growth of the value function in all directions at a given point, and are therefore an essential tool for comparative statics in static and dynamic models.

Nonconvexities, however, arise in many problems such as dynamic programming with discrete choices, mechanism design with Lipschitz primitive data, constrained lattice programming problems, incentive constrained dynamic programs, "bi-level"/Stackelberg games, to name a few. And to make matters more difficult, objectives and/or constraints are not even smooth in many of these problems. Clearly one then cannot expect envelopes to be simple derivatives as many technical difficulties arise simultaneously: Some constraints may be active, Lagrange multipliers may not be unique, and the absence of traditional derivatives (or gradients) mandates the use of some sort of generalized gradients which, typically, are not singletons.

Progress has been made in many cases but findings are spread out in many places (see, for instance Amir, Mirman and Perkin [1], Bonnisseau and LeVan [9], Askri and LeVan [5], Milgrom and Segal [31], Ricon-Zapatero and Santos [40]). We seek in this paper a comprehensive and unified way to present generalized envelope theorems for a large class of nonconvex and/or nonsmooth programs.

Fundamentally, we seek a class of programs constructed with continuous functions (objective and constraints) not necessarily convex and/or continuously differentiable, but with properties strong enough for the existence of some generalized derivatives. We show that Lipschitz programs meet this objective and provide a suitable environment for the generalization of existing classical envelope theorems.

An important feature of Lipschitz programs is the preservation of the Lipschitz property by maximization under relatively weak hypothesis, as demonstrated in Section 2. One of these hypothesis is a non-smooth constraint qualification related to the work of Hiriart-Urruty [23] and Auslender [6] and easily checked in applications. Combined with the Lipschitz objective and constraints and some geometric conditions on the choice domain used in Clarke [11], this non-smooth constraint qualification is sufficient for the value function to be Lipschitz.

Lipschitz value functions have well defined Dini derivatives, and our first result in Section 3 gives lower and upper bounds for these Dinis, a characterization which may prove useful in computational work, as well as for establishing the absolute continuity of the value function (an essential step in the proof of its supermodularity in one of the application later in the paper). Section 3 then consists in narrowing these bounds to sharpen the characterization of the rates of growth of the value function in specific directions. Gateaux differentiability (the existence of directional derivatives) is the next step, since it permits comparative statics in all directions at a specific point. The addition of Clarke regularity grants more power to Gateaux derivatives (at one point) who then behave almost like bounds on the rate of growth of the value function in a neighborhood of that point, just one step short of continuous differentiability. Clarke regularity also, in some conditions, is preserved under maximization (e.g. concave functions are lower Clarke regular).

Continuously differentiable functions are Lipschitz, and convex functions are (upper) Clarke regular, hence classical envelope theorems are just special cases of our more general results.

In Section 4 our results are applied to dynamic programming and to a proof of existence of equilibrium in a large class of dynamic models. Some brief but essential mathematical definitions and results are gathered in the Appendix.

2 Lipschitz programs

We 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 $D : S \rightrightarrows A$ the feasible correspondence defined as:

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

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

The choice set A and the state space (or parameter space) S are both are open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. In contrast to standard "smooth" optimization problems in which constraints and objectives are continuously differentiable (i.e., C^1 or "smooth") all functions f, g_i and h_j are initially only assumed to be Lipschitz[1] at every $(a, s) \in A \times S$. The function $V : S \to \mathbb{R}$, defined as $V(s) = \max_{a \in D(s)} f(a, s)$, is the value function, and the correspondence $A^* : S \rightrightarrows A$, defined as $A^*(s) = \arg\max_{a \in D(s)} f(a, s)$, is the optimal solution correspondence. The classical Lagrangian[2] 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.

2.1 Constraint Qualifications

Recall the definition of a KKT point:

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

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

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

Denoting by $K(a, s)$ the closed and convex (but possibly empty) set of vectors (λ, μ) satisfying the above "multiplier rule" at (a, s), the role of constraint qualifications is precisely to guarantee this set is no-empty and bounded. A standard CQ in smooth programs is the Mangasarian-Fromovitz CQ, defined as follows:

[1] Lipschitz in the sense of "locally Lipschitz" (see Appendix for definitions and some mathematical results).

[2] If A is closed, then the abstract constraint $a \in A$ induces an additional term in the Lagrangian (see, for instance, Clarke [11], Chapter 6).

Definition 2 *The Mangasarian-Fromovitz Constraint Qualification (MFCQ) is satisfied at* $a^*(s) \in A^*(s)$ *if there exists* $y \in \mathbb{R}^n$ *such that:*

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

where $I(a^*(s), s)$ *is the set of indexes of the active inequality constraints (those for which* $g_i(a^*(s), s) = 0$*), and the matrix* $\nabla_a h(a^*(s), s)$ *has full rank.*

Gauvin ([17], Lemma 1) proved that, in smooth programs, MFCQ at $a^*(s)$ is equivalent to the compactness of $K(a^*(s), s)$. Kyparisis [29] sharpened this result under the following slightly less general condition which simply treats active inequality constraints for which multipliers are strictly positive ("binding constraints") as equality constraints.

Definition 3 *The Strict Mangasarian-Fromoviz Constraint Qualification (SMFCQ) is satisfied at* $a^*(s) \in A^*(s)$ *if there exists* $y \in \mathbb{R}^n$ *such that:*

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

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

Kyparsis ([29], Proposition 1.1) showed that the SMFCQ is both necessary and sufficient for $K(a^*(s), s)$ to be a singleton in smooth programs.(see also Bonnans and Shapiro [8], Remark 4.49).

No classical gradients generically exist in Lipschitz programs, so we rely on a generalization of the MFCQ (referred to as the "Generalized MFCQ", or GMFCQ) introduced by Hiriart-Urruty [23] and stated in terms of Clarke's generalized gradients. Denoting by $\overline{g}(a^*(s), s)$ the vector of binding inequality constraints at $a^*(s)$, so that $\overline{g} : A \times S \to \mathbb{R}^{\overline{p}}$ (where $\overline{p} = Card(I(a^*(s), s))$), GMFCQ can be stated as follows:

Definition 4 *The Generalized Mangasarian-Fromovitz Constraint Qualification (GMFCQ) is satisfied at* $a^*(s) \in A^*(s)$ *if there exists* $y \in \mathbb{R}^n$ *such that:*

$$\forall (\gamma_a, \upsilon_a) \in \partial_a(\overline{g}, h)(a^*(s), s),\ \gamma_a \cdot y < 0,\ and\ \upsilon_a \cdot y = 0$$

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

Hiriart-Urruty ([23],Theorem 4.2) proved that the GMFCQ at some $a^*(s)$ implies the non-emptiness of $K(a^*(s), s)$.

Remark 5 *We note that:*

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

so this version of GMFCQ is slightly more general than that of Auslender ([6], Theorem 2.1).

2.2 Lipschitz value functions

The Lipschitz properties of both objective and constraints in Program (1) are, of course, not sufficient for the value function V to even be continuous, and additional restrictions are therefore needed to insure that V is Lipschitz. We adopt the following general hypothesis made in Clarke ([11], Hypothesis 6.5.1, page 241):

Criterion 6 *(Clarke's Hypothesis) $V(s)$ is finite and there exists a compact set Λ and a positive number ε_0 such that for all $s' \in \varepsilon_0 B(s)$ for which $V(s') \geq V(s) - \varepsilon_0$, necessarily $A^*(s') \cap \Lambda \neq \emptyset$.*

Although Clarke's Hypothesis is not expressed in terms of primitive data of the problem, we show that it is satisfied in three important cases. The first is a simple condition mentioned by Clarke, the second is the inf-compactness condition used by Bonnans and Shapiro [8] to derive stability results for differentiable programs, and the third is the uniform compactness condition central to the work of Gauvin and Dubeau [18].[3]

Proposition 7 *Clarke's Hypothesis in the following three cases:*

(i) (growth condition) $\forall r \in \mathbb{R}$, the sets $\{(a', s') \in A \times S,\ f(a', s) \geq r\}$ are compact;

(ii) (inf-compactness) there exists $r \in \mathbb{R}$ 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 a nonempty set contained in Ω;

(iii) (uniform compactness) there exists a neighborhood S' of s such that $cl\left[\cup_{s' \in S'} D(s)\right]$ is compact.

Proof. (i). Given any $s \in S$ necessarily there exists some r such that $G(s) = \{a \in A, f(a, s) \geq r\} \cap D(s) \neq \varnothing$; by hypothesis $G(s)$ is compact so $A^*(s) = \arg\max_{D(s)} f(a, s) = \arg\max_{G(s)} f(a, s)$ is non-empty and $V(s)$ is finite. Consider $s' \in \varepsilon_0 B(s)$ such that $V(s') \geq V(s) - \varepsilon_0$. Since $A^*(s')$ is nonempty and included in the compact (by hypothesis) set $\Lambda = \{(a, s') \in A \times S, f(a, s') \geq V(s) - 2\varepsilon_0\}$, Clarke's Hypothesis is satisfied.

(ii). By inf-compactness at s, there exists r such that if $s' \in \varepsilon_0 B(s)$ then $\{a' \in D(s'),\ f(a', s') \geq r\}$ is nonempty and included in the compact set Ω. Letting $\Lambda = \Omega, \forall s' \in \varepsilon_0 B(s)$ necessarily $A^*(s') \subset \Omega$ and $A^*(s') \subset \{a' \in D(s'),\ f(a', s') \geq r\} \subset \Omega$ and $A^*(s')$ is nonempty thus $A^*(s) \cap \Lambda \neq \varnothing$ hence Clarke's Hypothesis is satisfied.

(iii). Uniform compactness of D near s implies the existence of a neighborhood S' of s such that $cl\left[\cup_{s' \in S'} D(s)\right]$ is compact. As a result:

$$\exists \varepsilon_0 > 0,\ s' \in \varepsilon_0 B(s) \Longrightarrow\ D(s') \text{ is compact}$$

since $D(s')$ is a closed subset of the compact $cl\left[\cup_{s' \in S'} D(s)\right]$. This implies that $A^*(s')$ is nonempty (the objective f is continuous on the compact $D(s')$), and Clarke's Hypothesis is satisfied by letting $\Lambda = cl\left[\cup_{s' \in S'} D(s)\right]$. ■

When combined with the GMFCQ, any one of these conditions implies a very powerful result analogous to Berge's maximum theorem on the preservation of continuity under maximization: The value function is Lipschitz and A^* is upper hemicontinuous. The upper hemicontinuity of the optimal correspondence is a very important property, since it implies that as s_n converges to s the maxima of $f(., s_n)$ become arbitrarily close to some of the maxima of $f(., s)$.

---

[3] A fourth mild compactness condition is given later in this section.

Theorem 8 *If the GMFCQ holds at any $a^*(s) \in A^*(s)$ and if Clarke's Hypothesis is satisfied (hence whenever any of the condition of Proposition 7 is satisfied), V is Lipschitz at s and A^* is upper hemicontinuous at s. Moreover:*

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

Proof. The Lipschitz property (and therefore continuity) of V and the formula for the generalized gradient follow directly from Clarke [11] (Corollary 1, page 242). The upper hemicontinuity of A^* is established separately for the three conditions of Proposition 7.

(i). Given any $s_n \to s$ and any sequence $\{a_n\}$ such that $a_n \in A^*(s_n) \subset D(s_n)$, $V(s_n) = f(a_n, s_n) \to V(s)$ by continuity of V. Given $\varepsilon' > 0$ there exists N such that $\forall n \geq N$, $f(a_n, s_n) \geq V(s) - \varepsilon'$. By the growth condition the sequence $\{a_n, s_n\}_{n\geq N}$ belongs to a compact set, and therefore has a convergent subsequence to $\{a, s\}$. By continuity of f, $f(a_n, s_n) \to f(a, s)$, hence $f(a, s) = V(s)$, and by closeness of D at s, $a \in D(s)$, hence $a \in A^*(s)$.

(ii). Under the inf-compactness condition, there exists r such that the set $A_s = \{a' \in D(s'), f(a', s') \geq r\}$ 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 closeness of D implies the desired result.

(iii). 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 belong 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 from a direct application of Berge's maximum theorem since the feasible correspondence D is not necessarily continuous, even though all constraints are continuous. 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 same conditions sufficient for the preservation of the Lipschitz property under maximzation in Theorem (8) are shown below to also be sufficient for the derivation of specific bounds for the Dini derivatives of the value function. In the rest of this section, we impose several stronger conditions on the primitive data (such as concavity, differentiability, Clarke regularity, continuous differentiability) to derive sharper envelope theorems going beyond the simple existence of bounds all the way to C^1 envelopes.

3.1 A Central Result on Stability Bounds

Under the conditions of Theorem 8 the value function is Lipschitz, so its Dini derivatives exist. Our first result states specific bounds for these Dini derivatives, obtained as a consequence of Clarke [11] Corollary 4 (page 243) (see also Tarafdar [44] for an alternative proof independent of Clarke's results), and expressed in terms of the primitive data.

Theorem 9 *If GMFCQ holds at any $a^*(s) \in A^*(s)$ and under Clarke's Hypothesis, 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 proof, 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":

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

with its associated Lagrangian:[4]

$$L_m((a,a'),s,\lambda,\theta) = -f(a,a') + \lambda g(a,a') + \theta\left[-a'+s\right]$$

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, in the sense that $\lambda \in K(a^*(s),s)$ if and only if $(\lambda,\theta) \in K_m(a^*(s),s)$.

By Theorem 6.1.1 in Clarke [11], there exists $\lambda \geq 0$, and θ such that $\lambda g(a^*(s),s) = 0$ and:

$$0 \in \partial_{(a,a')} L((a^*(s),s),s,\lambda,\theta)$$

which 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:

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

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

and

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

The assumptions of Corollary 4 in Clarke [11] are satisfied (in Clarke's notations, if the GMFCQ holds at each $a^*(s)$ then $M^0(\sum) = \{0\}$) hence:

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

and:

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

Noticing that $(-V)^+(s;x) = -V_+(s;x)$ and that $(-V)_+(s;x) = -V^+(s;x)$, we obtain:

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

and in a similar manner:

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

which proves the desired result, noting that both sets $\partial_s(f - \lambda g)(a^*(s), s)$ and $A^*(s)$ are compact valued, so inf and sup become min and max, respectively. ■

3.2 Differentiability of the Value Function

3.2.1 De-constraining a program

Getting additional characterization of the rate of growth of the value function requires more than just the Lipschitz structure of both the objective and the constraints. Things would be simpler in the absence of constraints, so our first results concern Lipschitz programs satisfying a mild compactness condition which imply that.constraints can be locally (at least) ignored and that and Clarke's Hypothesis is automatically satisfied.

Proposition 10 *Suppose that there exist a compact set Λ and a neighborhood $N(s)$ of s such that $\forall s' \in N(s), A^*(s') \subset \Lambda \subset D(s')$, 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 Gateaux differentiable and:*

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

Proof. The hypothesis implies that:

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

hence by Berge's theorem of the maximum, V is continuous at s thus finite. Clarke's Hypothesis is satisfied by choosing any $\varepsilon_0 > 0$ such that $\varepsilon_o B(s) \subset N(s)$. Consequently, V is Lipschitz at s by Theorem 8, and:

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

The additional assumption of Clarke upper regularity permits squeezing together upper and lower Dini derivatives of V. Indeed, for all $a^*(s) \in A^*(s)$:

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

the last equality obtained from the Gateaux differentiability of f in s for each $a^*(s)$. In addition:

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

the last equality from the Clarke upper regularity of f in s for each $a^*(s)$. Upper and lower Dinis therefore coincide, hence V is Gateaux differentiable and:

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

We note also that $-V(s'-x) \leq V'(s;x)$. ■

Remark 11 *Since continuously differentiable functions are necessarily upper Clarke regular, V is Gateaux differentiable whenever the compactness condition of Proposition 10 holds and f is continuously differentiable in s for each $a^*(s) \in D(s)$. This is precisely Lemma 3.1 in Amir, Mirman and Perkins [1].*

To generate sharper differentiability properties beyond that of Gateaux differentiability, one must seek restrictions guaranteeing the preservation of some form of regularity (upper or lower) of the objective under maximization. One can for instance, follow Clarke [11] (specifically, Theorem 2.8.2) and impose conditions sufficient for V to inherit the upper Clarke regularity of the objective, as done for instance in Askri and LeVan [5].

Alternatively, assuming the choice domain is convex, concavity (which implies lower Clarke regularity) is preserved under maximization, a property we exploit next by combining it with differentiability.

Corollary 12 *Under the compactness condition of Proposition 10, if f is concave in (a,s) and differentiable at s for each a, and if $graphD$ is convex, then V is continuously differentiable.*

Proof. Concave functions are lower Clarke regular, and Clarke regular differentiable functions are in fact continuously differentiable (see Appendix) and therefore upper Clarke regular. Consequently, by Proposition 10 V is Gateaux differentiable and:

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

If $graphD$ is convex, then V is concave hence lower Clarke regular thus for all x:

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

and, therefore:

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

so V is continuously differentiable at s. ■

3.2.2 Constrained programs

Next we turn our attention to general Lipschitz programs with continuously differentiable objective and constraints and for which the SMFCQ guarantees that Clarke gradients are singletons and the the multiplier is unique. In that case, the value function can be shown to be at least Gateaux differentiable.

Proposition 13 *Under Clarke's Hypothesis, if the SMFCQ holds at every optimal solution $a^*(s) \in A^*(s)$, and the primitive data is continuously differentiable in s, then V is Gateaux differentiable at s and:*

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

Proof. Follows from Theorem 8 and Theorem 9 given uniqueness of multipliers. ■

An alternative to SMFCQ is to assume enough concavity to "squeeze" the lower and upper Dini bounds to obtain directional envelopes, as done in Milgrom and Segal ([31], Corollary 5). We derive such a result for a less restrictive setting.

Corollary 14 *Under Clarke's hypothesis and if the GMFCQ holds at every $a^*(s) \in A^*(s)$, if the primitive data is continuously differentiable in s, f, and $-g$ concave, h affine in a, the derivatives f_s, $-g_s$, h_s are upper semicontinuous in a, then V is Gateaux differentiable and:*

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

Proof. By Theorem 9:

$$\max_{a^*(s)\in K(a^*(s),s)} \min_{\lambda\in K(a^*(s),s)} L_2(a^*(s), s, \lambda, \mu) \leq D_+V(s;x)$$

Imposing additional conditions on the primitive data helps tighten the upper bound as follows. First, choose a sequence $\{t_n\} \downarrow 0$ such that:

$$\limsup_{t\downarrow 0} \frac{V(s+tx) - V(s)}{t} = \lim_{n\to\infty} \frac{V(s+t_n x) - V(s)}{t_n}$$

Next, consider a sequence $\{a^*(s+t_n x)\}$ with $a^*(s+t_n x) \in A^*(s+t_n x)$ for all $n \in \mathbb{N}$. By Clarke's Hypothesis, for n large enough all $a^*(s+t_n x)$ belong to a compact set, so without loss of generality we assume that $a^*(s+t_n x)$ converges to some a^*. By closeness of D, necessarily in $a^* \in D(s)$; by continuity of V, $V(s) = f(a^*, s)$. As a result $a^* \in A^*(s)$ is a global maxima. By strong duality, the Lagrangian has a global saddle point at (a^*, s, λ, μ) where $(\lambda, \mu) \in K(a^*, s)$. Thus, for any $(\lambda, \mu) \in K(a^*, s)$:

$$\begin{aligned}
&\lim_{n\to\infty} \frac{V(s+t_n x) - V(s)}{t_n} \\
&= \lim_{n\to\infty} \frac{L(a^*(s+t_n x), s+t_n x, \lambda_n, \mu_n) - L(a^*, s, \lambda, \mu)}{t_n}
\end{aligned}$$

where $(\lambda_n, \mu_n) \in K(a^*(s+t_n x), s+t_n x)$. Consequently for any $(\lambda, \mu) \in K(a^*, s)$:

$$\begin{aligned}
&\lim_{n\to\infty} \frac{V(s+t_n x) - V(s)}{t_n} \\
&\le \lim_{t_n\to 0^+} \frac{L(a^*(s+t_n x), s+t_n x, \lambda, \mu) - L(a^*, s, \lambda, \mu)}{t_n} \\
&\le \lim_{t_n\to 0^+} \frac{L(a^*(s+t_n x), s+t_n x, \lambda, \mu) - L(a^*(s+t_n x), s, \lambda, \mu)}{t_n} \\
&= L_s(a^*(s+t_n x), s, \lambda, \mu) \cdot x
\end{aligned}$$

The first and the second inequality follows from the fact that $(a_n^*(s+t_n x), s+t_n x, \lambda_n, \mu_n)$ and (a^*, s, λ, μ) are global saddle points of L for any $s+t_n x$ and s respectively.

As a result:

$$\begin{aligned}
&\limsup_{t\to 0^+} \frac{V(s+tx) - V(s)}{t} \\
&\le \min_{(\lambda,\mu)\in K(a^*(s),s)} L_s(a^*(s+t_n x), s, \lambda, \mu) \cdot x
\end{aligned}$$

Given that $a_n^*(s) \to a^* \in A^*(s)$ and that $L_s(.)$ is upper semicontinuous in its first argument, the above inequality implies:

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

Thus $D_+V = D^+V$ and the result follows. ■

Next, in the absence of equality constraints, we show that if the primitive data is jointly concave and continuously differentiable, and if the SMFCQ hold for every optimal solution, then the value function is once continuously differentiable.

Corollary 15 *Under Clarke's hypothesis, if the primitive data is continuously differentiable in (a, s), f, and $-g$ are jointly concave in (a, s), and SMFCQ holds at all $a^*(s)$ in $A^*(s)$, then V is continuously differentiable and:*

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

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

Proof. Under SMFCQ, and with continuously differentiable primitives, the multiplier is unique and by Corollary 14 V is Gateaux differentiable with:

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

in which $\left\{\lambda_{a^*(s)}\right\} = K(a^*(s),s)$. In addition:

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

The concavity of V (inherited from that of f and $-g$), together with its Gateaux 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:

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

hence V is continuously differentiable at s and:

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

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

We note that Corollary 15 does not require the set of optimal solutions to be singleton. However, any optimal solution along with its associated unique multiplier can be used to calculate the gradient of the value function.

4 Applications and Extensions

4.1 Lipschitz Dynamic Programming

We extend the results of Laraki and Sudderth [30] and Hinderer [22] on the preservation of Lipschitz continuity in recursive dynamic programs by weakening the global Lipschitz conditions on the primitive data to local Lipschitzness.

Consider the following dynamic program:

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

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

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

and its solution set $A^*_{n+1}(s) = \arg\max_{a\in D(s)} f(a,s) + \beta V_n(a)$. 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 8.

Proposition 16 *If the GMFCQ is satisfied for all* $a^*_{n+1}(s) \in A^*_{n+1}(s)$ *for each* n*, then if Clarke's Hypothesis is satisfied (or under any of the conditions in Proposition 7 the sequence* $\{V_n\}$ *is a sequence of Lipschitz functions, with Clarke gradients given by:*

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

In particular, if all constraints are inactive (i.e., $g(a^*_{n+1}(s), s) < 0$) then GMFCQ is trivially satisfied and:

$$\partial V_{n+1}(s) \subset co \left\{ \bigcup_{a^*_{n+1}(s) \in A^*_{n+1}(s)} \partial_s f(a^*_{n+1}(s), s) \right\}$$

whenever Clarke's Hypothesis holds.

It is well known that the sequence $\{V_n\}$ 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, since the Weistrass Approximation Theorem asserts that any continuous functions, Lipschitz or not, may be uniformly approximated by polynomials (which are Lipschitz).

Nevertheless, it is possible to prove that V is Lipschitz (and more) under certain conditions. One can, for instance, demand that a global Lipschitz condition be satisfied, as in Laraki and Sudderth [30] and Hinderer [22]. Alternatively, one can work impose sufficient structure on the primitive data to guarantee that all optimal solutions are located on a compact domain on which the objective is then globally Lipschitz, as in Askri and LeVan [5], in effect "de-constraining" the program.

Concavity, 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. We exploit the implications of concavity in dynamic programs next. requirement may be sufficient to guarantee that V^* is locally Lipschitz as discussed next.

4.1.1 Concave Dynamic Programming

Much more (than just Lipschitzness) can be revealed concerning the differentiability properties of the value function in the presence of concavity, providing the primitive data is also assumed to be differentiable with respect to s. 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 17 *Assuming that (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 MFCQ is satisfied at every optimal solution, and (iv) Clarke's Hypothesis is satisfied, then the Lipschitz value function* V *is concave and Gateaux differentiable with:*

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

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 Corollary 14 to this program implies that V is Gateaux differentiable with:

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

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 Corollary 15:

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

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

Our result on the continuous differentiability generalizes Benveniste and Scheinkman [7] 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 [40].

4.1.2 Differentiability of the Pareto Frontier

Consider the model in Kocherlakota and Koeppl (see also Rincon-Zapatero and Santos [40]) of an exchange economy in which two infinitely lived agents receive a stochastic endowment in each period which they mutually share under limited commitment. As in Koeppl, the endowment for agent $i = 1, 2$ in period t is (ω_s^1, ω_s^2) which is determined by the realization of θ_t. The stochastic process $\theta = \{\theta_1, \theta_2, ...\}$ is a sequence of iid random variables, each having finite support $\Theta = \{1, 2, ..., S\}$, and . The probability that θ_t equals s is denoted by $\pi_s = \Pr\{\theta_t = s \in \Theta\}$.

We will assume the following, in which we relax Koeppl's assumptions of strict monotonicity, strict concavity and C^2 utility function.

Assumption 4.1.2: The utility function $u : \mathbb{R}_+ \to \mathbb{R}$ is increasing, concave, continuously differentiable with $\lim_{c \to 0^+} u'(c) = \infty$, and $0 < \beta < 1$, and for each U_0, the feasible set is uniformly compact.

We characterize the incentive feasible allocations (see Koeppl for details).

$$V(U_0) = \max_{\{c_s, u_s\}_{s=1}^S} \sum_{s=1}^S \pi_s \left[u(\overline{\omega}_s - c_s) + \beta V(U_s)\right]$$

subject to

$$
\begin{aligned}
& U_0 - \sum_{s=1}^{S} \pi_s \left[u(c_s) + \beta U_s \right] \leq 0 \\
& u(\omega_s^1) + \beta U_{aut} - u(c_s) - \beta U_s \leq 0 \\
& u(\overline{\omega}_s - \omega_s^1) + \beta U_{aut} - u(\overline{\omega}_s - c_s) - \beta V(U_s) \leq 0 \\
& U_s \in [U_{aut}, U_{\max}]
\end{aligned}
$$

The set of optimal solutions is denoted $Y^*(U_0)$, and a typical element of this set is $\{c_s^*, U_s^*\}_{s=1}^S$; the KKT multiplier vector takes the form $(\lambda_1, \{\lambda_{2s}\}, \{\lambda_{3s}\}, \{\lambda_{4s}\}, \{\lambda_{5s}\}) \in K(\{c_s^*, U_s^*\})$. Given assumption *4.1.2*, and the joint concavity of the objective and the constraints in $(\{c_s, u_s\}, U_0)$, the hypothesis of Proposition 17 are met. Thus, by Proposition 17, if GMFCQ is satisfied for every optimal solution $(c^*(U_0), U^*(U_0)) \in Y^*(U_0)$, then V is concave with directional derivatives given by,

$$V'(U_0; x) = \max_{(c^*, U^*) \in Y^*(U_0)} \min_{\lambda \in K(c_s^*, U_s^*)} \{-\lambda_1 \cdot x\}$$

Further, if SMFCQ is satisfied for every optimal solution $(c^*(U_0), U^*(U_0)) \in Y^*(U_0)$, then V is concave and C^1 with derivative given by:

$$\nabla V(U_0) = -\lambda_1$$

for any $(\lambda_1, \{\lambda_{2s}\}, \{\lambda_{3s}\}, \{\lambda_{4s}\}, \{\lambda_{5s}\}) \in K(\{c_s^*, U_s^*\})$.

4.2 Optimization problems with discrete choice variables

Consider a Lipschitz program in which some of the decision variables (without loss of generality, a_1) is constrained to take only one of r possible values, that is:

$$\max_{a \in D(s)} f(a, s) \tag{2}$$

where $D(s) = \{g(a, s) \leq 0$, and $a_1 = b_j \ j = 1, ..., r\}$.. To apply our reusults, we rewrite the r equality constraints $a_1 = b_j$ as the C^1 equality constraint $h(a, s) = \prod_{j=1}^{r} (a_1 - b_j) = 0$. If $a^*(s) \in A^*(s)$, then $a_1^*(s)$ must equal some b_k, hence: $\nabla_a h(a^*(s), s) = (\prod_{j \neq k} (b_k - b_j), 0, ..., 0) \neq 0$.

Associated with the above maximization problem is the Lagrangian:

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

and $a^*(s) = (b_k, a_2^*(s), ..., a_n^*(s)) \in A^*(s)$ is a KKT point if there exists $\lambda \geq 0$ and $\mu \in \mathbb{R}$ such that:

$$\mu \nabla_a h(a^*(s)) \in \partial_a (f - \lambda g)(a^*(s), s))$$

Given the expression for $\nabla_a h(a^*(s), s)$, GMFCQ is defined as follows.

Definition 18 *A feasible point $a^*(s)$ satisfies the GMFCQ if there exists $y = (0, y_2, ..., y_n) \in \mathbb{R}^n$ such that:*

$$\forall \gamma \in \partial_a \overline{g}(a^*(s), s),\ \ \gamma \cdot \widetilde{y} < 0$$

By a straightforward application of Theorem 9 :

Proposition 19 *Under Clarke's hypothesis, and if the GMFCQ holds for every $a^*(s) \in A^*(s)$, then for any direction $x \in \mathbb{R}^m$:*

$$\liminf_{t \to 0^+} \frac{V(s+tx) - V(s)}{t} \geq \inf_{\lambda \in K(a^*(s),s)} \{ \min_{\theta \in \partial_s (f - \lambda g)(a^*(s),s)} \theta \cdot x \}$$

and:

$$\limsup_{t \to 0^+} \frac{V(s+tx) - V(s)}{t} \leq \sup_{\lambda \in K(a^*(s),s)} \{ \max_{\theta \in \partial_s (f - \lambda g)(a^*(s),s)} \theta \cdot x \}$$

V is locally Lipschitz with Clarke gradient:

$$\partial V(s) \subset co \left\{ \bigcup_{a^*(s) \in A^*(s)} \bigcup_{\lambda \in K^*(a^*(s),s)} \partial_s (f - \lambda g)(a^*(s), s) \right\}$$

Sharper characterization of the differentiability properties of V cannot rely on concavity since the domain is not convex.

Such result directly applies to the finite horizon (N periods) labor-leisure choice problem in which labor takes only the binary values $\{0, 1\}$. Thus we formulate a N period problem as,

$$V_n(k_n) = \max_{c_n, k_{n+1}, l_n} \{ u(c_n, 1 - l_n) + \beta V_{n+1}(k_{n+1}) \}$$

subject to

$$\begin{aligned} c_n + k_{n+1} - f(k_n, l_n) &\leq 0 \\ -c_n &\leq 0 \\ -k_{n+1} &\leq 0 \\ l_n(1 - l_n) &= 0 \end{aligned}$$

for all $n \leq N$, and $(k_{N+1}, V_{N+1}) = (0, 0)$. We also assume the following:

Assumption 4.2: Functions $u : \mathbb{R}^n_+ \times \mathbb{R}_+ \to \mathbb{R}$ and $f : \mathbb{R}^n_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ are locally Lipschitz with differential extensions on the boundaries, and strictly increasing in both arguments. The feasible set $D_n : \mathbb{R}^n_+ \rightrightarrows \mathbb{R}^n_+ \times \mathbb{R}^n_+ \times \mathbb{R}_+$ is nonempty-valued and uniformly compact near for all k_t.

Since u is strictly increasing the first inequality constraint always holds with an equality, and the associated Lagrangian is:

$$\begin{aligned} L_n(k_n) = {} & u(f(k_n, l_n) - k_{n+1}, 1 - l_n) + \beta V_{n+1}(k_{n+1}) + \lambda^1_n (f(k_n, l_n) - k_{n+1}) \\ & + \lambda^2_n k_{n+1} + \mu_n (l_n (1 - l_n)) \end{aligned}$$

Under assumption and if the GMFCQ holds for every optimal solution $y_n^*(k_n) \in Y_n^{;*}(k_n)$, a direct consequence of Proposition 19, is that for any direction $x \in R^n$:

$$\lim\inf_{t \to 0^+} \frac{V_n(k_n + tx) - V_n(k_n)}{t} \geq \inf_{\lambda_n \in K(y^*(k_n),k_n)} \{ \min_{\theta \in \partial_{k_n} L_n(y^*(k_n),k_n)} \theta \cdot x \}$$

and:

$$\lim\sup_{t \to 0^+} \frac{V_n(k_n + tx) - V_n(k_n)}{t} \leq \sup_{\lambda_n \in K(y^*(k_n),k_n)} \{ \max_{\theta \in \partial_{k_n} L_n(y^*(k_n),k_n)} \theta \cdot x \}$$

and V_n is locally Lipschitz for all $n = 1, .., N$.

4.3 Computing Markov equilibrium in growth models with nonsmooth nonconvex technologies

Recent work on optimal growth models with nonsmooth and nonconvex technologies have largely set aside the issue of existence of recursive and/or sequential equilibrium (see, for instance, Kamihigashi and Roy [26][27]), an issue we propose to address in this section by combining the envelope results of this paper with the lattice programming methodology of Mirman, Morand, and Reffett [33].

In these models, nonconvexities typically arise when a consumer's decisions depends on the aggregate or per capita state K in addition to its own individual state k. Our strategy is to impose sufficient conditions on the primitive data such that our search for a recursive equilibrium can be restricted to a large class of monotone functions. It is therefore related to the work of Hopenhayn and Prescott [25], extended in Morand, Reffett, and Tarafdar [36], on the existence of monotone controls on standard stochastic optimal growth models.

To simplify we work here in a deterministic setup, but, just as in Hopenhayn and Prescott [25] and central in Mirman, Morand and Reffett [33], we use a key result from lattice programming relating the supermodularity of a program's objective to the monotonicity (in (k, K)) of optimal decisions. It is in the proof of the supermodularity of the value function that our generalized envelope results play a critical role in models where the primitive data is not continuously differentiable.

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 entering period $t = \{0, 1, 2, ...\}$ with an individual stock of capital k_t and supplying inelastically one unit of time to firms. Common in the literature (e.g., Coleman [12], Greenwood and Huffman [19]), and consistent with recent work (e.g., [43] [26] [27]) we use a "reduced-form" production function $F(k, n, K, N)$ where k and n are, respectively, the firm's capital and labor inputs. Since $n = N = 1$ we use the notation $f(k, K, z) = F(k, 1, K, 1, z)$ 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, 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. Noting that the function $u(f(k,.) - y)$ is Lipschitz at any K satisfying $f(k, K) - y > 0$, and therefore almost everywhere differentiable at such points, we state this assumption as follows:

Assumption (iv) At points where $u(f(k,.) - y)$ *is differentiable, the function* $u'(f(k,K) - y)f_1(k,K)$ *is increasing in* K.

A consumer seeks to maximize utility given by:

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

given initial state $k_0 = K_0 > 0$. Period t choices of consumption and investement must satisfy:

$$c_t + k_{t+1} \leq f(k_t, K_t)$$

and the consumer is constrained to use a law of motion h to recursively compute the sequence $\{K_t\}$ of future per capita capital stocks as $K_{t+1} = h(K_t)$.

We require h to belong to the set $\boldsymbol{B}$ defined as:

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

and note that $(\boldsymbol{B}, \leq)$ is a complete lattice, where $\leq$ is the pointwise partial (see for instance Davey and Priestley [15]), and that the set of subsets of $\boldsymbol{B}$ endowed with the induced set order $\leq_a$ is also a complete lattice.

For a given $h \in \mathbf{B}$, by a standard argument there exists a unique value function V satisfying:

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

where $\Gamma(k,K) = \{y \in \mathbf{K}, 0 \leq y \leq f(k,K)\}$ is non-empty, compact and convex, and the continuity of f, implies that Γ is a continuous correspondence. Denote by Y^* the set of solutions to the above program, that is:

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

and by $\vee Y^*$ and $\wedge Y^*$ its greatest and least selection, respectively.

Interpreting a selection from the optimal correspondence Y^* as a "best response" to the law of motion h, we define a recursive equilibrium as a law of motion that is best response to itself, that is:

Definition 20 *A recursive equilibrium is a element* h^* *in* $\boldsymbol{B}$ *such that* $h^*(k) \in Y^*(k, k; h^*)$ *for all* $k \in \mathbb{K}$.

4.3.2 Existence of Recursive equilibrium

Because recursive equilibria are precisely the fixed point of mapping A defined as:

$$h \in \boldsymbol{B} \rightrightarrows Ah = \{h' \in \boldsymbol{B}, \forall k \in \mathbb{K}\ h'(k) \in Y^*(k, k; h')\},$$

our existence proof relies on the order-preserving properties of this mapping, as well as the related operator $\overline{A}$ defined as:

$$h \in \boldsymbol{B} \to \overline{A}h = \vee Y^*(k, k; h)$$

Both operators share the following property:

Lemma 21 *Both $A : (\boldsymbol{B}, \leq) \to (2^{\boldsymbol{B}} \backslash \varnothing, \leq_a)$ and $\overline{A} : (\boldsymbol{B}, \leq) \to (\boldsymbol{B}, \leq)$ are isotone mappings.*

Proof. The proof follows precisely the argument in Mirman, Morand and Reffett [32] (Lemma 7) except for the proof of supermodularity of V. We set aside this (important) detail for now and turn to the main result of this section. ■

Proposition 22 *The set of recursive equilibria is a non-empty complete lattice. The sequence $\{\overline{A}^n f\}$ pointwise converges to the greatest recursive equilibrium.*

Proof. See Mirman, Morand and Reffett [33] (Theorems 6 and 9). ■

4.3.3 Proof of supermodularity of the value function

We finally turn to the proof of supermodularity of V, for which our generalized envelope results are needed because the traditional lattice theoretic argument (Theorem 2.7.6 in Topkis [45]) on the preservation of supermodularity under maximization is not applicabel (or limited simply to Leontieff production functions, as in Hopenhayn and Prescott [25]). Neither does the argument in Mirman, Morand and Reffett [33] since it requires the primitive data to be at least continuously differentiable.

Proof. Given $V_0 = 0$ we prove by induction that each element of the sequence $\{V_n = T^{(n)}V_0\}$ is supermodular, so that V inherits that property as the pointwise limit of that sequence. Fix $h \in \boldsymbol{B}$ and assume then that V_n is Lipschitz and supermodular, and consider the Lipschitz program:

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

Given that h is increasing, $\beta V_n(y, h(K))$ has increasing differences in $(y; K)$ (on $\mathbb{R}^2$ supermodularity and increasing differences are equivalent properties) while $u(f(k, K) - y)$ has increasing differences in $(y; (k, K))$ since u is concave and f is increasing. Consequently, the objective in the above program has increasing differences in $(y; (k, K))$, while the choice correspondence $[0, f(k, K)]$ is strong set order ascending, hence the optimal choice set Y^*_{n+1} is strong set order ascending and $\vee Y^*_{n+1}$ and $\wedge Y^*_{n+1}$ are both isotone selections in (k, K) by Theorems 2.8.1 and 2.8.3 in Topkis [45]. Note that by the same argument both $f - \vee Y^*_{n+1}$ and $f - \wedge Y^*_{n+1}$ are also isotone selections in (k, K).

Inada conditions imply interiority of solutions (so that all multipliers are 0), so it follows from Theorem 9 that V_{n+1} is Lipschitz and that:

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

and:

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

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

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

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

and also, given any $\overline{k}\in\mathbb{K}$, for all $\widehat{k}\geq k\geq\overline{k}$:

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

Thus $\forall\widehat{k}\geq k\geq\overline{k}$ 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 above on any interval $[\overline{k},\widehat{k}]$. A symmetric argument holds for the direction $x<0$, thus proving that on any interval $[\overline{k},\widehat{k}]$ both Dinis are uniformly bounded. This implies that imply $k\to V_{n+1}(k,K)$ is absolutely continuous on $[\overline{k},\widehat{k}]$ for any $0<\overline{k}<\widehat{k}$.

This absolute continuity together with the properties that V_{n+1} is increasing and continuous in it first argument, imply that V_{n+1} is absolutely continuous on $\mathbb{K}=[0,\widehat{k}]$ (see Problem 37 in Royden [42]). By the fundamental theorem of integral calculus (Theorem 10, Chapter 6 in Royden [42]), $k\to V_{n+1}(k,K)$ is therefore almost everywhere differentiable and, for all $k\in\mathbb{K}$:

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

At the points where $k\to V_{n+1}(k,K)$ is differentiable, by definition both Dinis must coincide, hence:

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

Note that for any $K' > K$, where the derivative exists:

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

the first inequality resulting from the curvature assumption (Assumption (iv)), and the second from the isotonicity of $\wedge Y^*_{n+1}$ previously established. In light of (3) this proves the desired supermodularity of V_{n+1}. By induction, the supermodularity property is true for all n, and the sequence of functions $\{V_n\}_{n=0}^{\infty}$ is a collection of supermodular functions in (k, K). Its pointwise limit, precisely V, must therefore inherit that property. ■

4.4 Payoff equivalence with multidimensional types

In this last section, we establish the "Payoff Equivalence" result in mechanism design theory with multidimensional type and nonsmooth utility function. Initially derived in the context of single dimensional types and continuously differentiable utility by Myerson [37] payoff equivalence has recently being extended by Krishna and Maenner [28] to models with convex sets of types and convex utility functions, as well as models in which the allocation rule, the payment function and utility functions are globally Lipschitz and upper Clarke regular in all arguments, and utility functions are increasing.

In this section, we impose additional structure on the set of types to apply our directionally differentiable envelope theorems, but extend their result to a considerably weaker set of hypothesis. Specifically, we drop all but one of the Clarke regularity assumptions, we weaken the Lipschitz requirements to locally Lipschitz, and we do not require the monotonicity assumption on utilities.

We first recall the problem. Let $\mathbb{X}$ denote the set of social alternatives, where $\mathbb{X}$ is a subset of $\mathbb{R}^n$. There are $\mathbb{I}$ agents and each $i \in \mathbb{I}$ has a k-dimensional type $t_i \in \mathbb{T}_i \subset \mathbb{R}^k$. The set $\mathbb{T}= \mathbf{\Pi}_{j\in\mathbb{I}}\mathbb{T}_i$ is the product of the sets of types. Agent i's payoff function takes a quasilinear form, $u^i(x, t_i) - \mu^i$, where $x \in \mathbb{X}$ is the alternative chosen by the planner and μ_i is a monetary transfer to the planner. A mechanism is a pair (χ, μ) where $\chi : \mathbb{T}\to\mathbb{X}$ is the allocation rule and $\mu : \mathbb{T}\to\mathbb{R}^I$ is the payment rule. Thus, if agent i, reports type s_i, for all i in $\mathbb{I}$, the social planner chooses alternative $\chi(s_1, s_2, ..s_I)$ and the transfer payment of agent i is $\mu^i(s_1, s_2, ..s_I)$.

We maintain the following assumptions in this section:

Assumption 4.3.1: The type set $\mathbb{T}_i$ is open, separable, and connected.

Assumption 4.3.2: The types $t_i \in \mathbb{T}_i$ are independently distributed across agents according to a probability measure $m(t_i)$ over $\mathbb{T}_i$.

Assumption 4.3.3: For each i (a) *the utility function u^i is locally Lipschitz in (x, t_i) and Clarke regular in t_i; (b) the payment function μ^i is locally Lipschitz in reported types $(s_1, s_2, ..., s_I)$, for all i; (c) the allocation rule χ* is locally Lipschitz,

Assumption 4.3.4: The utility function u_i admit a differential extension on the boundary for the first argument for all i in $\mathbb{I}$.

Given a mechanism (χ, μ), the expected payoff to agent i (of type t_i) from reporting s_i when all other agents truthfully reveal their types is:

$$E_{t_{-i}} u^i(\chi(s_i, t_{-i}), t_i) - E_{t_{-i}} \mu^i(s_i, t_{-i})$$
$$= U^i(s_i, t_i) - \overline{\mu}^i(s_i)$$

The mechanism (χ, μ) is incentive compatible if all agents reveal their true type, that is if $\forall i \in \mathbb{I}$ and $\forall t_i \in \mathbb{T}_i$:

$$V_i(t_i) = U_i(t_i, t_i) - \overline{\mu}^i(t_i) = \max_{s_i \in T_i} \left\{ U^i(s_i, t_i) - \overline{\mu}^i(s_i) \right\} \tag{4}$$

Now we show the expected utility and expected payment functions preserves Lipschitzian properties of the utility and payment function respectively.

Proposition 23 *Under assumptions 4.3.1-4.3.4, $U^i(s_i, t_i)$ is locally Lipschitz in (s_i, t_i) and Clarke Regular in t_i and $\overline{\mu}^i(s_i)$ is locally Lipschitz in s_i for all i.*

Proof. Under Assumption 4.3.1-4.3.4, given $u^i(\chi(s_i, t_{-i}), t_i)$ is locally Lipschitz in (s_i, t_{-i}, t_i), since u^i and χ are also locally Lipschitz, and $\mathbb{T}_i$ is separable subset of $\mathbb{R}^k$, by Clarke [10] Theorem 1, implies that

$$E_{t_{-i}} u^i(\chi(s_i, t_{-i}), t_i) = \int u^i(\chi(s_i, t_{-i}), t_i) m(dt_{-i})$$

is locally Lipschitz in (s_i, t_i). Similarly, $\mu^i(s_i, t_{-i})$ is locally Lipschitz in $(s_i, t_{-i}) \in \mathbb{T}_i \times \mathbb{T}_{-i}$ (and, therefore $\overline{\mu}^i(s_i) = E_{t_{-i}} \mu^i(s_i, s)$ is locally Lipschitz in s_i). For any s_i, as $u^i(\chi(s_i, t_{-i}), t_i)$ is Clarke regular in t_i, by the Clarke's result, $\int u^i(\chi(s_i, t_{-i}), t_i) m(dt_{-i})$ is Clarke regular in t_i (as for any s_i, $u^i(\chi(s_i, t_{-i}), t_i)$ is Clarke regular in t_i). ■

The next proposition is an important technical result that follows from Krishna and Maenner [28] that we shall use in deriving the main result of this section.

Proposition 24 *If $W : C \to \mathbb{R}$ is a Lipschitz and Clarke regular function, C is a connected set in $\mathbb{R}^n$, and $\varsigma_w \in \partial W$ is a measurable selection, then for any smooth path α joining a to b in C, we have*

$$W(b) - W(a) = \int \varsigma_w d\alpha$$

Proof. See Krishna and Maenner ([28], Theorem 1). ■

For an application of Proposition 24 we require the function W to be lipschitz and not just locally Lipschitz on C. The next proposition show that any locally Lipschitz function on C is Lipschitz on a cpmpact subset of C.

Proposition 25 *Any locally Lipschitz function $W : C \to \mathbb{R}$, is Lipschitz on any compact set $D \subset C$*

Proof. f is locally Lipschitz on D, thus for all $a \in D$ there exist a open neighborhood of a, N_a such that

$$|f(a') - f(a'')| \leq K(N_a) \|a' - a''\|$$

for all $a', a'' \in N_a$. Since D is compact there will be finite number of open sets $N_1 N_2, ..., N_T$ such that $D \subset \cup_i N_i$ with

$$|f(a') - f(a'')| \leq K(N_i) \|a' - a''\|$$

for all $a', a'' \in N_i$ for all $i = 1, 2, ..., T$.

Any $a \in D$ *also satisfies* $a \in N_i$ for atleast some $i = 1, 2, ...T$. Since N_i is a open cover, there exist a $\delta \in \mathbb{R}$ such that if $\|a' - a''\| < \delta$ then $a', a'' \in N_i$ for at least some i. Thus, if $\|a' - a''\| < \delta$

$$|f(a') - f(a'')| \leq \max_i \{K(N_i)\} \|a' - a''\|$$

where $i = 1, 2, ..., T$.

Now suppose $\|a' - a''\| \geq \delta > 0$. Here,

$$\begin{aligned} |f(a') - f(a'')| &\leq \left| \max_{a \in D} f(a) - \min_{a \in D} f(a) \right| \\ &= \frac{|f(a^*) - f(a_*)|}{\delta} \delta \\ &\leq \frac{|f(a^*) - f(a_*)|}{\delta} \|a' - a''\| \end{aligned}$$

Here, $f(a^*)$ and $f(a_*)$ are the maxima and minima that f achives on the compact set D by Weierstrass Theorem.

For $K = \max\{K(N_1), K(N_2), ..., K(N_T), \frac{|f(a^*) - f(a_*)|}{\delta}\}$ the result follows. ■

Now we will provide alternative characterization of the utility and payment functions for "Payoff Equivalence" result by applying the directional differentiable envelope theorems for an unconstrained problem.

Theorem 26 *Under 4.3.1-4.3.4, for any incentive compatible mechanism* (χ, μ)*,* $V_i(t_i)$ *is locally Lipschitz and Clarke regular, with Gateaux derivative (directional derivative) given by*

$$V_i'(t_i; d) = \max_{s_i^*(t_i) \in S_i^*(t_i)} U_2^{i'}(s_i^*(t_i), t_i; d)$$

Further, $V_i(t_i)$ *is determined by* χ *up to an additive constant. Finally, for all* $t_i, t_i' \in \mathbb{T}_i$ *and any smooth path* α *joining* t_i *and* t_i' *in* $\mathbb{T}_i$

$$V_i(t_i) = V_i(t_i') + \int \varsigma_{u_2^i} d\alpha$$

for any measurable selection $\varsigma_{u_2^i} \in \partial U_2^i(t_i, t_i)$*.*

Proof. By Proposition 23, the objective of problem (4) is locally Lipschitz and Clarke regular in t_i. Thus, by Proposition 10, V_i is locally Lipschitz and Clarke regular with Gateaux derivative:

$$V_i'(t_i; d) = \max_{s_i^*(t_i) \in S_i^*(t_i)} U_2^{i'}(a_i^*(t_i), t_i; d)$$

The Gateaux derivative derivative and hence the Clarke gradient of V_i only depends on the expected utility from allocation and not on the expected payment.

Now for any $t_i, t_i' \in \mathbb{T}_i$ and any smooth path α joining t_i and t_i' in $\mathbb{T}_i$, let $D(t_i, t_i', \alpha)$ be any proper, open, bounded subset of $\mathbb{T}_i$ containing t_i, t_i' and α. Further, $\overline{D}(t_i, t_i', \alpha)$ is the closure of $D(t_i, t_i', \alpha)$. By Proposition 25 V_i is Lipschitz on $\overline{D}(t_i, t_i', \alpha)$ and hence $D(t_i, t_i', \alpha)$.

Further, for any incentive compatible mechanism (χ, μ), Clarke regularity of V_i imply

$$\begin{aligned} \partial V_i(t_i) &= \cup_{s_i^*(t_i) \in S_i^*(t_i)} \partial U_2^i(s_i^*(t_i), t_i) \\ &\supseteq \partial U_2^i(t_i, t_i) \end{aligned}$$

since, $t_i \in S_i^*(t_i)$. As a result, on the open set $D(t_i, t_i', \alpha)$, Proposition 24 implies that for any measurable selection $\varsigma_{u_2^i} \in \partial U_2^i(t_i, t_i)$

$$V_i(t_i) = V_i(t_i') + \int \varsigma_{u_2^i} d\alpha$$

This shows that V_i is determined by $\varsigma_{u_2^i} \in \partial U_2^i(t_i, t_i)$ and hence by the allocation χ (and not the payment function μ) up to an additive constant. ■

5 Appendix: Mathematical Tools

5.1 Derivatives and subgradients

Given an open set $\Omega \subset \mathbb{R}^n$, the function $f : \Omega \to \mathbb{R}^m$ is said to be Lipschitz at $x \in \Omega$ if $\exists k > 0$ and $\exists \delta > \delta$ such that:

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

where $B(x)$ is the open ball of radius 1 centered on x. If the modulus can be chosen independently of x on an open subset of Ω, f is said to be globally Lipschitz on that subset.

If f is Lipschitz at x then the upper and lower Dini derivatives, respectively defined as the functions:

$$\begin{aligned} d &\longmapsto D^+ f(x; d) = \limsup_{t \to 0^+} \frac{f(x + td) - f(x)}{t} \quad \text{and:} \\ d &\longmapsto D_+ f(x; d) = \liminf_{t \to 0^+} \frac{f(x + td) - f(x)}{t} \end{aligned}$$

always exist.

Function f is said to be (a) Gateaux (directionally) differentiable at x if both Dini derivatives coincide for all d, in which case the Gateaux derivative is:

$$f'(x;d) = \lim_{t\to 0^+} \frac{f(x+td)-f(x)}{t},$$

and (b) differentiable at x if it is Gateaux differentiable at x and if $f'(x;d) = \nabla f(x)\cdot d$. Note that the function $x \to |x|$ is directionally differentiable but not differentiable at 0. However, by Rademacher's theorem, 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.

Lipschitz functions also have the property that the upper and lower Clarke derivatives, respectively defined as the functions:

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

also always exist. Note that:

$$f^{-o}(x;d) \leq -f^{-o}(x;-d) = f^o(x;d) \tag{5}$$

If f is Lipschitz and Gateaux differentiable at x, then clearly:

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

Gateaux differentiable Lipschitz functions (at some x) are said to be upper Clarke regular at x if $f^o(x;d) = f'(x;d)$ (and lower Clarke regular at x if. $f^{-o}(x;d) = f'(x;d)$). Function f is said to be strictly differentiability at x if both upper and lower Clarke derivatives coincide. In finite dimensional spaces, strict differentiability and continuous differentiability are equivalent.

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

$$\partial f(x) = \overline{co}\,\{\lim \nabla f(x_i) : x_i \to x, x_i \notin \Theta, x_i \notin \Omega_f\}$$

where $\overline{co}$ denotes the convex hull[5], 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 [11] (Proposition 2.1.5) shows that $x \rightrightarrows \partial f(.)$ is an upper hemicontinuous correspondence. Clarke [11] (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.

It is important to note the following important properties of convex functions $f:\Omega \to \mathbb{R}^m$:

[5] In the formula, either co or $\overline{co}$ will do since we work with finite dimensional spaces.

Lemma 27 *Suppose $f : \Omega \to \mathbb{R}^m$ is convex. Then:*

(i) f is Lipschitz at every $x \in \Omega$.

(ii) f is upper Clarke regular at every $x \in \Omega$.

(iii) the Clarke gradient of f at x coincides with the subgradient of convex analysis, i.e. the set of $p \in M_{m\times n}$ satisfying $\forall d, p \cdot d \leq f(x_0 + d) - f(x_0)$.

(iv). If f is also differentiable at x, then f is continuously differentiable at x.

Proof. (i), (ii) and (iii) are well-known. We prove now that differentiability together with upper (or lower) Clarke regularity implies continuous differentiability. Upper Clarke regularity and differentiability imply that:

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

hence, using (5) above:

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

which proves that f is strictly differentiable (thus continuously differentiable) at x. A similar argument clearly applies if f is lower (rather than upper) Clarke regular. ■

5.2 Properties of Correspondences

We work in metric spaces, so we can state topological properties of correspondences exclusively in terms of sequences.

Definition 28 *Given $A \subset R^n$ and $S \subset R^m$, a non-empty 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 $D(s) = \{a \in A,\ g_i(a, s) \leq 0,\ i = 1, ..., p\}$, in which the g_i are locally Lipschitz (and thus continuous), is necessarily closed at s. The same property holds true in the presence of locally Lipschitz equality constraints.

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

Definition 29 *A non-empty valued correspondence D is said to be uniformly compact near s if there exists a neighborhood S' of s such that $cl \left[\cup_{s' \in S'} D(s)\right]$ is compact.*

We note the result in Hogan [24] 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 upperhemicontinuity 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 30 *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. ■

5.3 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 [46], 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 $\forall (a, b) \in A_1 \times A_1$, $a \wedge b \in A_2$ 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., $>$) $f(x) + f(y)$. An important property of the class of supermodular functions is they are closed under pointwise limits. (Topkis, [45], 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 non-decreasing 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. L. Mirman, and W. Perkins. 1991, One-sector nonclassical optimal growth: optimality conditions and comparative dynamics. *International Economic Review,* 32, 625-644.

[2] Amir, R. 1996. Sensitivity analysis of multisector optimal economic dynamics. *Journal of Mathematical Economics, 25, 123-141*

[3] E. Antoniadou, Lattice programming and economic optimization, Ph.D. Dissertation, Stanford University, 1996.

[4] E. Antoniadou, Comparative statics for the consumers problem, *Econ. Theory* 31 (2007) 189-203.

[5] K. Askri, C. LeVan, Differentiability of the value function of nonclassical optimal growth models, *Journal of Optimization Theory and Applications* 97 (1998), 591-604.

[6] A. Auslender, Differentiable stability in non convex and non differentiable programming, *Math. Programming Stud.* 10 (1979) 29-41.

[7] Benveniste, L. and J. Scheinkman, On the differentiability of the value function in dynamic models of economics, Econometrica 47 (1979) 727-32.

[8] Bonnans, J. Frederic and A. Shapiro. Perturbation Analysis of Optimization Problems. Springer-Verlag, New York, Inc. 2000.

[9] J.M. Bonnisseau, and C. LeVan. On the subdifferential of the value function in economic optimization problems. *Journal of Mathematical Economics.* (1996), 55-73.

[10] F. Clarke, Generalized gradients of Lipschitz functionals, Advances in Mathematics 40 (1981) 52-67.

[11] F. Clarke, *Optimization and Nonsmooth Analysis*. SIAM 1983.

[12] Coleman, W. J., II. 2000. Uniqueness of an equilibrium in infinite-horizon economies subject to taxes and externalities , *Journal of Economic Theory* **95** (2000), 71-78.

[13] J. Danskin, *The Theory of Max-Min*, Springer, 1967.

[14] Dugundji, J and A. Granas. *Fixed Point Theory.* Warsaw Scientific Press. 1982

[15] Davey, B and H. Priestley. *Introduction to Lattices and Order, Second Edition.* Cambridge Press, 2002.

[16] G. Fontanie, Subdifferential stability in Lipschitz programming, MS, Operations Research and Systems Analysis Center, University of North Carolina1980.

[17] J. Gauvin. A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming, *Mathematical Programming* 12 (1979) 136-138.

[18] J. Gauvin and F. Dubeau. Differentiable properties of the marginal function in mathematical programming. Mathematical Programming Study 19 (1982) 101-119.

[19] Greenwood, J. and G. Huffman. On the existence of nonoptimal equilibria in dynamic stochastic economies, *Journal of Economic Theory,* (1995), 65, 611-623

[20] E. G. Gol'stein, *Theory of Convex Programming*, Translations of Mathematical Monographs, Providence: American Mathematical Society, 1972.

[21] Grinold, R. 1970. Lagrangian subgradient. *Management Science*. 185-188.

[22] Hinderer, K. Lipschitz Continuity of Value Functions in Markovian Decision Processes. *Math. Meth. Oper. Res.*, (2005), 62, 3–22.

[23] J.B. Hiriart Urruty, Refinement of necessary optimality conditions in nondifferentiable programming, *Applied Mathematical Optimization* 5 (1979) 62-82.

[24] W. Hogan, Point-to-set maps in mathematical programming, *SIAM Review* 15 (1973) 591-603.

[25] H. Hopenhayn, and E. Prescott. 1992. Stochastic monotonicity and stationary distributions in dynamic economies. *Econometrica,* 60, 1387-1406.

[26] T. Kamihigashi, and S. Roy. Dynamic optimization with a nonsmooth, nonconvex technology: the case of a linear objective function. *Economic Theory,* (2006) 29:325-340.

[27] T. Kamihigashi, and S. Roy. A nonsmooth, nonconvex model of optimal growth. *Journal of Economic Theory,* (2007) 132: 435-460.

[28] V. Krishna, E. Maenner, Convex potentials with an application to mechanism design, Econometrica 69 (2001) 1113-1119.

[29] J. Kyparisis, On the uniqueness of Kuhn-Tucker Multiplier in nonlinear programming, *Mathematical Programming* 32 (1985) 242-246.

[30] Laraki, R. and W. Sudderth. The preservation of continuity and Lipschitz continuity of optimal reward operators. *Mathematics of Operations Research,* (2004), 29, 672-685.

[31] P. Milgrom I. Segal, Envelope theorems for arbitrary choices, *Econometrica* 70 (2002) 583-601.

[32] L. Mirman, R. Ruble, Lattice theory and the consumer's problem. *Mathematics of Operations Research* 33 2008: 301-314.

[33] L. Mirman, O. Morand, K. Reffett. 2008. A qualitative approach to Markovian equilibrium in infinite horizon economies with capital. *Journal of Economic Theory,* 2008: 75-98.

[34] L. Montrucchio. Lipschitz continuous policy functions for strongly concave optimization problems. *Journal of Mathematical Economics,* (1987), 16, 259-273.

[35] L. Montrucchio. Thompson metric, contractive property, and differentiability of the policy function. *Journal of Economic Behavior and Organization.* (1998), 33, 449-466.

[36] O. Morand, K. Reffett, S. Tarafdar, A nonsmooth approach to Envelope Theorems MS, Arizona State University, 2015.

[37] R. Myerson, Optimal Auction Design, Mathematics of Operations Research 6 (1981) 58-73.

[38] V.H. Nguyen, J.J. Strodiot, R. Mifflin, On conditions to have bounded multipliers in locally Lipschitz programming, *Mathematical Programming* 18 (1980) 100-106.

[39] J. Quah, The comparative statics of constrained optimization problems, *Econometrica* 75 (2007), 401-431.

[40] J. Rincon-Zapatero, M. Santos, Differentiability of the value function without interiority assumptions *Journal of Econ. Theory* 144 (2009), 1948-1964.

[41] R.T. Rockafellar, R J-B Wets, *Variational Analysis* Springer.

[42] H. Royden. *Real Analyis.* MacMillan Press. 1968

[43]

[44] S. Tarafdar, Optimization in economies with nonconvexities, Ph.D. Dissertation, Arizona State University, 2010.

[45] D. Topkis, *Supermodularity and Complementarity*, Princeton 1998.

[46] A. Veinott, *Lattice programming: qualitative optimization and equilibria*, MS, Stanford 1992.