# A Nonsmooth Approach to Envelope Theorems*

Olivier Morand† Kevin Reffett‡ Suchismita Tarafdar§

First Draft: August 2009
This Draft: February, 2013

**Abstract**

We develop a nonsmooth approach to envelope theorems that applies across a broad class of parameterized nonlinear optimization problems that arise typically in economic applications. Our methods include envelope theorems for both the convex and nonconvex case, and allow for noninterior solutions. We develop conditions under which the value function has (i) differential local Lipschitz bounds, as well as (ii) Clarke differentiable, (iii) directionally differentiable, and (iv) once-continuously differentiable ($C^1$). We relate our results to the existing literature, and present applications of the results to monotone controls in stochastic growth models with Markov shocks, constrained lattice programming, and payoff equivalence in mechanism design.

*We thank Rabah Amir, Bob Becker, Madhav Chandrasekher, Bernard Cornet, Manjira Datta, Amanda Friedenberg, Karl Hinderer, Rida Laraki, Cuong LeVan, Len Mirman, Debasis Mishra, Manuel Santos, Arunava Sen, Tridib Sharma, Yiannis Vailakis, and Lukasz Woźny for helpful discussions during the writing of this paper. Special thanks go to Andrzej Nowak for directing us to the work of Hinderer, Laraki, and Sudderth. Kevin Reffett thanks the Centre d'Economie de la Sorbonne (CES) and the Paris School of Economics for arranging his visits during the Spring terms of both 2008 and 2009, as well as financial support from the 2010 Dean's Summer Grant program at ASU. The usual cavaets apply.

†Department of Economics, University of Connecticut
‡Department of Economics, WP Carey School of Business, Arizona State University
§Planning Unit, Indian Statistical Institute (ISI), Delhi

# 1 Introduction

Since the work of Viner [49] and Samuelson [43], the envelope theorem has been a standard tool in economic analysis. Aside from its extensive use in optimization theory and general equilibrium (e.g., duality and pricing), it has also become a key ingredient in arguments that characterize the structure of equilibrium in many fields in economics including dynamic contract theory, public finance, economic growth, consumer and producer theory, game theory, dynamic programming, and lattice programming. In the "classical" case, the envelope theorem is simply a continuous derivative of a value function in a parameter. In the original work on the classical envelope theorem, sufficient conditions for its existence required a great of mathematical structure, including strong convexity assumptions on the primitive data, strong interiority assumptions, as well as the continuous differentiability of objectives and constraints (e.g., see the results of Samuelson [43], Rockafellar [42], Mirman and Zilcha [33], Beneveniste and Scheinkman [7]).

Although some of these assumptions have been relaxed (e.g., conditions on the structure of the choice sets as in unconstained problems such as studied in Milgrom and Segal [31]), and the interiority of optimal solutions as in Rincon-Zapatero and Santos [41])), none of this work has considered the case of envelope theorems in economic models with nonconvexities and constraints.[1] As in a great deal of recent work in economics, nonconvexities have played an important role (e.g., numerous paper in macroeconomics, new trade theory, labor economics, economic growth, Mirrleesian public finance, and industrial organization, among others), the case with nonconvexities per the existence of envelope theorems is an interesting one to study. This is especially true as much of this recent work has taken place in the context of dynamic economies where dynamic programming arguments are being invoked, where the lack of useful envelope theorems can become a critical impediment in the development rigorous approaches to characterizing and constructing optimal solutions for agent decision problems.[2]

[1]Its worth noting the proof in Mirman and Zilcha ([33], Lemma 1) works for the noninterior solution case if utility function is $C^1$ at zero consumption. In that case, their result is a special case of main theorems in Rincon-Zapatero and Santos [41]). Also, the results reported in Milgrom and Segal [31] for envelopes for the case of arbitrary choice sets are closely related (yet distinct from) an important theorem in Clarke (e.g., [9], Theorem 2.1).

[2]In general, the value function is nondifferentiable in optimization problems with nonconvexities even if all the primitive data of the optimization problem are infinitely smooth.

In this paper, we provide a catalog of envelope theorems (often in a "generalized" sense) that are appropriate in economic models with nonconvexities. We build our approach in such a manner that it includes both the convex and nonconvex case within a single mathematical framework. Methodologically, we take a "nonsmooth" approach to the problem, developing a theory from the perspective of generalized envelope theorems which seem most useful in optimization problems with nonconvexities, yet reduce to the classical envelope theorems become a special case of the theory in some convex cases. More specifically, we begin with a description of a collection of parameterized "local Lipschitzian" optimization problems in finite dimensional Euclidean spaces. This class of problems includes in its canonical form many optimization problems that have arisen recently many literatures in economics.[3] To allow for dynamic programming applications, we allow objective functions to be locally Lipschitzian, but not necessarily differentiable. As in most interesting applications constraints are once-continuously differentiable constraints (even in dynamic programming), we consider the case of smooth constraints.[4]

For our problems, we develop conditions under which progressively stronger conditions on the structural properties of the primitives lead to progressively sharper characterizations of the differentiable properties of the value function. In the end, we provide conditions under which value functions admit differential bounds, are Clarke differentiable, directionally differentiable, and once-continuously differentiable ($C^1$). We also show the relationship between existing results in the large literature on envelope theorems for convex problems, and our results for the nonconvex case. We conclude by presenting important applications of our results. In particular, we show how to use our methods to obtain monotone controls in stochastic growth models with Markov shocks, standard monotone comparative statics in constrained lat-

See, for example, Amir, Mirman and Perkins [1] for a nice discussion of this fact in the context of a one-sector growth model.

[3] One class of problems with nonconvexities that we do not consider in this paper is so-called "mixed-integer programming" problems, where some choice variable, for example, is discrete. We dedicate a separate paper to that case, although many if the ideas in this paper do extend to that setting (e.g., the results related to Clarke [9] and Milgrom and Segal [31]). See Morand, Reffett, and Tarafdar [35] for discussion.

[4] The nonsmooth constraint case is also interesting. For example, in incentive constrained dynamic programming, one cannot expect that value function to be $C^1$ (although conditions for this case are given in Rincon-Zapatero and Santos [41], for example). Our problems with nonsmooth constraints are being studied in Morand, Reffett, and Tarafdar [36].

tice programming via a new duality approach, as well to study important problems in implementation and mechanism design.

When considering the question of the envelope theorem in the nonconvex case, at least three important issues arise immediately. First, as continuously differentiable envelope theorems are not expected for the nonconvex case, one would like to propose an appropriate alternative notion of a “generalized” envelope that can also be easily be applied in applications as a "substitute" for the classical envelope. Second, as the nonconvexities are introduced in optimization problems in the applied literature to study "extensive" vs. "intensive" margins (e.g., in macroeconomics and new trade theory), the proposed approach must work in settings with both equality and inequality constraints, as well as situation with optimal solutions are not necessarily noninterior.[5] Finally, as the Slater condition from convex analysis is known to be inappropriate in nonconvex settings, we need to identify new and useful classes of constraint qualifications that are both (i) strong enough to allow for simple (and often exact) calculations of the generalized envelopes (as well as construct bounds for local Lipschitz constants of the value function (the latter is often all that is needed in applications)[6], while (ii) not being too strong to be appropriate for the notion of "differentiability" under consideration. Our approach in the paper addresses all three of these questions.

# 2 Existing Literature

To place our work in the context of the existing literature, it proves useful to trace the development of the classical envelope theorem to the pioneering work of Samuelson [43]. In his book *Foundations of Economic Analysis,* Samuelson discusses conditions under which $C^1$ envelopes exist in optimization problems with smooth objectives, feasible sets that (in effect) do not vary with the parameter, with interior optimal solutions. Generalizations of his classical envelope theorem have been numerous, with perhaps the most notably contributions found in Danskin [14], Rockafellar [42], Gol'stein [21],

[5] In this literature, it appears an "extensive" margin is simply a decision to move from the boundary to an interior optimal solution; an "intensive" margin measures how much an interior optimal solution changes.

[6] In dynamic programming, differential bounds are sufficient to establish the requisite equicontinuity needed to close arguments are infinity, or to extend Santos and Vigo's results [44] on uniform errors bounds via discretization for numerical approximations to dynamic programming problems with nonconvexities.

Mirman and Zilcha [33], and Beneveniste and Scheinkman [7]. In many of these discussions, interiority of unique optimal solutions plays a key role, so these results are difficult to apply to programming problems with general constraints and/or noninterior optimal solutions (even for convex programs).

An important extension of this classical literature appeared recently in an important paper by Milgrom and Segal [31]. In this paper, the authors produce a number of new classical and directional envelope theorems under different sets of sufficient conditions. In the most recent contribution on the classical envelope theorem, Rincon-Zapatero and Santos [41] show how to extend the classical $C^1$ envelope theorem to infinite horizon dynamic programming settings for the case where optimal solutions are allowed to be noninterior. This extension is also important in very general problems considering envelope theorems (even in single stage programming problems). To the best of our knowledge, Rincon-Zapatero and Santos [41] have the most general classical envelope theorem in the existing literature, and also show how it can be applied to some important problems in the literature on incentive constrained dynamic programming (e.g., recursive saddlepoint methods). In all of these papers, the authors assume standard convexity conditions, often with interiority, and/or smooth convex concave objectives with strictly quasiconcave constraints. When constraint qualifications are used, they use versions of the Slater condition (e.g., often very strong versions of the Slater condition such as Linear independence constraint qualification (LICQ) in Rincon-Zapatero and Santos [41]), which along with assumptions about the existence of differential extensions over the boundary plays a key role in the case of noninterior solutions.

In essence, the objective of this paper is to show that many of the results obtained in the existing literature for the convex case can be generalized (in an suitably defined manner) to the nonconvex case. First, as in Rincon-Zapatero and Santos [41], we develop envelope theorems that allow for boundary solutions. What is different is that in the convex case, we weaken their constraint qualification, yet still obtain classical $C^1$ envelopes. Further, for our results in the nonconvex case, we allow noninterior solutions.[7] We show that relaxing the so called "linear independence" constraint qualification (LICQ) that is used in their work in all cases matters, and we

[7] For example, in Askri and LeVan [5], the authors construct nonsmooth envelopes in the context of dynamic programming. Their proofs are based essentially on Clarke's version of Danskin's theorem (e.g., Clarke [9], Theorem 2.1). That requires interiority.

give an simple example of a consumers problem where the value function is smooth in income, the LICQ does not hold, yet our results actually do apply. As in Milgrom and Segal [31], our calculations of directional and generalized envelopes for the nonconvex are basically the same as in the convex case they study (i.e., they simply require standard "min" and "max" calculations across optimal solutions, and are read off envelopes of the Lagrangian dual program). So our results nicely complement both of these papers. Further, we also calculate generalized Clarke gradients as well as differential bounds for the value function under very weak constraint qualifications (thus generalizing the subdifferential results that are implied in all of these papers for the convex case).

Finally, from a methodological perspective, our approach is closely related to the work on nonsmooth envelopes of Gauvin and Tolle [20] and Gauvin and Dubeau [19], among others. One key difference between this body of work, and ours, is that in this literature, objectives are assumed to be $C^1$. For many applications (e.g., dynamic programming in economic models with nonconvexities), this assumption is difficult to check. So we relax this assumption in many of our results. More specifically, building on the work of Auslender [6] and Fontanie [18], we are able to improve upon this literature by obtaining similar results obtained for the $C^1$ objective case for the locally Lipschitz objective case, but typically under a weaker constraint qualifications than used in much of this work (e.g., we do not use the LICQ). Again, by examples, we show that this extension can be important in applications.

The remainder of the paper is organized as follows. In the next section, we describe the class of parameterized optimization problems under study, and define the relevant constraint qualifications used in the paper and mention the Lagrange multiplier rule used in many of our proofs. The duality scheme we use throughout the paper is standard in the literature, and we take a standard perturbation approach to the problem, allowing for both "right-hand" and "left-hand" side perturbations in our programs. In Section four, we present a catalog of generalized (and classical) envelope theorems under different conditions on the primitive data of our optimization problems. The final section of the paper concludes with three important applications of our results. In particular, we use our results to provide (i) new conditions for monotone controls in stochastic growth models with Markov shocks (first considered in Hopenhayn and Prescott [24]), (ii) a new duality method for characterizing normal demand in the consumer problem (and discuss how the method can be used for more general constrained lattice programming problems as studied in Quah [40]), and (iii) new results for a mechanism design problem studied in Krishna and Maenner ([27]). The appendix presents definitions of the important mathematical terms used in the paper.

# 3 Lipschitz Optimization Problems

Throughout the paper, we consider the following family of *Parameterized Lipschitz Optimization Problems*. Let $A$ be space of actions (or controls), $S$ the parameter space, and $f : A \times S \to \mathbf{R}$ the objective function. We seek to characterize envelope theorems in the following parameterized program:

$$V(s) = \max_{a \in D(s)} f(a, s) \tag{3.0.1}$$

where the feasible correspondence is given by

$$D(s) = \{a | g^i(a, s) \leq 0, \quad i = 1, ........, p, \; h^j(a, s) = 0, \quad j = 1, ........, q\}$$

and the optimal solution correspondence is:

$$A^*(s) = \arg \max_{a \in D(s)} f(a, s)$$

We shall often refer to program (3.0.1) as the "primal" problem.

We will maintain some baseline assumptions throughout the paper:

*Assumption 1:*

*(a) The metric spaces $(A, \rho_A)$ and $(S, \rho_S)$ are each convex in* $\mathbf{R}^n$ and $\mathbf{R}^m$*, respectively,*

*(b) the constraints $g^i$, $i = 1, .., p$ and $h^j = 0$, $j = 1, .., q$ are jointly $C^1$, and $n \geq q$,*

*(c) the objective function $f : A \times S \to \mathbf{R}$ is continuous in $(a, s)$,*

*(d) the feasible correspondence $D : S \rightrightarrows A$ is nonempty, continuous and compact-valued.*

*Assumption 2: The objective $f(a, s)$ and the constraints $g(a, s)$* and $h(a, s)$ *both admit differential extensions over the boundary of $A \times S$.*

Assumption 2 is needed to obtain many of the results in the paper, as we shall allow for optimal solutions to exist on the boundary (as, for example, in Rincon-Zapatero and Santos [41]).

The Lagrangian duality scheme we use is standard, albeit we study its use in the context of nonlinear duality setting. Of course, the cost of relaxing the convexity assumptions is immediate, as we will not, in general, have optimal solutions that satisfy strong duality (e.g., zero duality gap, saddle-point stability, etc. Therefore, unlike the convex case, constructing both a necessary and sufficient first order theory for optimal solutions will be in general impossible. Fortunately, as we shall show, this fact does not prevent us from providing mild conditions under which sharp characterizations of simple nonsmooth envelope theorems for the primal problem can be read off linearizations of the Lagrangian dual at its optimum.

Let the dual variables (or *Karash-Kuhn-Tucker (KKT) multipliers*) be given as a vector $\sigma = (\lambda, \mu) \in \Omega \subset \mathbf{R}^n_+$. We conjugate the problem (3.0.1) with a classical Lagrangian duality scheme as follows:

$$L(a, \lambda, \mu; s) = \begin{cases} f(a, s) - \lambda^T g(a, s) - \mu^T h(a, s) & \text{if } a \in A, \sigma \in \Omega \\ -\infty, & \text{if } a \notin A, \sigma \in \Omega \\ \infty, & \text{otherwise} \end{cases}$$

where

$$g(a, s) = [g^1(a, s), ..., g^p(a, s)]^T; \; h(a, s) = [h^1(a, s), ..., h^q(a, s)]^T$$

The dual program to (3.0.1) is, therefore, the following:

$$V^D(s) = \inf_{(\lambda,\mu)\in\Omega} \sup_{a\in A} L(a, \lambda, \mu; s) \tag{3.0.2}$$

We say a point $a \in D(s)$ is a *Karush-Kuhn-Tucker (KKT) point* for the dual program (3.0.2) if there exist vectors $\lambda_i \geq 0$, $i = 1, ..., p$ and $\mu \geq 0$, $j = 1, ..., q$ such that:

$$0 \in \partial_a f(a, s) - \left(\lambda^T g_1 + \mu^T h_1\right)(a, s)$$

We shall denote the set of these KKT points associated with any feasible point $a \in D(s)$ by $K(a, s) \subset \Omega$.

In constrained optimization, it is well-known that constraint qualifications play a critical role in guaranteeing that dual programs are useful in characterizing the structure of the value function and of the optimal solutions in the primal problem. For example, in the convex case, the Slater condition is critical for proving strong duality between the dual and primal programs[8]. More specifically, constraint qualifications are needed to guarantee that (i) optimal solutions are KKT points, and (ii) the feasible region around such points will not vanish under local perturbations of the parameters. As shall become clear in the sequel, in a nonconvex setting, the question of what is an "appropriate" constraint qualification is more complicated, and will play a critical role in the *form* of the envelope theorem that can be constructed (e.g., differential bounds, Clarke envelopes, directional derivatives, etc.).

We now take a moment to discuss a few important constraint qualifications that have played a key in the literature on smooth nonconvex optimization. The strongest constraint qualification is the so-called "linear independence constraint qualification" (henceforth, LICQ), which has been the focus of the recent work of Gauvin and Dubeau [19] and Rincon-Zapatero and Santos [41], for example. This constraint qualification is defined as follows:

**Definition 1** *A feasible point $a \in D(s)$ satisfies the Linear Independence Constraint Qualifier (LICQ) if, the following vectors are linearly independent,*

$$g_1^i(a,s),\ i \in I, \quad h_1^j(a,s),\ j = 1,...,q$$

*where $I = \{i : g^i(a,s) = 0\}$*

In this paper, we only mention LICQ for the sake of comparing our results with those in the existing literature. For our purposes, it will turn out that LICQ is too strong, and it will play no role in our arguments.

In our analysis, we shall use one of two other constraint qualifications, depending on the nature of the envelope theorem sought. For example, if we seek to only bound Dini directional derivatives and Clarke gradients, we will appeal to the Mangasarian-Fromowitz Constraint Qualification (MFCQ), which is given by:

**Definition 2** *A feasible point $a \in D(s)$ satisfies the Mangasarian-Fromovitz Constraint Qualifier (MFCQ) if:*

*(i) the following vectors are linearly independent,*

$$h_1^j(a,s),\ j = 1,...,q$$

[8] In concave problems, a constraint system satisfies a *Slater condition* if there exists $a \in D(s)$ such that $h(a,s) = 0$ and $g_i(a,s) < 0$ for all constraints, as Milgrom and Segal [31].

*(ii) there exists a* $y \in \mathbf{R}^n$ *such that,*

$$g_1^i(a,s)y < 0,\ i \in I;\ \ h_1^j(a,s)y = 0, \quad j = 1, ..., q$$

*where* $I = \{i : g^i(a,s) = 0\}$.

It is well-known that the MFCQ implies the Slater condition of the convex case. One key limitation of MFCQ is that at any optimal solution the MFCQ can only guarantee the existence of a compact set of KKT multipliers, but not uniqueness of the KKT multipliers.[9] This makes obtaining directional derivatives and sharp characterizations of the generalized gradients very difficult. To obtain unique KKT multipliers for each optimal solution, we impose as slightly stronger constraint qualification than MFCQ (along with some standard regularity conditions on $D(s)$, e.g., uniform compactness), namely the Strict Mangasarian-Fromowitz Constraint Qualification (henceforth, SMFCQ), which is defined as follows:

**Definition 3** *A feasible point* $a \in D(s)$ *satisfies the Strict Mangasarian-Fromovitz Constraint Qualifier (SMFCQ) if:*
*(i) the following vectors are linearly independent,*

$$g_1^i(a,s),\ i \in I_b;\ \ h_1^j(a,s),\ j = 1, ..., q$$

*(ii) there exist* $y \in \mathbf{R}^n$ *such that,*

$$\begin{gathered} g_1^i(a,s)y < 0,\ i \in I_s;\ \ g_1^i(a,s)y = 0,\ i \in I_b \\ h_1^j(a,s)y = 0, \quad j = 1, ..., q \end{gathered}$$

*where* $I_b = \{i \in I : \lambda_i > 0\}$, $I_s = \{i \in I : \lambda_i = 0\}$ *and* $I = \{i : g^i(a,s) = 0\}$.

Clearly, SMFCQ implies MFCQ, but not the converse. It is also important to note that LICQ implies SMFCQ, but not the converse, and we shall provide examples of this critical difference later in the paper (e.g., when discussing $C^1$ differentiability of the value function with noninterior optimal solutions).

[9] In the case of the program (3.0.1) with smooth primitive data, MFCQ with some additional regularity conditions on $D(s)$ is in fact *equivalent* to a compact set of KKT multipliers. See Gauvin ([20]) and Gauvin and Dubeau [19]).

## 3.1 Lagrange Multiplier Rule

With these definitions in mind, we now turn to the question of the existence of Lagrange multiplier rules for our classical Lagrangian duality scheme . This discussion is needed to understand the nature of our first order characterization of optimal solutions in the dual program (3.0.2). The Lagrange multiplier rule we shall use is standard, and proven in Clarke ( [11] Theorem 6.1.1), and we state this result as a proposition:

**Proposition 4** *Under Assumptions 1 and 2, if $f$ is locally Lipschitz at the optimal solution $(a^*(s), s)$, and problem (3.0.1) satisfies MFCQ, then, there exist $\lambda \in \mathbf{R}^p$ and $\mu \in \mathbf{R}^p$ such that*

*(a)* $\lambda^i \geq 0$
*(b)* $\lambda^i g^i(a^*(s), s) = 0$
*(c)* $0 \in \ \partial_a f(a^*(s), s) - \left(\lambda^T g_1 + \mu^T h_1\right)(a^*(s), s)$

We remark that under one additional condition in Proposition 4, the first order condition in result (c) can be replaced by the expression

$$f_a^{-o}(a^*(s), s; x_a) - \left(\lambda^T g_1 - \mu^T h_1\right)(a^*(s), s) \cdot x_a \leq 0$$

for each direction $x_a \in \mathbf{R}^n$. We shall use this version in many of our arguments in the sequel, and it is stated in the next proposition:

**Proposition 5** *Under Assumptions 1 and 2 hold, if $f$ is directionally differentiable at an optimal solution $(a^*(s), s)$, and problem (3.0.1) satisfies MFCQ, then, there exist a multiplier $\lambda \in \mathbf{R}^p$ and $\mu \in \mathbf{R}^p$ such that for each direction $x_a \in \mathbf{R}^n$*

$$f_a'(a^*(s), s; x_a) - \left(\lambda^T g_1 - \mu^T h_1\right)(a^*(s), s) \cdot x_a \leq 0$$

*If, additionally, $f$ is $C^1$ in $a$, then,*

$$0 = \left( \ f_1 - \lambda^T g_1 + \mu^T h_1\right)(a^*(s), s)$$

The last condition of the proposition is the "standard" first order condition for the dual program (i.e., the standard KKT theorem). This condition shall play a role in characterizing conditions under which KKT points are unique for each optimal solution.

# 4 Differentiability of the Value Function

With this first order theory in hand, we now turn to the question of differentiability of the value function. In this section, we provide successively stronger characterizations of the envelopes (of course, under successively stronger sufficient conditions). In reading these results, it is important to keep in mind that the objective will only be assumed to be locally Lipschitz (not smooth) in many of the cases. We first discuss the question of differential bounds.

## 4.1 Differential Bounds and Clarke Envelope

For the class of optimization problems in (3.0.1), under Assumptions 1 and 2, in a companion paper we prove an important result providing bounds on the lower and upper Dini derivatives of the value function (e.g., Morand Reffett and Tarafdar [35] and Tarafdar [45], chapter 2 Theorem 4.3). We state this result now without proof, and refer the reader to our companion paper for a comprehensive discussion.

**Lemma 6** *Under Assumptions 1 and 2, if additionally $f$ is locally Lipschitz, $D(s)$ is nonempty and uniformly compact near $s$, and MFCQ holds at the optimal solution $a^*(s) \in A^*(s)$, then, in any direction $x \in \mathbf{R}^m$, we have:*

$$\liminf_{t \to 0^+} \frac{V(s+tx) - V(s)}{t} \geq \inf_{(\lambda,\mu) \in K(a^*(s),s)} L_s^{-o}(a^*(s), s, \lambda, \mu; x)$$

$$\limsup_{t \to 0^+} \frac{V(s+tx) - V(s)}{t} \leq \sup_{(\lambda,\mu) \in K(a^*(s),s)} \{L_s^{o}(a^*(s), s, \lambda, \mu; x)\}$$

*therefore,*

$$\begin{aligned}
&\sup_{a^*(s) \in A^*(s)} \min_{(\lambda,\mu) \in K(a^*(s),s)} \left\{ L_s^{-o}(a^*(s), s, \lambda, \mu; x) \right\} \\
&\leq D_+ V(s; x) \leq D^+ V(s; x) \\
&\leq \sup_{a^*(s) \in A^*(s)} \max_{(\lambda,\mu) \in K(a^*(s),s)} \{L_s^{o}(a^*(s), s, \lambda, \mu; x)\}
\end{aligned}$$

*where*

$$\begin{aligned}
L_s^{-o}(a^*(s), s, \lambda, \mu; x) &= \min_{\varsigma_s \in \partial_s f(a^*(s),s)} \left[ \left( \varsigma_s - (\lambda^T g_2 + \mu^T h_2)(a^*(s), s) \right) \cdot x \right] \\
L_s^{o}(a^*(s), s, \lambda, \mu; x) &= \max_{\varsigma_s \in \partial_s f(a^*(s),s)} \left[ \left( \varsigma_s - (\lambda^T g_2 + \mu^T h_2)(a^*(s), s) \right) \cdot x \right]
\end{aligned}$$

Notice that although Lemma 6 only provides differential bounds, these bounds are easily calculated directly from the Lagrangian at the optimum. Under MFCQ, tighter (or exact) bounds are not possible, as one loses uniqueness of the KKT multipliers for each optimal solution and therefore cannot pin down the values of the directional derivatives in any direction.

In the next section, we will sharpen this characterization of bounds, but at the cost of using a stronger constraint qualification. Also, one important implication of Lemma 6 is that the value function is locally Lipschitz (Tarafdar [45], chapter 2). This implies that it has a well-defined Clarke gradient, and we later use this fact to provide bounds of the Clarke envelope under the assumptions of Lemma 6. This is true without any additional conditions guaranteeing interior solutions (as, for example, in Clarke [9] or Askri and LeVan [5]).

We now use the lemma above to establish our first result on differential bounds for the Clarke generalized gradient (or, "Clarke envelope") of the value function:

**Theorem 7** *Let Assumptions 1 and 2 hold, and further assume that $f$ is locally Lipschitz, $D(s)$ is nonempty and uniformly compact near $s$, and that MFCQ holds for every optimal solution $a^*(s) \in A^*(s)$. Then, the Clarke envelope $\partial V(s)$ satisfies:*

$$\partial V(s) \subset \left\{ \cup_{a^*(s) \in A^*(s)} \cup_{(\lambda,\mu) \in K(a^*(s),s)} \partial_s L(a^*(s), s) \right\}$$

*where*

$$\partial_s L(a^*(s), s) = (\partial_s f - \lambda^T g_2 - \mu^T h_2)(a^*(s), s)$$

**Proof.** Let $\varsigma \in \partial V(s)$, and consider the sequence $\{s_n\} \to s$ such that $\nabla V(s_n) \to \varsigma$. By Morand Reffett and Tarafdar [35], or Tarafdar [45], Theorem 10(ii), for any direction $x$, as $a_n^*(s_n) \in A_n^*(s_n)$ and $\lambda_n \in K(a_n^*(s_n), s_n)$, we have:

$$\begin{aligned} \nabla V(s_n) \cdot s &= D^+ V(s_n; x) \\ &\leq \max_{\varsigma_{s_n} \in \partial_s f(a_n^*(s_n), s_n)} \left[ \left( \varsigma_{s_n} - \lambda_n^T g_2(a_n^*(s_n), s_n) - \mu_n^T h_2(a_n^*(s_n), s_n) \right) \cdot x \right] \quad (4.1.1) \end{aligned}$$

To simplify notation in the rest of the proof, we denote the object $\varsigma_{s_n}(a_n^*(s_n), s_n)$ by $\varsigma_{s_n}$. As $D(s)$ is nonempty and uniformly compact, and MFCQ holds at all optimal points, by a standard argument (e.g., Gauvin and Tolle [20], Theorem 2.9 and Gauvin and Dubeau [19], corollary 3.6), there exist a subsequence $\{a_n^*(s_n), s_n, \lambda_n\}$, $a^*(s) \in A^*(s)$ and $(\lambda, \mu) \in K(a^*(s), s)$ such that $a_n^*(s_n) \to a^*(s)$, $s_n \to s$ and $\lambda_n \to \lambda$. Taking the limit of (4.1.1), we get,

$$\begin{aligned} \varsigma \cdot s &\leq \max_{\varsigma_s \in \partial_s f(a^*(s),s)} \left[ \left( \varsigma_s - \lambda^T g_2(a^*(s), s) - \mu^T h_2(a^*(s), s) \right) \cdot x \right] \\ &\leq \max_{(\lambda,\mu) \in K(a^*(s),s)} \max_{\varsigma_s \in \partial_s f(a^*(s),s)} \left[ \left( \varsigma_s - \lambda^T g_2(a^*(s), s) - \mu^T h_2(a^*(s), s) \right) \cdot x \right] \end{aligned}$$

Next, $\forall\ a^*(s) \in A^*(s)$, $K(a^*(s), s)$ is compact; thus, we have the following:

$$\varsigma \cdot s \leq \sup_{a^*(s) \in A^*(s)} \max_{(\lambda,\mu) \in K(a^*(s),s)} \max_{\eta \in \partial_s L(a^*(s),s)} [\eta \cdot x]$$

As the Clarke generalized gradient of the Lagrangian, $\partial_s L(a^*(s), s)$, is nonempty and convex by a standard argument, we conclude,

$$\varsigma \cdot s \leq \max_{(\lambda,\mu) \in K(a^*(s),s)} \left[ \eta \cdot x : \eta \in co\{\cup_{a^*(s) \in A^*(s)} \partial_s L(a^*(s), s, \lambda, \mu)\} \right]$$

Thus,

$$\varsigma \in co\{\cup_{a^*(s) \in A^*(s)} \eta : \eta \in \partial_s L(a^*(s), s)$$

implying the result. ■

A few remarks on this theorem. First, the inclusion for the Clarke envelope in the theorem becomes an equality when $V(s)$ is directionally differentiable. We shall discuss sufficient conditions for this in the next section of the paper. Second, Theorem 7 is useful in many applications . For example, this result is a key step in constructing bounds in Lemma 6 to study structural properties of the limiting (stationary) value function in an infinite horizon stochastic dynamic programming problem (e.g., to calculate local Lipschitz bounded needed to construct equicontinous families of value functions at each point $s \in S$ where each iterate of the Bellman operator is an element of this space). Such an approach follows the exact line of argument pursued in Rincon-Zapatero and Santos [41], but only for the nonsmooth case. We shall show how such an argument works in section 4.1 of the paper, but also see Morand, Reffett, and Tarafdar [37] for a related discussion in the context of a much more general class of stochastic dynamic programming problems.

## 4.2 Directional Differentiability

Often, in applications, one needs a sharper form of the envelope than the mere existence of differential bounds. To obtain a stronger characterization of the differential structure of the value function, such as the existence of directionally differentiable envelopes, we need to make a stronger assumption on the constraint qualification. In particular, as previously suggested, we need to impose a constraint qualification that guarantees the uniqueness of the KKT multiplier at each optimal solution. Uniqueness of the KKT multipliers in our arguments will be tied to SMFCQ via an important result of Kyparisis [28] which we state below without proof: [10]

**Proposition 8** *Assume that $f$ is differentiable and Assumptions 1 and 2. The SMFCQ condition holds for every optimal solution $a^*(s)$ iff the KKT multiplier at $a^*(s)$ is unique (i.e., $K(a^*(s), s)$ is a singleton).*

This is an important result. For example, even in the convex case, Rincon-Zapatero and Santos [41] emphasized the importance of having conditions under which the KKT points are *globally* unique in their proof of $C^1$ differentiability of the value function (e.g., see [41], Theorem 3.2). Certainly, LICQ is one such condition, along with enough smoothness in the primitive data of program (3.0.1), to guarantee such uniqueness. However, as the Kyparisis proposition implies, in many contexts, this uniqueness result per the KKT multipliers and optimal solutions is equivalent to SMFCQ, and hence we adopt this constraint qualification in our work.

**Theorem 9** *Under Assumption 1 and 2, if $f$ is $C^1$, $D(s)$ is nonempty and uniformly compact near $s$, and SMFCQ holds for every optimal solution $a^*(s) \in A^*(s)$, then for any direction $x \in \mathbf{R}^m$, the directional envelope is given by*

$$V'(s,x) = \max_{a^*(s)\in A^*(s)} \{L_2(a^*(s), s; \lambda, \mu) \cdot x\}$$

*where*

$$L_2(a^*(s), s, \lambda, \mu) = \left(f_2 - \lambda^T g_2 - \mu^T h_2\right)(a^*(s), s)$$

**Proof.** As MFCQ holds for all $a^*(s) \in A^*(s)$, from Lemma 6 (since $L_s^{-o}$ and $L_s^o$ are identical functions), we have

$$\begin{aligned} &\sup_{a^*(s)\in A^*(s)} \min_{(\lambda,\mu)\in K(a^*(s),s)} \{L_2(a^*(s), s, \lambda, \mu; x)\} \qquad (4.2.1)\\ &\leq D_+V(s;x) \leq D^+V(s;x)\\ &\leq \sup_{a^*(s)\in A^*(s)} \max_{(\lambda,\mu)\in K(a^*(s),s)} \{L_2(a^*(s), s, \lambda, \mu; x)\} \end{aligned}$$

[10] With smooth constraints, the differentiability of the objective is sufficient to prove Kyparisis' result.

By Kyparisis' result in Proposition 8, SMFCQ implies the correspondence $K(a^*(s), s)$ is singleton for each $(a^*(s), s)$; hence, the min and max above coincide. Therefore,

$$\sup_{a^*(s)\in A^*(s)} L_2(a^*(s), s; \lambda, \mu) \cdot x = D_+V(s; x)$$
$$= D^+V(s; x) \sup_{a^*(s)\in A^*(s)} L_2(a^*(s), s; \lambda, \mu) \cdot x$$

Noting that the supremum will be attained Gauvin and Dubeau (e.g., [19], Corollary 4.3), we have:

$$D_+V(s; x) = D^+V(s; x) = \max_{a^*(s)\in A^*(s)} \{L_2(a^*(s), s; \lambda, \mu) \cdot x\}$$

Thus, the upper and lower right Dini derivative of the value function coincide, and the result follows. ■

We make an important remark on Theorem 9, namely, that the hypotheses of the theorem are only sufficient, not necessary. To see this, consider a simple version of the consumer's problem with three goods, where the utility function is Lipschitzian, but not smooth.

**Example:** Consider the following maximization problem:

$$\max_{a_1,a_2,a_3} \{a_1 + \min\{a_2, a_3\}\}$$

subject to

$$\begin{aligned} pa_1 + a_2 + a_3 &\leq m \\ -a_1 &\leq 0 \\ -a_2 &\leq 0 \\ -a_3 &\leq 0 \end{aligned}$$

The value function:

$$V(m, p) = \begin{cases} \frac{m}{p}, & p \leq 2 \\ \frac{m}{2}, & p > 2 \end{cases}$$

with optimal solution:

$$(a_1^*(m, p), a_2^*(m, p), a_3^*(m, p)) = \begin{cases} (\frac{m}{p}, 0, 0), & p < 2 \\ (a^*, m - 2a^*, m - 2a^*),\ a^* \in [0, \frac{m}{2}], & p = 2 \\ (0, \frac{m}{2}, \frac{m}{2}), & p > 2 \end{cases}$$

Clearly, at $p = 2$, the directional derivatives with respect to $p$ exist, and are given by:

$$V_2'(m, p; d)\,|_{p=2} = 0, \qquad d > 0$$
$$= -\frac{m}{4} \cdot d, \qquad d < 0$$

Therefore, clearly, one can construct examples where directionals for the value function exist, yet the objective is not differentiable. The problem with establishing more general results for this case is the following: if $f$ is only locally Lipschitz (e.g., locally Lipschitz is $a$, for each $s$, but $C^1$ in $s$, for each $a$), we can no longer guarantee (through Kyparisis' result) the uniqueness of the KKT multipliers for each optimal solution $a^*(s) \in A^*(s)$. Therefore, we can only calculate the bounds of Dini derivatives, as in Lemma 6. Unfortunately, general conditions on the primitives for ensuring that the multiplier set is a singleton for every optimal solution when the objective is locally Lipschitz and/or directionally differentiable in the choice variable $a$ are not known.

Along these lines, we do state what can be shown when the $K(a^*(s), s)$ is a singleton in the corollary below:

**Corollary 10** *Under Assumption 1 and 2, in problem (3.0.1), if, $f$ is locally Lipschitz in $a$ and $C^1$ in $s$, $D(s)$ is nonempty and uniformly compact near $s$, MFCQ hold for every optimal solution $a^*(s) \in A^*(s)$, and $K(a^*(s), s)$ is a singleton, then for any direction $x \in \mathbf{R}^m$, the directional envelope is given by:*

$$V^{'}(s; x) = \max_{a^*(s) \in A^*(s)} L_2(a^*(s), s, \lambda, \mu) \cdot x$$

We now can return to the example above to discuss this corollary. Let $\lambda_j$ be the associated KKT multiplier with the $j^{th}$ constraint. At any $m$ and $p = 2$ there are multiple optimal solutions, but the multipliers are unique corresponding to each optimal solution. Thus, the directional derivative with respect to $p$, from Corollary 10 is:

$$V_2'(m, p; d) = \max_{a^*(m,p) \in A^*(m,p)} \nabla_5 L(a_1^*, a_2^*, a_3^*, m, p, \lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*; d)$$

Notice at $p = 2$

$$\begin{aligned} V_2'(m,p;d)\,|_{p=2} &= \max_{a_i^*}\{-\lambda_1^* a_1^* d\} \\ &= \max_{a_i^*}\left\{-\frac{1}{2}a_1^* d\right\} \end{aligned}$$

Further, since $\lambda_1^* = 1/p = 1/2$ for all $a^*(m,p) \in A^*(m,p)$ where $p = 2$, we have

$$V_2'(m,2;d) = \begin{cases} 0, & d > 0 \\ -\frac{m}{4} \cdot d, & d < 0 \end{cases}$$

as implied by Corollary 10.

Recall that in convex problems, there is no need for unique multipliers to ensure directional differentiability of the value function at points where the value function is finite (e.g., see also Gol'stein [21] and Gauvin and Tolle [20], corollary 3.5). So, rather than relaxing the smoothness assumption on the objective as done above and imposing conditions for uniqueness of multipliers, we could also seek to relax the constraint qualification from SMFCQ to MFCQ by assuming some convexity. This is done in Milgrom and Segal [31], but for problems with only inequality constraints, and in which the parameter set is one dimensional (the interval $[0,1]$) . We generalize this in the following corollary to Lemma 6 for the convex case:

**Corollary 11** *Let Assumption 1 and 2 hold, and assume additionally the following: $f$ is $C^1$, concave in $a$, $g$ is convex, $h$ is affine in $a$, $D(s)$ is nonempty and uniformly compact near $s$, and MFCQ holds for every optimal solution $a^*(s) \in A^*(s)$. Then, for any direction $x \in \mathbf{R}^m$, the directional envelope is given by:*

$$V'(s;x) = \max_{a^*(s)\in A^*(s)} \min_{(\lambda,\mu)\in K(a^*(s),s)} L_2(a^*(s),s,\lambda,\mu)\cdot x$$

**Proof.** Lemma 6 part (i) provides the lower bounds of The Dini derivative. With additional structure the upper limit is tighter. Let us first choose a sequence $\{t_n\}$ converging to 0 such that:

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

Since $D(s)$ is uniformly compact near $s$, for $n$ large, there exists $a_n^*(s) \in D(s+t_nx)$ such that $V(s+t_nx) = f(a_n^*(s), s+t_nx)$. Since the sequence $\{a_n^*(s)\}$ is in a compact domain, there exists a convergent subsequence. Therefore, without loss of generality, assume that $a_n^*(s) \to a^*(s) \in D(s)$. Then, by the continuity of $V$, $V(s) = f(a^*(s), s)$. As any $a^*(s) \in A^*(s)$ is a global maxima, appealing to strong duality, the Lagrangian also has a global saddle point at $(a^*(s), \lambda, \mu)$ where $(\lambda, \mu) \in K(a^*(s), s)$. Then, for any $((\lambda, \mu) \in K(a^*(s), s)$ , we have the following,

$$
\begin{aligned}
&\limsup_{t\to 0^+} \frac{V(s+tx)-V(s)}{t} \\
&= \lim_{n\to\infty} \frac{V(s+t_nx)-V(s)}{t_n} \\
&= \lim_{n\to\infty} \frac{L(a_n^*(s)+t_n\overline{y}, s+t_nx, \lambda, \mu) - L(a^*(s), s, \lambda, \mu)}{t_n}
\end{aligned}
$$

As $n \to \infty$, $t_n \to 0^+$, and thus,

$$
\begin{aligned}
&\limsup_{t\to 0^+} \frac{V(s+tx)-V(s)}{t} \\
&\leq \lim_{t_n\to 0^+} \frac{L(a_n^*(s)+t_n\overline{y}, s+t_nx, \lambda, \mu) - L(a^*(s), s, \lambda, \mu)}{t_n} \\
&= L_1(a^*(s), s, \lambda^*(s), \mu^*(s)) \cdot \overline{y} + L_2(a^*(s), s, \lambda^*(s), \mu^*(s)) \cdot x \\
&= L_2(a^*(s), s, \lambda^*(s), \mu^*(s))
\end{aligned}
$$

where last equality follows from the first order condition. Note, the above hold, for all $(\lambda, \mu) \in K(a^*(s), s)$, Therefore

$$
\begin{aligned}
&\limsup_{t\to 0^+} \frac{V(s+tx)-V(s)}{t} \\
&\leq \inf_{\lambda,\mu} [L_2(a_I^*(s), s, \lambda, \mu) \cdot x]
\end{aligned}
$$

Thus, the upper bound of Lemma 6 Part (ii) is more tight for all optimal solutions $a^*(s) \in A^*(s)$. Noting that for any $a^*(s) \in A^*(s)$, the lower and upper bound of Lemma 6 parts (i) and (ii) coincide, thus as in Lemma 6 (iii), we have

$$\begin{aligned}
&\sup_{a^*(s)\in A^*(s)} \min_{(\lambda,\mu)\in K(a^*(s),s)} L_2(a^*(s),s,\lambda,\mu)\cdot x \\
&= D_+V(s;x) = D^+V(s;x) \\
&= \sup_{a^*(s)\in A^*(s)} \min_{(\lambda,\mu)\in K(a^*(s),s)} L_2(a^*(s),s,\lambda,\mu)\cdot x
\end{aligned}$$

Hence the result. ■

As shown in the previous theorem, the MFCQ is sufficient to guarantee directional differentiable envelopes for convex programming problems. The crucial condition here is is that for any $a^*(s) \in A^*(s)$, we have a global maxima, so that any (local) saddle point of the classical Lagrange is also a *global* saddle point. We believe a simple version of this theorem might be available for quasiconcave programming problems (as, for example, Sion's theorem, for example, guarantees the existence of saddlepoints for quasiconvex-quasiconcave saddlefunctions in a minimax problem). The issue that remains is developing an appropriate conjugation scheme (or surrogate duality scheme) such that the set of local maxima and global maxima coincide in the dual program.

It is also useful to illustrate Theorem 11 with an example of a case where directional differentiable envelopes do exist in spite of non-unique multipliers for each optimal solution. For this, we pick a concave programming problem with a convex constraint set.

**Example:** (*Consumer's Problem with value function directionally differentiable, but not* $C^1$). Preferences are $u : \mathbf{R}^2_+ \to \mathbf{R}$, with $u(a_1, a_2) = a_1 + a_2$. The prices are parametric, and given as follows:

$$\begin{aligned}
p_1 &= 1 \text{ if } a_1 \le 5 \\
&= 2 \text{ if } a_1 > 5
\end{aligned}$$

$$\begin{aligned}
p_2 &= 1 \text{ if } a_2 \le 5 \\
&= 2\text{if } a_2 > 5
\end{aligned}$$

Consider for instance income $m \in [8, 12]$. Then, the budget correspondence can described as:

$$a_1 + a_2 - m \le 0 \tag{4.2.2}$$

$$2a_1 - 5 + a_2 - m \le 0 \tag{4.2.3}$$

$$a_1 + 2a_2 - 5 - m \le 0 \tag{4.2.4}$$

$$2a_1 - 5 + 2a_2 - 5 - m \le 0 \tag{4.2.5}$$

The consumer's version of problem (3.0.1) is given as follows:

$$\max_{a_1,a_2}\{a_1 + a_2\}$$

subject to the constraints (4.2.2) to (4.2.5). Denote the KKT multipliers corresponding to the above constraints, respectively, as $\lambda^i$, for $i = 1, 2, 3, 4$. The Lagrangian is given by:

$$\begin{aligned} L(a_1, a_2, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = {} & a_1 + a_2 - \lambda_1 (a_1 + a_2 - m) \\ & - \lambda_2 (2a_1 - 5 + a_2 - m) - \lambda_3 (a_1 + 2a_2 - 5 - m) \\ & - \lambda_4 (2a_1 - 5 + 2a_2 - 5 - m) \end{aligned}$$

The (unique) optimal solution of the Lagrangian problem $m = 10$ is given as:

$$\begin{aligned} (a_1^*(m), a_2^*(m), \lambda_1^*(m), \lambda_2^*(m), \lambda_3^*(m), \lambda_4^*(m)) = {} & \left(\frac{m}{2}, \frac{m}{2}, \lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*\right) \\ & \lambda_1^* + 2\lambda_2^* + \lambda_3^* + 2\lambda_4^* = 1,\ \lambda_2^* = \lambda_3^* \end{aligned}$$

At $m = 10$, all constraints are binding; both LICQ and SMFCQ are violated. However, MFCQ does hold. Further, the objective is concave and the inequality constraints are convex; therefore, by applying Theorem 11, the directionally differentiable envelopes are given by the Lagrangian in direction $d$,

$$V^{'}(m; d) = \min_{\lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*} \{L_2(a_1^*(m), a_2^*(m), m; \lambda_1^*, \lambda_2^*, \lambda_3^*, \lambda_4^*)d\}$$

At $m = 10$,

$$V'(m; d)\,|_{m=10} = \begin{cases} 0.5d, & d > 0 \\ 0, & d < 0 \end{cases}$$

Alternatively, since the value function is given by,

$$V(m) = \begin{cases} m, & m \leq 10 \\ \frac{m+10}{2}, & m > 10 \end{cases}$$

the directional derivatives can be calculated directly from the value function, and they exactly coincide with the directional derivative given by Theorem 11.

In the example, we picked a concave utility function, but this is not important. We would get very similar results for a strictly quasiconcave objective (e.g., $u = x \cdot y$). In that case, the optimal solution at $m = 10$ remains the same, and is a global maxima. The directionally differentiable envelope can thus be computed from the Lagrangian for this quasiconcave programming, similar to the envelope result of Theorem 11. This is the basis for our conjecture that conditions for ordinal versions of concavity might be available.

Finally, we note that our results on nonsmooth envelopes collapse to something closely related to Clarke [9] Theorem 2.1 for unconstrained problems. Evidently, for an unconstrained problem there is no need for KKT theory. Thus, getting directionally differentiable envelopes does not require strong assumptions on the objective like $C^1$ in $s$; rather, Clarke regularity will suffice. The following corollary gives the conditions under which an unconstrained problem has directionally differentiable envelopes.

**Corollary 12** *Under Assumption 1 and 2, if $f$ is jointly locally Lipschitz and Clarke regular in $s$, and optimal solutions are interior, then for every optimal solution $a^*(s) \in A^*(s)$ and any direction $x \in \mathbf{R}^m$, the directional envelope is given by:*

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

**Proof.** We will improve the lower bound in Lemma 6 (i),

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

Since $f$ is directionally differentiable in $s$ the last equality follows. By Clarke regularity, the upper bound in Lemma 6 (ii) is:

$$\begin{aligned}
&\limsup_{t\to 0^+} \frac{V(s+tx) - V(s)}{t} \\
&\leq f_s(a^*(s), s; x)
\end{aligned}$$

Similarly, by Lemma 6 (iii), for any direction $x \in \mathbf{R}^m$, we have

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

■

Versions of this result (and Clarke's) have found extensive applications in a number of papers. One important example is the work of Askri and LeVan [5] on envelope theorems in nonclassical multisector growth models with constraints, but where strong interiority assumptions of optimal solutions are required for Clarke's theorem directly to apply. A somewhat related approach to envelopes is taken in Amir, Mirman, and Perkins [1] and Amir [2] in which supermodularity is used to obtain a standard Euler equation that is is necessary for all interior optimal solutions (although, the value function is only Clarke regular). Dechert and Nishimura [15] also discuss directionally differentiability results in a one-sector non classical optimal growth model, but only under a single equality constraint. Very importantly, all these papers require interiority, and none of their arguments work for the general case of inequality constraints. Finally, Mirman and Zilcha [33] and some of the results in Milgrom and Segal [31] are also special cases of our result, as for their problems, the authors consider a convex programs with one dimensional parameters restricted to subsets of the real line and single inequality constraint.

## 4.3 Generalized Gradient of the Value Function

Whenever the Dini derivatives of the value function are bounded above and below, the value function is locally Lipschitz and the Clarke generalized directional derivatives and gradients exist. The only remaining issue is how to calculate them. Recall that in Section 3.1 of the paper we established bounds for the directional derivative and the Clarke envelopes. Next, with some additional structure in Section 3.2, we obtained directional envelopes. Now, Corollary 13 below gives condition under which the Clarke gradient of the value function exactly coincides with the Clarke gradient of the Lagrangian at the optimal solution.

**Corollary 13** *If the assumptions of Theorem 9, Corollary 10, and Corollary 12 hold for every optimal solution $a^*(s) \in A^*(s)$, then the Clarke envelope is* *calculated as:*

$$\partial V(s) = \left\{ \cup_{a^*(s) \in A^*(s)} \partial_s (f - \lambda^T g - \mu^T h)(a^*(s), s) \right\}.$$

**Proof.** From Corollary 7 we know,

$$\partial V(s) \subset \left\{ \cup_{a^*(s) \in A^*(s)} \partial_s (f - \lambda^T g - \mu^T h)(a^*(s), s) \right\} \tag{4.3.1}$$

Theorem 9, Corollary 10, and Corollary 12 states,

$$V^{'}(s;x) = \max_{a^*(s) \in A^*(s)} \left\{ \left( f_2(a^*(s), s) - \lambda^T g_2(a^*(s), s) - \mu^T h_2(a^*(s), s) \right) \cdot x \right\}$$

For Corollary (12) consider the set of active inequality constraints (i.e holds with equality at the optimal solution) to be a null set, and no equality constraints. Now, consider any $\eta$, such that

$$\eta \in \cup_{a^*(s) \in A^*(s)} \left\{ f_2(a^*(s), s) - \lambda^T g_2(a^*(s), s) - \mu^T h_2(a^*(s), s) \right\} \tag{4.3.2}$$

Therefore, for any direction $x$, from Theorem 9, Corollary 10, and Corollary 12 and relation of directional derivative with generalized directional derivative,

$$\begin{aligned} \eta \cdot x &\leq \max_{a^*(s) \in A^*(s)} \left\{ \left( f_2(a^*(s), s) - \lambda^T g_2(a^*(s), s) - \mu^T h_2(a^*(s), s) \right) \cdot x \right\} \\ &\leq V^{'}(s;x) \\ &\leq V^o(s;x) \end{aligned}$$

where $\varsigma_s \in \partial_s f(a^*(s), s)$. From the definition of the generalized gradient, any $\eta$ satisfying (4.3.2) is also an element of the generalized gradient of $V$,

$$\eta \in \max_{a^*(s) \in A^*(s)} \partial V(s)$$

The above and result (4.3.1) gives the result. ■

Thus, an advantage of the result is it provide an simple method to calculate the Clarke gradient and the directional derivative using standard linearizations of the Lagrangian.

## 4.4 C¹ Differentiability of the Value Function

The most general result in the literature concerning the existence of a classical $C^1$ envelope theorem appears in a recent paper by Rincon-Zapatero and Santos [41] (e.g., see Theorems 3.1 and 3.2), in which the authors allow for non-interior optimal solutions and inequality constraints. In this section, we use a corollary of our main result on directional differentiability to relate our results to their findings. We also present a simple example in which the value function is $C^1$, but for which the hypothesis of the main theorems in Rincon-Zapatero and Santos [41] do not apply.

First, we give conditions for $C^1$ envelope theorems. Our result is essentially a special case of the Clarke envelope and the directional differentiability envelopes of the previous section. In addition to requiring the KKT multiplier set to be a singleton (through SMFCQ, as in the directionally differentiable case), we also need uniqueness of optimal solutions for each $s$. To achieve the latter condition, we assume the objective to be strictly quasiconcave and the constraints quasiconcave.[11]

**Theorem 14** *Under Assumptions 1 and 2, if (i) $f$ is $C^1$, $f$ is strictly quasiconcave, (ii) $D(s)$ is quasi-convex and uniformly compact, (iii) SMFCQ holds for every optimal solution $a^*(s) \in A^*(s)$, then, $V(s)$ is strictly differentiable (i.e., $C^1$) in $s$, and the classical envelope is given by:*

$$\nabla V(s) = \max_{a^*(s) \in A^*(s)} L_2(a^*(s), s, \lambda, \mu)$$

*where*

$$L_2(a^*(s), s, \lambda, \mu) = \left(f_2 - \lambda^T g_2 - \mu^T h_2\right)(a^*(s), s).$$

**Proof.** From Theorem 9, we know under SMFCQ, for every $a^*(s) \in A^*(s)$, for any direction $x \in \mathbf{R}^m$, the directional derivative exists and is given by

$$V^{'}(s; x) = \max_{a^*(s) \in A^*(s)} \{L_2(a^*(s), s; \lambda) \cdot x\}$$

[11] Relaxing the conditions on the objective in Rincon-Zapatero and Santos [41] to quasiconcavity is not the focal point of our argument; rather, it is the role of the SMFCQ vs. LICQ. As Rincon-Zapatero and Santos [41] are studying a discounted dynamic programming problem, clearly only concavity of objective is needed (e.g., as, in general, the sum of two quasiconcave functions is not quasiconcave). Nothing changes in their arguments for 1 stage programs if quasiconcavity is assumed.

By the strict quasi-concavity hypothesis for $f$,quasi-convexity of $g$, and $h$ affine, the set of optimal solutions $A^{*}(s)$ is singleton for each $s \in S$. Hence, in any direction $x \in \mathbf{R}^{m}$

$$V^{'}(s;x) = \{L_2(a^{*}(s), s; \lambda) \cdot x\}$$

Consequently,

$$D_s V(s) = L_2(a^{*}(s), s; \lambda)$$

Finally since $V : \mathbf{R}^{m} \to \mathbf{R}$, $m$ finite, the value function is $C^1$.

$$\nabla V(s) = \max_{a^{*}(s) \in A^{*}(s)} L_2(a^{*}(s), s, \lambda, \mu)$$

■

The following example shows that the hypotheses of Rincon-Zapetaro and Santos [41] can be strong in some applications. For this example, we again modify the preferences in our previous consumer theory example, and configure the constraints so LICQ fails, but SMFCQ holds.

**Example:** (Consumer's problem under price rationing with a C[1] value function but LICQ fails). We reconsider $u : \mathbf{R}^2_+ \to \mathbf{R}$ is given by $u(a_1, a_2) = a_1 \cdot a_2^2$. The price are as follows:

$$p_1 = \begin{cases} 1 \text{ if } a_1 \leq 5 \\ 2 \text{ if } a_1 > 5 \end{cases}$$

$$p_2 = \begin{cases} 1 \text{ if } a_2 \leq 5 \\ 2 \text{ if } a_2 > 5 \end{cases}$$

Income $m \in [8, 12]$. The budget correspondence is identical to the constraints given in the Example of section 3.2 and is similarly enumerated as:

$$a_1 + a_2 - m \leq 0 \tag{4.4.1}$$

$$2a_1 - 5 + a_2 - m \leq 0 \tag{4.4.2}$$

$$a_1 + 2a_2 - 5 - m \leq 0 \tag{4.4.3}$$

$$2a_1 - 5 + 2a_2 - 5 - m \leq 0 \tag{4.4.4}$$

As the feasible set is convex, and the objective is strictly quasi-concave, the optimal solutions are unique. We denote the KKT multipliers corresponding to the constraint $j$ again by $\lambda_j$. The optimal solutions associated with the Lagrangian of this problem can be verified to be:

$$(a^*(m), \lambda^*(m)) = \begin{cases} (m-5, 5, 100-10m, 0, 10m-75, 0)\,, & m < 10 \\ \left(\frac{m+5}{3}, \frac{m+5}{3}, 0, 0\frac{(m+5)^2}{9}, 0\right), & m = 10 \\ \left(5, \frac{m}{2}, 0, 0, \frac{20m-m^2}{4}, \frac{m^2-10m}{4}\right), & m > 10 \end{cases}$$

For $m < 10$, constraints (4.4.1) and (4.4.3) are binding, and LICQ is satisfied. Dually at $m > 10$, constraints (A3) and (A4) are binding, and LICQ is satisfied. Interestingly, at $m = 10$, all constraints are binding, yet only constraint (4.4.3) is active, while all other constraints are saturated. The vector $r = [-2, 1]^T$ satisfies the conditions for SMFCQ; yet at $m = 10$, LICQ is clearly violated. Hence, Rincon-Zapatero and Santos main theorem ([41], Theorem 3.1) does not apply; yet, we still have unique multipliers under Theorem (9). The value function of this problem is given by,

$$V(m) = \begin{cases} (m-5)25, & m < 10 \\ \frac{(m+5)^3}{27}, & m = 10 \\ \frac{5m^2}{4}, & m > 10 \end{cases}$$

is continuously differentiable with derivative:

$$V'(m) = \begin{cases} 25, & m < 10 \\ \frac{(m+5)^2}{9}, & m = 10 \\ \frac{5m}{2}, & m > 10 \end{cases}$$

# 5 Applications

Our first application is to develop new sufficient conditions for monotone controls in stochastic growth models with Markov shocks. This question was asked in Hopenhayn and Prescott [24], for example. We show our results on envelopes can be used weaken a very strong condition required in Hopenhayn and Prescott's argument. We next reconsider monotone comparative statics in constrained optimization problems. For the sake of concreteness, we study the consumer's problem and the question of normal demand. We propose an aggregation procedure, where a critical single-crossing property is calculated using our envelope theorems. Our results complement some important recent results obtained in Quah [40]. We conclude with an application of the results to payoff equivalence in multidimensional mechanism design problems. In particular, we extend some important results obtained in Krishna and Maenner [27] to a more general nonsmooth setting.

### 5.1 Monotone Controls in Stochastic Growth Models with Markov Shocks

One set of sufficient conditions to establish stochastic stability in stochastic growth models is to impose conditions under which agents have (i) monotone controls for savings/investment decisions, along with conditions on (ii) shocks, preferences, and technologies sufficient to guarantee "tightness" of invariant/ergodic distributions, and (iii) optimal decision rules have Euler inequalities that are necessary (e.g., see Mirman [32] and Nishimura and Stachurski [39], among many other papers). Results on the existence of monotone controls with Euler inequalities for the stochastic growth models with iid shocks are well known. We now consider the case with Markov shocks.

Uncertainty enters in the form of a random production shock $z \in \mathbf{Z} = [z_{\min},\ z_{\max}] \subset \mathbb{R}^{++}$ following a first order Markov process with stationary transition function $Q$. We impose precisely the same assumptions[12] as Hopenhayn and Prescott ([24], section 6, part B), except for a critical one: we do not restrict attention to production functions $f(k, z)$ that imply that the correspondence $D(s) = [0, f(k, z)]$ has sublatticed graph. That is, in Hopenhayn and Prescott ([24], Proposition 2), to develop their dynamic programming argument, they require that the feasible correspondence $D(s)$ have a sublatticed-graph to apply a critical theorem in Topkis on preserving supermodularity to the value function at each stage of the dynamic program (e.g., Topkis [46], Theorem 2.7.6). In practice, this condition excludes all but Leontief production.[13] This is clearly a major shortcoming. We avoid this problem by using our envelope results to give show that we can preserve supermodularity of the value function under maximization, as is needed in

[12] Including standard restrictions on the production sufficient to guarantee compactness of the state space.

[13] To see that the sublatticed graph restriction on $f(k, z)$ in Hopenhayn and Prescott [24] is very strong, consider the production function $f(k, z) = zk$. Letting $s = (k, z)$, $s_1 = (2, 1)$, $s_2 = (1, 2)$, for $D(s)$ to have a sublatticed graph, we need for $a_1 \in D(s_1)$ and $a_2 \in D(s_2)$ imply that $a_1 \wedge a_2 \in D(s_1 \wedge s_2)$ and $a_1 \vee a_2 \in D(s_1 \vee s_2)$. Take $a_1 = a_2 = 2$ but $2 = a_1 \wedge a_1 \notin D(s_1 \wedge s_2)$.

any dynamic programming argument.

For this section, we assume the Hopenhayn and Prescott [24], excepting requiring $f(k,z) = f(s)$ to be such that $D(s) = \{a|a \in [0, f(s)]$ has a sublatticed graph:

*Assumption 4.1.1: The utility function* $u : K \to \mathbb{R}_+$ *is strictly increasing, concave, smooth, and* $u'(0) = +\infty$*;* $0 < \beta < 1$ *with* $u(0) = 0$*.*

*Assumption 4.1.2: The production function* $f : K \times Z \to \mathbb{R}_+$ is *increasing, smooth, supermodular and* $f(0,z) = 0$*. In addition, there exists* $k_{\max} > 0$ *such that* $\forall k > k_{\max}$, $f(k, z_{\max}) < k$*.*

*Assumption 4.1.3:* $Q$ *is a stochastically increasing transition function.*

*Assumption 4.1.4: The utility function and the production function satisfy the curvature condition* $u'' f_1 f_2 + u' f_{12} \geq 0$*.*

Given the boundedness assumptions in Assumption 4.1.1-4.1.2, appealing to a standard argument, the planner's optimal growth problem can be solved for a unique value function $V^*$ satisfying the Bellman equation:

$$V^*(k,z) = \max_{0 \leq k' \leq f(k,z)} \left\{ u(f(k,z) - k') + \beta \int V^*(k', z') Q(z, dz') \right\} \quad (5.1.1)$$

If $V_0$ is any continuous bounded value function, then the fixed point of (5.1.1) may be obtained as the pointwise limit of the sequence of functions $\{V_n\}_{n=1}^{\infty}$, where $V_n$ is defined recursively by the Bellman operator $V_n = T^n(V_0)$:

$$T^n(V_0)(k,z) = \max_{0 \leq y \leq f(k,z)} \left\{ u(f(k,z) - k') + \beta \int T^{n-1}(V_0))(y, z') Q(z, dz') \right\} \quad (5.1.2)$$

Denoting by $Y_n^*(k,z)$ the optimal solutions in the above problem, a standard application of Berge's Theorem of the Maximum establishes that $Y_n^*(.,z)$ is a non-empty, compact-valued and upper hemicontinuous correspondence. We denote $y_n^*(k,z) = \vee Y_n^*(k,z)$ the maximal optimal solution.[14]

Our proof of supermodularity of $V$ goes as follows. First, given any initial locally Lipschitz supermodular value function $V_0$ (e.g., $V_0 = 0$), we show that the sequence of function $\{V_n\}_{n=1}^{\infty}$ is such that given any $z$, each $V_n(.,z)$ is Lipschitz on $[\theta, k_{\max}]$ with $\theta$ arbitrarily close to 0 (lemma (15) below). We then use this result to show that each $V_n(.,z)$ is absolutely

[14] As in all of these arguments, optimal solutions are compact subchains, the sup is achieved by the max (respectively, the inf is achieved by the min).

continuous on its domain $[0, k_{\max}]$ (Lemma 16), which implies that it must be equal to its indefinite integral (Proposition 17). Finally, we show that at each stage $n$, Assumptions 4.1.1 - 4.1.4 implies that this indefinite integral is increasing in $z$, which proves that each $V_n(k, z)$ is a supermodular function (Proposition 18). As supermodularity is preserved under pointwise limits, the value function $V^*$ inherits this property (Theorem 19). We therefore we obtain our result for a much broader set of production functions than Hopenhayn and Prescott [24].

**Lemma 15** *Under assumptions 4.1.1 - 4.1.3 for any* $0 < \theta < k_{\max}$ *any* $z$, $V_n(., z)$ *is globally Lipschitz on* $[\theta, k_{\max}]$.

**Proof.** A recursive argument will do. Fix $z$ in $Z$ and consider any function $V_p : K \times Z \to R$ such that $V_p(., z)$ is locally Lipschitz in $k$ in $)0, k_{\max}]$. Clearly $V_{p+1}(., z)$ is increasing and therefore almost everywhere differentiable, and Corollary (10) implies that where $V_{p+1}$ is differentiable (with respect to its first argument), necessarily:

$$V'_{p+1}(k, z) = \max_{y \in Y^*_{p+1}(k,z)} \{u'(f(k, z) - y) f_1(k, z)\}.$$

But since $u$ is strictly concave it must be the case that:

$$V'_{p+1}(k, z) = u'(f(k, z) - y^*_{p+1}(k, z)) f_1(k, z)$$

where $y^*_{p+1}(k, z) = \max Y^*_{p+1}(k, z)$ (i.e., the envelope is evaluated at the maximal optimal selection).

We now prove that $f(k, z) - y^*_{p+1}(k, z)$ is bounded away from 0 on any $[\theta, k_{\max}]$ where $\theta > 0$. Suppose this is not the case, that is, suppose that given any sequence $\{\varepsilon_n\}_{n \in \mathbb{N}}$ converging to 0, there exists a sequence $\{k_n\}_{n \in \mathbb{N}}$ of elements of $[\theta, k_{\max}]$ such that:

$$0 \leq f(k_n, z) - y^*_{p+1}(k_n, z) < \varepsilon_n$$

that is:

$$\lim_{n \to +\infty} y^*_{p+1}(k_n, z) = f(k_n, z).$$

Since $[\theta, k_{\max}]$ is compact, from the sequence $\{k_n\}_{n \in \mathbb{N}}$ , one can extract a converging subsequence; we denote this sequence by $\{k'_n\}_{n \in \mathbb{N}}$, and $k'$ its limit (which belongs to $[\theta, k_{\max}]$). As the correspondence $Y^*_{p+1}(., z)$ is upper hemicontinuous, since $\{k'_n\} \to k'$ (and by definition $y^*_{p+1}(k'_n, z) \in Y^*_{p+1}(k'_n, z)$), there exists a convergence subsequence of $\{y^*_{p+1}(k'_n, z)\}_{n\in\mathbb{N}}$ whose limit point is in $Y^*_{p+1}(k', z)$. But, by this construction, the sequence $\{y^*_{p+1}(k'_n, z)\}_{n\in\mathbb{N}}$ converges to $f(k', z)$ so all its subsequences converge to $f(k', z)$. That is, upper hemicontinuity requires $f(k', z)$ to belong to $Y^*_{p+1}(k', z)$, which is impossible since it violates interiority of optimal solutions. Hence it must be that $f(k, z) - y^*_{p+1}(k, z)$ is bounded away from 0 on any $[\theta, k_{\max}]$ where $\theta > 0$, that is, there exists $\eta_{p+1}(\theta) > 0$ such that $f(k, z) - y^*_{p+1}(k, z) > \eta_{p+1}(\theta) > 0$. Therefore, for any $k \in [\theta, k_{\max}]$, we have

$$0 \leq u'(f(k, z) - y^*_{p+1,\max}(k, z))f_1(k, z) \leq u'(\eta_{p+1}(\theta))f'(\theta, z).$$

This establishes that for any given $z$, the derivative $V'_{p+1}(., z)$ exists almost everywhere, and is bounded on $[\theta, k_{\max}]$. As $V_{p+1}(., z)$ continuous, differentiable almost everywhere, and with a bounded derivative where it exists, it is therefore necessarily (globally) Lipschitz on $[\theta, k_{\max}]$. A recursive argument , then, establishes the Lipschitz property of each $V_n(., z)$ on any $[\theta, k_{\max}]$ with $0 < \theta$. ■

We now use this lemma to prove another lemma on the absolute continuity of the value function.

**Lemma 16** *Under assumptions 4.1.1-4.1.3, for all $z$, $V_n(., z)$ is absolutely continuous on $[0, k_{\max}]$.*

We need to establish that for any $\varepsilon > 0$, there exists $\delta > 0$ such that for any finite collection of non-overlapping intervals $\{]x_i, x'_i[\}_{i=1}^n$ satisfying:

$$\sum_{i=1}^{N} |x'_i - x_i| < \delta$$

it necessarily must be the case that:

$$\sum_{i=1}^{N} |V_n(x'_i, z) - V_n(x_i, z)| < \varepsilon. \tag{5.1.3}$$

Given $\varepsilon > 0$ we divide the collection of non-overlapping intervals in two disjoint subcollections, one near 0 and one away from it. On the first subcollection, we exploit the continuity and monotonicity of $V_n$ to show that the corresponding first part of the sum of (5.1.3) can be made arbitrarily small. Then we exploit the Lipschitzness of $V_n$ on any $[\theta, k_{\max}]$ established in the previous lemma to show that the second part of the sum in (5.1.3) (corresponding to the second subcollection) can also be made arbitrarily small. The details of this proof are as follows.

Since $V_n(.,z)$ is continuous, for any $\varepsilon/2$, there exists $k_0 > 0$ such that $V_n(k_0,z) - V_n(0,z) < \varepsilon/2$. Consider $\theta = \min(k_0, \alpha)$ (the specific $\alpha$ is the one introduced in the previous lemma). Without loss of generality, there exists $P \geq 1$ such that $(x_i, x_i')$ belong to $(0,\theta)$ for $i = 1, ..., P$ and $(x_i, x_i')$ belong to $[\theta, k_{\max}]$ for $j = P+1, ..., N$. On $[\theta, k_{\max}]$ we exploit the property established in the previous lemma that $V_n(.,z)$ is Lipschitz, so that there exists $M_n(z)$ such that:

$$|V_n(k',z) - V_n(k,z)| \leq M_n(z)\,|k'-k|$$

As a result:

$$\sum_{i=j+1}^{n} |V_n(x_i',z) - V_n(x_i,z)| \leq M_n(z) \sum_{j=1}^{n} |x_i' - x_i| < M_n(z)\delta$$

On $[0,\theta]$, we exploit the property that $V_n$ is increasing in its first argument so that:

$$\sum_{i=1}^{j} |V_n(x_i',z) - V_n(x_i,z)| \leq V_n(\theta,z) - V_n(0,z) \leq V_n(k_0,z) - V_n(0,z) < \varepsilon/2$$

Together these two inequalities imply that:

$$\sum_{i=1}^{n} |V_n(x_i',z) - V_n(x_i,z)| < M_n(z)\delta + \varepsilon/2$$

and therefore $\delta = \varepsilon/(2M_n(z))$ establishes the desired result.

Having thus established that all the functions $V_n(.,z)$ are absolutely continuous on $\mathbf{K}$, and given that increasing functions are almost everywhere differentiable, we show that each $V_n$ is in fact equal to its indefinite integral, as stated in the following proposition.

**Proposition 17** *Under assumptions 4.1.1-4.1.3, for all $n \geq 1$, $V_n$ is equal to its indefinite integral, that is:*

$$V_n(k,z) = \int_0^k V_n'(x,z)dx$$

*where:*

$$V_n'(k,z) = u'(f(k,z) - y_n^*(k,z))f_1(k,z)$$

*and* $y_n^*(k,z) = \vee Y_n^*(k,z)$

**Proof.** We have shown above that for each $z$, $V_n(.,z)$ is absolutely continuous on $[0,k_{\max}]$. By Theorem 5.14 in Royden it is equal to its indefinite integral, that is:

$$V_n(k,z) = \int_0^k V_n'(x,z)dx.$$

and we have established also that:

$$V_n'(k,z) = u'(f(k,z) - y_n^*(k,z))f_1(k,z)$$

where $y_n^*(k,z) = \vee Y_n^*(k,z)$. ■

We now show the value functions for the finite horizon problem are each supermodular, with the optimal solutions are ascending in the strong set order (with the top and bottom elements increasing selections).

**Proposition 18** *Under assumptions 4.1.1-4.1.4, for all* $n \geq 1$, $V_n$ *is supermodular,* $Y_n^*((k_t,z_t)$ *is strong set order ascending, and* $\vee Y_n^*$ *and* $\wedge Y_n^*$ *are increasing in* $(k_t,z_t)$*.*

**Proof.** Assume that $V_{n-1}$ is supermodular and recall that:

$$V_n(k,z) = \max_{0 \leq k' \leq f(k,z)} \{u(f(k,z) - k') + \beta E V_{n-1}(k',z')\}.$$

By supermodularity of $V_{n-1}$ in $(k',z')$, $\beta E V_{n-1}(k',z')$ has increasing differences in $(k',z)$. Also, $u(f(k,z)-k')$ has increasing differences in $(k',z)$ since $u'' \leq 0$ and $f_2 \geq 0$. Under assumption 4.1.4, the objective above has increasing differences in $(k,z)$, so by Topkis (Theorem 2.8.3), $Y_{n+1}^*$ is strong set order ascending, and $\vee Y_{n+1}^*$ and $\wedge Y_{n+1}^*$ are increasing in $(k,z)$. Recalling that:

$$V_n^{'}(k,z) = u'(f(k,z) - y_n^*(k,z))f_1(k,z)$$

the curvature assumption 4.1.4 implies that:

$$-u''(f(k,z) - y_n^*(k,z))f_1(k,z) + u'(f(k,z) - y_n^*(k,z))f_{12}(k,z) \geq 0$$

which, together with $y_n^*(k,z) = \vee Y_n^*$ increasing in $z$ implies in turn that $V_n'(k,.)$ is an increasing function of $z$. Thus, for $k' > k$ and $z' > z$ and any $n \geq 1$, by the previous proposition (17):

$$V_n(k',z') - V_n(k,z') = \int_k^{k'} V_n'(x,z')dx$$

$$\geq$$

$$\int_k^{k'} V_n'(x,z)dx = V_n(k,z) - V_n(k,z)$$

i.e., $V_n$ is a supermodular function in $(k,z)$. Given that the proposition is true for $n = 0$ (simply choose $V_0 = 0$), a recursive argument proves the desired result. ■

Finally, we prove our main theorem of this section. Namely, we use Proposition 18 along with a standard limiting argument to show that the stationary solution to the Bellman equation also has the properties of the Proposition.

**Theorem 19** *Under assumptions 4.1.1-4.1.4, the value function $V$ is supermodular, $Y^*(k,z)$ is strong set order ascending, and the extremal optimal investment policies $\vee Y^*$ and $\wedge Y^*$ are increasing in $(k,z)$.*

**Proof.** $V$ is the pointwise limit of a sequence of supermodular functions, and therefore also supermodular, and the properties of the optimal investment correspondence $Y^*$ follows from a standard result in Topkis (e.g., [46],Theorem 2.8.3). ■

We mention two additional important implications of these results. First, the results in this section (e.g., Lemma 15) are of independent interest in the literature on Lipschitzian stochastic dynamic programming. That is, noting the Inada condition $u'(0) = +\infty$ (which implies that $u$ is not Lipschitz everywhere on $[0, k_{\max}]$), it is not the case under our assumptions that $V_{p+1}(.,z)$ is Lipschitz everywhere on all its domain $[0, k_{\max}]$. Therefore, for example, the results of Hinderer [23] are not applicable here.[15] Yet, this result will be sufficient to characterize the local Lipschitz stucture of the value function in

[15] We choose the primitives in this example to satisfy Inada conditions, in part, so the

our setting on the interval $[\theta, k_{\max}]$, and we can construct *explicit* estimates of the Lipschitz modulus at eacjh point $k \in [\theta, k_{\max}]$. That prove sufficient to close the argument at all such points at infinity (as this pointwise bound can be used to construct an equicontinuous set of value functions that are converging uniformly to a Lipschitz limit on $[\theta, k_{\max}]$. For many applications (e.g., constructing error bounded for approximate solutions to the Bellman equation via discretization), this result is sufficient (See Morand, Reffett, and Tarafdar [37] for discussion, and for additional extensions of the results in Hinderer [23])).

Second, as we have monotone controls in this problem, at least for the Inada condition case, we can construct exactly the same Euler equation found in Amir, Mirman, and Perkins [1] for the stochastic growth case with Markov shocks (e.g., see Amir, Mirman, and Perkins ([1], Theorem 3.2a)). This is because the essential step in proof of the existence of a standard Euler equation for growth models (stochastic or deterministic) with nonclassical technologies involves using the fact that maximal and minimal selections for investment are isotone. So standard Euler equations (necessary, but not sufficient) are available for this model. Now, this fact is important as the existence of standard ($C^1$) Euler equation is a critical ingredient in the arguments in the literature for conditions for the existence of stochastic stability (e.g., Nishimura and Stachurski [39], secton 4, especially Proposition 4.4). Such stochastic stability arguments can now be easily adapted to the case of stochastic growth model with Markov shocks (as our envelope theorems can be used to prove the existence of an Euler equation, which in turn can be used to construct the requisite Foster-Lyapunov function needed to extend their arguments for stochastic bounds to the case of Markov shocks).

## 5.2 Aggregation and Constrained Lattice Programming

We next apply our results to a class of constrained lattice programming problems. For example, it is well-known that standard lattice programming techniques (e.g., Topkis [46]) in the consumer's problem to obtain conditions

---

example will violate all of Hinderer's Lipschitz conditions on period returns. We can easily dispense with Inada conditions, yet still obtain monotone controls via our lattice programming argument (e.g., see Morand, Reffett, and Tarafdar [37]).

Also, the Inada condition often plays a key role in stochastic stability arguments (e.g., see Nishimura and Stachurski [39]), although also see Kamihigashi [25] for the case without Inada conditions.

for normal demand for a single good (as ordered changes in income do not induce strong set ordered changes in the budget correspondence in the standard Euclidean order $\mathbf{E}^n$,[16] so the standard lattice programming approach does not apply).[17] In recent work, many authors have each proposed interesting new approaches that do obtain conditions for some version of normal demand.[18] In many of these cases, the comparative statics obtained are "weak" comparative statics (e.g., one identifies conditions under which there exists a monotone selection for demand, as opposed to "strong" comparative statics where one has sharp characterizations of monotone selections that are typically obtained in standard lattice programming formulations). In this section, we use our nonsmooth envelope theorems and duality to propose a new method to provide answers to this question to obtain strong comparative statics.

To understand our dual approach to constrained lattice programming, consider a standard monotone comparative statics in the consumer's problem (in particular, income effects). Let $A = C \subset \mathbf{E}^n_+$ be the commodity space, $C$ a nonempty sublattice, $s \in S = [0, \infty)$ be the level of income. For this entire section, fix the vector of prices to be $p \in \mathbf{S}_{n-1}$, with $p >> 0$, and $\mathbf{S}_{n-1}$ is the $n-1$ dimensional simplex. When formalizing our aggregation procedure, we will need to work on sections of the commodity space $C$. Therefore, define following sections of $C$ : (i) $C_i = \{c \in C | c_{-i} = (c_1, c_2, ..., c_{i-1}, c_{i+1}, ..c_n)$ fixed$\} \subset \mathbf{E}_+$, and (ii) $C_{-i} = \{c \in C | c_i \in \mathbf{E}_+$ fixed$\} \subset \mathbf{E}^{n-1}$. Further, as we shall seek strong monotone comparative statics, we need to define the following Veinott sublattice powerdomains (iii) $\mathcal{L}(E_i) = \{L_i \subset E_+ | L_i$ a nonempty sublattice of $E_+)$, and (iv) $\mathcal{L}(E_{-i}) = \{L_{-i} \subset C_{-i} | L_{-i}$ a nonempty sublattice of $E^{n-1}_+\}$. Endow all these sets with Veinott's strong set order.[19]

We shall assume the cardinal complementarity conditions in Quah [40].

[16] The space $\mathbf{E}^n$ is the pair $(\mathbf{R}^n, \geq_e)$ where $\geq_e$ is the standard componentwise (product) Euclidean order.

[17] Although in this section we consider the consumer's problem, nothing in our arguments change if instead we assume the constraint set is generated by smooth function $g : C \to \mathbf{R}$ that is increasing, convex, submodular, and satisfies a Slater condition in $c_{-i}$. This problem is among those studied in Morand, Reffett, and Tarafdar [36]. Also, direct extensions are available for the case that commodity spaces are Riesz spaces. We simply need to conjugate Lagrangian dual programs in such a setting. See Ekeland and Teman [16] for discussion.

[18] See, for example, [3][4][40][34], among others.

[19] That is, we seek standard monotone comparative statics. Also, note we can define similar Veinott sublattice powerdomains can be defined for the dual variable $\lambda \subset \mathbf{E}_+$ in the Lagrangian problem we write down in a moment.

That is, we assume the consumer's preferences can be represented by a real-valued utility function $u : C \to \mathbf{R}$, where $u(c)$ satisfies the following assumption:

*Assumption 4.2.1: ($C_i$−supermodularity). The function $u(c)$ is (a) $C^1$ in $c_{-i}$* for each $c_i$, *(b) supermodular on $C \subset \mathbf{E}^n$, (c) concave in $c_{-1}$ for each $c_1$ (where $c_{-1}$ denotes the vector $c$ with the ith coordinate deleted), (d) $u(c)$ increasing, (e) $u(c)$ is bounded below.*

The conditions in Assumption 4.2.1(a)-(d) are the cardinal complementarity conditions for constrained optimization problems first proposed in Quah [40], and we shall refer to them as the *Quah cardinal complementarity conditions.*[20]

We ask the following question: for a fixed price vector $p >> 0$, what are sufficient conditions for the consumer's demand correspondence for good $c_i$ to be isotone in income $s \in S$ in *Veinott's strong set order* to $\mathcal{L}(C_i)$. For the cardinal case, one answer for these *strong* comparative statics will be Quah's cardinal complementarity conditions.

The consumer's problem can be stated as follows: for fixed price vector $p = (p_1, p_2, .., p_{i-1}, 1, p_{i+1}..., p_n) >> 0$, $p \in \mathbf{S}_{n-1}$, at income level $s \in S$, the consumer solves[21]

$$V(s) = \max_{c \in D(s)} u(c) \tag{5.2.1}$$

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 optimal solutions for the problem are

$$C^*(s) = \arg\max_{c \in D(s)} u(c) \tag{5.2.2}$$

[20] Quah requires only locally Lipschitz structure of payoffs in $c_{-i}$ in his cardinal complementarity conditions (as he imposes partial concavity in the cardinal case). We do not need $C^1$ either (and can just work with directionals and subdifferentials for the value functions, if needed). The point of this section is to show the structure of our dual argument. Hence, for the sake of exposition, we assume smooth primitive data.

[21] 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.

Under Assumption 4.2.1, clearly this program is well-defined (e.g., via a standard application of Berge's theorem). We shall refer to program (5.2.1) as the *primal problem.*

Define the $C_i-$*aggregated primal* (or, simply the "*aggregated primal*") as follows: for any $\infty > s > 0$, and $c_i \in [0, s]$, solve:

$$V^{ap}(c_i, s) = \max_{c_{-1} \in D(c_i, s)} u(c_{-i}; c_i) \tag{5.2.3}$$

The optimal solutions for the aggregated primal are

$$C_{-i}^{ap*}(c_i, s) = \arg \max_{c_{-1} \in D(c_i, s)} u(c_{-i}, c_i)$$

We can now recast this monotone comparative statics problem (under aggregation) as the following problem:

$$V(s) = \max_{c_i \in [0,s]} V^{ap}(c_i, s) \tag{5.2.4}$$

where $V(s)$ is the value function of the primal problem (5.2.1). Notice, program (5.2.4) is a *standard* lattice programming problem if for each $p \in \mathbf{S}_{n-1}$ with $p >> 0$, the value function $V^{Ap}(c_i, s)$ is supermodular in $(c_i, s)$ (as the feasible correspondence in program (5.2.4) is Veinott strong set order isotone in $s$).

To study the complementarity structure of the value function $V^{ap}$ in (5.2.3), first notice as the feasible correspondence is generated by a single constraint $g(c; p) = p \cdot c$ that is smooth, convex (and increasing and submodular also), the feasible correspondence $\Psi(s) = \{c_i | c_i \in [0, s]\}$ has a convex-graph. Therefore, as the objective function in the aggregated primal is concave in the choice variables $c_{-i}$ for each $c_i \in [0, s]$, the value function $V^{ap}(c_i, s)$ is partially concave function in $s$ for each $c_i \in [0, s]$.[22] Further, as the Slater condition is trivially satisfied for this case, for each $c_i \in [0, s]$, the value function $V^{ap}(c_i, s)$ has a non-empty partial subdifferential in $s$ (and, actually, it is directionally differentiable by Corollary 11). So our monotone comparative statics question reduces to characterizing a (nonsmooth) envelope theorem from Corollary 11 in the aggregated problem (5.2.3).

To characterize this nonsmooth envelope, we first conjugate the aggregated primal with a classical Lagrangian dual. Under Assumption 4.2.1, as

[22] e.g., see Fiacco and Kyparisis ([17], Proposition 2.5).

the aggregated primal in (5.2.3) is a standard convex programming problem in $c_{-i}$, we use the following standard duality scheme: for each $c_i \in [0, s]$, and $s > 0$, define

$$\begin{aligned} L(c_{-i}, \lambda; c_i, s) &= u(c_{-i}; c_i) - \lambda(p \cdot c - s) \text{ if } c_{-i} \in C_{-i}, \phi \in \Omega \\ &= +\infty \text{ if } c_{-i} \in C_{-i}, \phi \notin \Omega \\ &= -\infty \text{ if } c_{-i} \notin C_{-i} \end{aligned} \tag{5.2.5}$$

Using this Lagrangian, define the following "$C_i-$aggregated dual" problem (or, simply "aggregated dual") as follows: for $c_i \in [0, s]$, $s > 0$,

$$V^{ad}(c_i, s) = \inf_{\lambda} \sup_{c_{-i}} L(c_{-i}, \lambda; c_i, s) \tag{5.2.6}$$

where $\lambda \in \Lambda \subset \mathbf{R}_+$, $c_{-i} \in C_{-i}$, and $V^{ad}(c_i, s)$ is the value function for the aggregated dual. By strong duality, we have the following in place: (i) zero duality gap (i.e., $V^{ap}(c_i, s) = V^{ad}(c_i, s)$), (ii) standard necessary and sufficient Lagrange multiplier rules, and (iii) saddlepoint stability for optimal solutions, among other properties [23]. The objective function in the Lagrangian dual is

$$V^L(\lambda, c_i, s) = \sup_{c_{-i}} L(c_{-i}; \lambda, c_i, s) \tag{5.2.7}$$

and its set of optimal solutions for the Lagrangian dual is given by

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

We first characterize the correspondence $C_{-i}^{L*}(\lambda, c_i, s)$ using standard lattice programming. We have the following result for $p >> 0$:

**Lemma 20** $C_{-i}^{L*} : \Omega \times C_i \times S \times \mathbf{S}_{n-1} \to \mathcal{L}(\mathcal{C}_i)$ *is (i) descending in* $\lambda \in \Omega$ *(with all selections antitone), for each* $(c_i, s) \in C_i \times S$*, and (ii) ascending in* $c_i$ *(with all selections isotone), each* $(\lambda, s) \in \Omega \times S$.

**Proof.** As $C_{-i} \subset \mathbf{E}^{n-1}$ is a section of $C$, $C_{-i}$ is a sublattice (e.g., Topkis ([46], p. 16). Further, for each $(\lambda, c_i, s) \in \Omega \times C_i \times S$, noting the presence of an indicator function for the sublattice $C_{-i}$ in the definition of the Lagrangian in (5.2.7), we first consider the parameterized Lagrangian $L_{(\lambda, c_i, s)}$:

[23] See Rockafellar [42] for discussion.

$\mathbf{E}^{n-1} \mapsto \mathbf{R}^*_- = \mathbf{R} \cup -\infty$. As (i) $(\mathbf{R}^*_-, *)$ is a posemigroup (with identity element) for $* = +$, and (ii) the indicator function for any sublattice $C_i^{'} \subset C_i$ super$*$ (e.g., Veinott [48], Chapter 6, p. 18) with $*$=+ properly increasing in this posemigroup, under assumption 4.2.1, the Lagrangian is super$*$ in $c_i$ for $* = +$ (and, hence, $L_{(\lambda, c_i, s)}(c_{-i})$ supermodular in $c_{-i}$, for each $(\lambda, c_i, s)$). Further, $L_{(\lambda, c_i, s)}(c_{-i})$ has (i) strict decreasing differences between $(c_{-i}; \lambda)$ for each $(c_i, s)$, and increasing differences between $(c_{-i}; c_i)$ each $(\lambda, s)$ (and is a valuation in $(c_i; s)$ for each $(\lambda, c_i)$). Therefore, by Topkis' theorem (Topkis [46], 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 $\lambda$, each $c_i$, and ascending in $c_i$, each $\lambda$. Noting as in each case, we have strict decreasing (respectively, increasing) differences, $C_{-i}^{L*}(\lambda, c_i)$ has every selection (a) antitone in $\lambda$, for each $c_i$, and (b) isotone in $c_i$, for each $\lambda$ (and independent of $s$). This proves the result.

■

By a standard argument, the value function in $V^L(\lambda, c_i, s)$ convex in $\lambda$, each $(c_i, s)$. It is also submodular (noting that for the posemigroup $(\mathbf{R}^*_+, *)$ with $\mathbf{R}^*_+ = \mathbf{R} \cup \infty$ and $* = +$, indicator functions of sublattices in the space are sub$*$ with $*$ properly increasing). Further, by the convexity and submodularity of $V^L$ in $\lambda$, the effective domain for $V^L$ (i.e., the subset $\Lambda^e \subset \Lambda$, $\Lambda^e = \{\lambda | V^L(\lambda, c_i, s) < \infty\}$ is a convex sublattice. Finally, by a version of Danskin's theorem (e.g., Grinold [22], Lemma, p. 186), we have the right directional envelope $V^L(\lambda, c_i, s)$ in $\lambda$ in direction $e^+ > 0$ given by

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

Further, 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 21** *Under Assumption 4.2.1, in the dual problem (5.2.7), the KKT multipliers $\Lambda^*(c_i, s)$ are (1.a) ascending in Veinott's strong set order in $c_i$, each $s \in S$ with $\wedge\Lambda^*(c_i, s)$ and $\vee\Lambda^*(c_i, s)$ isotone selections in $c_i$, each $s \in$* $[0, s]$; *(1.b) descending in Veinott's strong set order in $s$, each $c_i \in [0, s]$ with $\wedge\Lambda^*(c_i, s)$ and $\vee\Lambda^*(c_i, s)$ antitone selections in $s$, each $c_i \in [0, s]$. Finally, (1.c) the aggregated primal $V^{ad}(c_i, s)$ has increasing differences (and, hence, is supermodular) in $(c_i, s)$.*

**Proof.** Noting the definition of the (partial) directional derivative $V_\lambda^L(\lambda, c_i, s)$ in any direction $e$, as $C_{-i}^L$ is ascending in $c_i$, each $s$, and independent of $s$, each $c_i$, we conclude $V^L(\lambda, c_i, s)$ has (i) decreasing differences between $(\lambda, c_i)$ each $s$,and (ii) increasing differences between $(\lambda, s)$, each $c_i$. Therefore, (1.a) and (1.b) follows directly from Topkis's theorem (Topkis, [46], 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 (5.2.1) using strong duality and Corollary 11 as the following right directional 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. Further, as the selection $\wedge\Lambda^*(c_i, s)$ is also isotone in $c_i$, each $s$, by duality we have increasing differences in $V^{ap}(c_i, s)$ for any right perturbation. Dually, from the left directional is

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

which is also well-defined as $\Lambda^*(c_i, s)$ is subchained. Again, as $\wedge\Lambda^*(c_i, s)$ is also isotone in $c_i$, each $s$, we have increasing differences for $V^{ap}(c_i, s)$ for perturbations from the left. Therefore, $V^{ap}(c_i, s)$ has increasing differences in $(c_i, s)$, and hence, is supermodular. ■

We now have the main comparative statics result of this section:

**Theorem 22** *Under Assumption 4.2.1, (1.a) $C_i^* : S \to \mathcal{L}(C_i)$ is ascending in Veinott strong set order, (1.b) $\vee C_i^*(s)$ and $\wedge C_i^*(s)$ isotone selections. Further, if Assumption 4.2.1 holds for all $i \in \{1, 2, .., n\}$, then statements (1.a) and (1.b) are true for all $i$, and $\vee C_i^*(s)$ and $\Lambda C_i^*(s)$ each Lipschitz selections of modulus 1.*

**Proof.** Claims (1.a) and (1.b) follow from Topkis' theorem (e.g., Topkis ([46], 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 ([13], 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 22. First, Theorem 22 delivers a nonsmooth version of Chipman's result (e.g., Chipman ([8]). Second, a similar approach could be taken to examining monotone comparative statics per price effects. That is, in our setting, issues of (gross) substitutability or complementarity in demand reduces to sufficient condition in the dual program (5.2.7), $\lambda^*(c_i, p)$ is falling or rising in $p_i$ (for own substitution effects or cross substitution effects). This comparative static is rather easily incorporated into our dual formulation, but appears somewhat difficult to accommodate in Quah [40]. Third, under Assumption 4.2.1, the optimal solutions $C^*(s) \notin \mathcal{L}(\mathbf{E}^n)$ (i.e, sublattices in $\mathbf{E}^n$); rather, the optimal solutions are only chain subcomplete in $\mathbf{E}_+^n$. Actually, under strict mononicity assumptions for $u(c)$, $C^*(s)$ is *antichained-valued* (as the optimal solutions each are elements of the budget line, and, hence, an antichain for each $s$). Fourth, if we only need good $i$ strong set order isotone in $s$, we do not need the choice space for $C_i$ to be convex. That is, the section $C_i \in \mathbf{E}_+$ could be a discrete lattice. It is not clear how Quah's flexible set ordering approach can handle this case. Also, conditions for all the selections for good $i$ normal simply require strict increasing differences between $(x_{-i}; x_i)$. Then this result 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 using dual methods involves conjugating quasiconcave programs. For this, we need to pick a different duality scheme with an "ordinal" Lagrangian (e.g., for discussion, see Crouzeix [12] or Trach [47], and references therewithin).

## 5.3 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. This result was first established in Myerson [38] for single dimensional type smooth utility and payoff functions. Recently, Krishna and Maenner [27] have relaxed these smoothness assumptions by either assuming convex sets of types and convex utility functions, or by requiring that the allocation rule, payment function and utility functions be Lipschitz and Clarke regular in all arguments with utility functions 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 $\mathbf{X}$ denote the set of social alternatives, where $\mathbf{X}$ is a subset of $\mathbf{R}^n$. There are $\mathbf{I}$ agents and each $i \in \mathbf{I}$ has a $k$-dimensional type $t_i \in \mathbf{T}_i \subset \mathbf{R}^k$. The set $\mathbf{T} = \mathbf{\Pi}_{j \in \mathbf{I}} \mathbf{T}_i$ 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 $x \in \mathbf{X}$ is the alternative chosen by the planner and $\mu_i$ is a monetary transfer to the planner. A mechanism is a pair $(\chi, \mu)$ where $\chi : \mathbf{T} \to \mathbf{X}$ is the allocation rule and $\mu : \mathbf{T} \to \mathbf{R}^I$ is the payment rule. Thus, if agent $i$, reports type $s_i$, for all $i$ in $\mathbf{I}$, the social planner chooses alternative $\chi(a_1, a_2, ..a_I)$ and the transfer payment of agent $i$ is $\mu^i(a_1, a_2, ..a_I)$.

We maintain the following assumptions in this section:

*Assumption 4.3.1: The type set* $\mathbf{T}_i$ *is compact, separable, and connected.*

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

*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* $\mu^i$ *is locally Lipschitz in reported types* $(s_1, s_2, ..., s_I)$*, for all* $i$*; (c) the allocation rule* $\chi$ is locally Lipschitz.

*Assumption 4.3.4: The utility function* $u^i$ *and the payment function* $\mu^i$ *admit a differential extension on the boundary for all* $i$ *in* $\mathbf{I}$.

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

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

The mechanism $(\chi, \mu)$ is incentive compatible if for all agents $i \in \mathbf{I}$ and all corresponding types $t_i \in \mathbf{T}_i$:

$$V_i(t_i) = \max_{s_i} \left\{ U^i(s_i, t_i) - \overline{\mu}^i(s_i) \right\} \tag{5.3.1}$$

and

$$t_i \in \arg\max_{s_i} \left\{ U^i(s_i, t_i) - \overline{\mu}^i(s_i) \right\}$$

$S_i^*(t_i) = \arg\max_{s_i} \left\{ U^i(s_i, t_i) - \overline{\mu}^i(s_i) \right\}$ denotes the set of optimal solutionof agent $i$, ,

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 $\chi$ are also locally Lipschitz, and $\mathbf{T}_i$ is a compact, separable subset, 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 $(s_i, t_{-i}) \in \mathbf{T}_i \times \mathbf{T}_{-i}$ (and, therefore $\overline{\mu}^i(s_i) = E_{t_{-i}} \mu^i(s_i, t_{-i})$ 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 [27] that we shall use in deriving the main result of this section.

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

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

**Proof.** See Krishna and Maenner ([27], Theorem 1) ■

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

**Theorem 25** *Under 4.3.1-4.3.4, for any incentive compatible mechanism* $(\chi, \mu)$*,* $V_i(t_i)$ *is locally Lipschitz and Clarke regular, with the directional differentiable envelope 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* $\chi$ *up to an additive constant. Finally, for all* $t_i, t_i' \in \mathbf{T}_i$ *and any smooth path* $\alpha$ *joining* $t_i$ *and* $t_i'$ *in* $\mathbf{T}_i$

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

*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 (5.3.1) is locally Lipschitz and Clarke regular in $t_i$, and the set $\mathbf{T}_i$ is compact. From our unconstrained result, the directional differentiable envelope of the indirect utility function is:

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

The directional derivative and hence the Clarke gradient of $V_i$ only depends on the expected utility from allocation and not on the expected payment. Further $V_i$ is locally Lipschitz and Clarke regular. Thus for an incentive compatible mechanism $(\chi, \mu)$,

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

since, $t_i \in S_i^*(t_i)$. As a result, 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 $\chi$ (and not the payment function $\mu$) up to an additive constant. ■

# 6 APPENDIX: Mathematical Terminology

## 6.1 Metric Spaces and Lipschitz Functions

Given the metric spaces $(X, \rho_X)$ and $(Y, \rho_Y)$, a function $f : X \to Y$ is locally *Lipschitz near $x$ with modulus* $k(x)$, $0 \le k < \infty$, if for all $x', x''$ in a neighborhood of $x$,

$$\rho_Y(f(x''), f(x') \le k(x)\rho_X(x'', x').$$

It is *Lipschitz* if $k(x)$ is independent of $x$. If $X$ is a convex set, a real valued function $f : X \to \mathbf{R}$ is *(strictly) convex* if for all $x, y \in X$, and all $\lambda \in (0, 1)$

$$f(\lambda x + (1 - \lambda)y) \le (<)\lambda f(x) + (1 - \lambda)f(y)$$

A *proper convex function* is a real valued *convex function.* A proper convex function is locally Lipschitz for any open set in $X$.

There are several notions of differentiability for a Lipschitz function $f : I \subset \mathbf{R}^n \to \mathbf{R}$ of modulus $k$. In particular, given $x_0 \in I$ and a direction $d \in \mathbf{R}^n$, we define the *upper radical right Dini derivative as:*

$$D^+f(x_0; d) = \lim\sup_{t \to 0^+} \frac{f(x_0 + td) - f(x_0)}{t},$$

and the *lower radical right Dini derivative as:*

$$D_+f(x_0; d) = \lim\inf_{t \to 0^+} \frac{f(x_0 + td) - f(x_0)}{t}.$$

The left Dini derivatives are defined similarly simply changing $t \to 0^+$ to $t \to 0^-$.

The *right directional derivative at* $x_0 \in X$ in the direction $d \in \mathbf{R}^n$ is:

$$f'(x_0; d) = \lim_{t \to 0^+} \frac{f(x_0 + td) - f(x_0)}{t}$$

Similarly the *left directional derivative is defined by* changing $t \to 0^+$ to $t \to 0^-$.

*Clarke upper and lower generalized directional derivatives* at $x_0$ in the direction $d \in \mathbf{R}^n$ are respectively:

$$f^o(x_0; d) = \lim \sup_{\substack{y \to x_0 \\ t \to 0^+}} \frac{f(y + td) - f(y)}{t}$$

$$f^{-o}(x_0; d) = \lim \inf_{\substack{y \to x_0 \\ t \to 0^+}} \frac{f(y + td) - f(y)}{t}$$

It is important to note that Clarke generalized derivatives of Lipschitz functions always exist, while directional derivatives of such functions need not. We say a function $f$ is *Clarke regular* if its upper Clarke generalized directional derivative equals its right directional derivative in all directions $d$.

A function $f$ is *differentiable at* $x_0 \in X$ if its directional derivatives exist in all direction and $f'(x_0; d) = \nabla_x f(x_0) \cdot d$, in which case the derivative is

$$\nabla_x f(x_0) = \lim_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h}$$

The function $f$ has a *strict derivative* at $x_0$, denoted by $D_s f(x_0)$, when for all $d \in \mathbf{R}^n$:

$$\langle D_s f(x_0), d \rangle = \lim_{\substack{x \to x_0 \\ t \to 0^+}} \frac{f(x + td) - f(x)}{t}$$

Finally, we say that $f$ is *continuously differentiable* if $D_s f(x) : \mathbf{R}^n \to \mathbf{R}$ is continuous at $x_0$. In a finite dimensional domain, a strictly differentiable function is continuously differentiable.

Finally, recall that the subgradient of a convex function $f$ is the set of $p \in \mathbf{R}^n$ satisfying:

$$p \cdot d \leq f(x_0 + d) - f(x_0)$$

for all directions $d \in \mathbf{R}^n$. The subgradient of Lipschitz functions may not necessarily exist, but the Clarke generalized gradient always does, and is defined as:

$$\partial f(x_0) = co \{\lim \nabla f(x_i) : x_i \to x_0, x_i \notin S, x_i \notin \Omega_f\}$$

where $co$ denotes the convex hull, $S$ is any set of Lebesgue measure zero in the domain, and $\Omega_f$ is a set of points at which $f$ fails to be differentiable.

## 6.2 Lattices and Supermodularity

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 $B$. A lattice is *complete* if any subset $B$ of $X$ has a least upper bound and a greatest lower bound. At lattice is $\sigma - complete$ if every sequence has a least upper bound and greatest lower bound. If every chain in $X$ is complete, then $X$ is referred to as a *chain complete poset.* If every countable chain in $X$ is complete, then $X$ is referred to as a *countably chain complete poset.*

Let $(X, \geq_X)$ and $(Y, \geq_Y)$ be Posets. A mapping $f : X \to Y$ is *isotone* on $X$ if $f(x') \geq_Y f(x)$, when $x' \geq_X x$, for $x, x' \in X$. If $f(x') \geq_Y f(x)$ when $x' >_X x$ for $x$, $x' \in X$, then we say $f$ is *increasing.* If $f(x') >_Y f(x)$ when $x' >_X x$, then we say $f$ is *strictly increasing.* 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 (e.g, see [48], Chapter 4). Let $L(Y) = \{A | A \subset Y$, $A$ a nonempty sublattice$\}$. We order $2^Y$ 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., $>$) $f(x) + f(y)$.[24] An important property of the class of supermodular functions is they are closed under pointwise limits. (Topkis, [46], 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

[24]In some cases, it proves useful to let $f$ be extended real valued (e.g., section 4.2 of this paper per indicator functions). In such a case, for our applications, following Veinott ([48], Chapter 6), its best to think of $f$ as a super$*$ function from a lattice to a possemigroup $(\mathbf{R}^*_-, *)$ where $*=$"+" with $\mathbf{R}^*_- = \mathbf{R} \cup \{-\infty\}$. In this case, the idea of a supermodular function can easily be extended to the case of extended real valued objectives.

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] R. Amir L. Mirman, W. Perkins, One-sector nonclassical optimal growth: optimality conditions and comparative dynamics, Int. Econ. Rev. 32 (1991) 625-644.

[2] R. Amir, Sensitivity analysis of multisector optimal economic dynamics, J. Math. Econ. 25 (1996) 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, and 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] J. Chipman, An Empirical Implication of Auspitz–Lieben–Edgeworth–Pareto Complementarity, Journal of Econ. Theory 14 (1977)228–231.

[9] F. Clarke, Generalized gradients and applications, Transactions of the American Mathematical Society 205 (1975) 247-262.

[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] Crouzeix, J.P. 1977. Contribution á l'étude des fonctions quisiconvexes" Thése, Université de Clemont.

[13] L. Curtat, Markov equilibria of stochastic games with complementarities, Games Econ. Behav. 17 (1996) 177-199.

[14] J. Danksin, The Theory of Max-Min, Springer, 1967.

[15] W. Dechert, K. Nishimura, A complete characterization of optimal growth paths in an aggregative model with a non-concave production function, Journal of Econ. Theory 31 (1983) 332-354.

[16] Ekeland, I. and R. Teman. 1976. *Convex Analysis and Variational Problems.* North-Holland Press.

[17] Fiacco, A. and J. Kyparisis. 1986. Convexity and concavity properties of optimal value functions in parameteric nonlinear programming. *Journal of Optimization Theory and Applications.* 95-126.

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

[19] J. Gauvin, and F. Dubeau, Differential properties of the marginal function in mathematical programming, Math. Programming Stud. 19 (1982) 101-119.

[20] J. Gauvin, J.W. Tolle, Differential stability in nonlinear programming, SIAM Journal of Control and Optimization 15 (1977) 294-311.

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

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

[23] K. Hinderer, Lipschitz continuity of value functions in Markovian decision processes, Math. Methods Operations Res. 62 (2005) 3–22.

[24] H. Hopenhayn, E. Prescott, Stochastic monotonicity and stationary distributions for dynamic economies, Econometrica 60 (1992) 1387-1406.

[25] Kamihigashi, T. 2007. Stochastic optimal growth with bounded and unbounded utility and bounded and unbounded shocks. *Journal of Mathematical Economics.* 43. 477-500.

[26] J. Kelley, General Topology, Van Nostrand, 1955.

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

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

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

[30] M. Li Calzi, A. Veinott, Jr, Subextremal functions and lattice programming, MS, Stanford University 1991.

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

[32] Mirman, L. 1970. *Two essays on Uncertainty and Economics.* Phd thesis. University of Rochester.

[33] L. Mirman, I. Zilcha, On optimal growth under uncertainty, Journal of Econ. Theory 11 (1975) 329-339.

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

[35] O. Morand, K. Reffett, S. Tarafdar, Duality and optimization in economies with nonconvexities, MS, Arizona State University, 2010a.

[36] O. Morand, K. Reffett, S. Tarafdar, Recursive aggregation, duality, and constrained lattice programming, MS, Arizona State University, 2010b.

[37] O. Morand, K. Reffett, S. Tarafdar, Nonsmooth methods in Lipschitzian stochastic dynamic programming, MS, Arizona State University, 2011.

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

[39] Nishimura, K. and J. Stachurski. 2005. Stability of stochastic growth models. *Journal of Economic Theory.* 122. 100-118.

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

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

[42] R.T. Rockafellar, Convex Analysis, Princeton, 1970.

[43] P. Samuelson, Foundations of Economic Analysis, Cambridge, 1947.

[44] M. Santos, J. Vigo-Aguirar, Analysis of a numerical dynamic programming algorithm applied to economic models, Econometrica, 66 (1998) 409-426.

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

[46] D. Topkis, Supermodularity and Complementarity, Princeton 1998.

[47] Trach, P.T. 1991. Quasiconjugates of functions, duality relationship between quasiconvex minimization under reverse convex constraint and quasiconcave maximization under a convex constraint, and applications. *Journal of Mathematical Analysis and Applications,* 159, 299-322.

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

[49] J. Viner, Cost curves and supply curves. Zeitschrift fur Nationalokonomie 3, 1931: Reprinted in Readings in price theory. Homewood, Il. Richard D. Irwin, 1951