The B.E. Journal of Theoretical Economics

Topics

Volume 7, *Issue* 1 2007 *Article* 35

Monotone Comparative Statics in Ordered Vector Spaces

Martin K. Jensen*

*University of Birmingham, m.k.jensen@bham.ac.uk

Recommended Citation
Martin K. Jensen (2007) "Monotone Comparative Statics in Ordered Vector Spaces," *The B.E. Journal of Theoretical Economics*: Vol. 7: Iss. 1 (Topics), Article 35.
Available at: http://www.bepress.com/bejte/vol7/iss1/art35

Monotone Comparative Statics in Ordered Vector Spaces*

Martin K. Jensen

Abstract

This paper considers ordered vector spaces with arbitrary closed cones and establishes a number of characterization results with applications to monotone comparative statics (Topkis (1978), Topkis (1998), Milgrom and Shannon (1994)). By appealing to the fundamental theorem of calculus for the Henstock-Kurzweil integral, we generalize existing results on increasing differences and supermodularity for C^1 or C^2 functions. None of the results are based on the assumption that the order is Euclidean. As applications we consider a teamwork game and a monopoly union model.

KEYWORDS: increasing differences, supermodularity, ordered vector space, vector lattice, comparative statics, non-smooth analysis, Henstock-Kurzweil integration

*Martin Kaae Jensen, Department of Economics, 90 Vincent Drive, University of Birmingham, Edgbaston, Birmingham B15 2TT, Birmingham, UK. Thanks are due to Chris Shannon and an anonymous referee for their comments and suggestions. Any remaining errors are the responsibility of the author.

1 Introduction

There is a large literature on monotone comparative statics and such assumptions on parameterized objective functions as *supermodularity* and *increasing differences* are quickly becoming standard. The strength of these assumptions is their intuitive appeal and, of course, the method of which they are a part. Their weakness is that they are not always easy to check in concrete applications. As a matter of fact, the only really powerful tools toward this end are the results of Topkis (1978) concerning C^1- or C^2-smooth functions defined on $\mathbb{R}^N$ with the Euclidean product order. But how does one check these assumptions if the order is not Euclidean or the function is not sufficiently smooth ? This and related questions have motivated this paper.

Specifically, the paper considers finite dimensional ordered vector spaces and establishes characterization results similar to Topkis', but more general in two respects: they relax the smoothness conditions and they hold for arbitrary closed vector orders.[1]

The generalization to arbitrary vector orders is useful whenever an objective function's domain is not a lattice/sublattice with respect to the Euclidean order. An example is the first application in section 5 which considers a teamwork game where the team's members have multiple tasks. The agents' choice sets are not Euclidean sublattices. But using results from section 2, a complete characterization of those orders for which they are, is obtained. Conditions for increasing differences and supermodularity can then be established.[2]

The generalization to non-smooth functions of Topkis' results turns out to be particularly easy to apply for two classes of functions: One is nearly everywhere differentiable functions, *i.e.*, functions which are differentiable at all except, perhaps, at most a countable number of points. The other class consists of those functions which are Lipschitz continuous and therefore admit a generalized derivative in the sense of Clarke (1983). A typical application within the first class is in situations where Inada-type boundary conditions are imposed on objective functions (see section 3.2). The second application in section 5 (a monopoly union model) illustrates the results on functions with generalized derivatives. The paper also contains a much more general result which applies to functions which are Henstock-Kurzweil integrable (theorem

[1] A closed vector order is a vector order whose positive cone is closed. See section 2 for mathematical preliminaries.

[2] Of course the motivation for this whole exercise would be to establish existence of a pure strategy Nash equilibrium via Tarski's fixed point theorem or to study comparative statics. In this paper, no attention is given to these issues (see Topkis (1998) for a comprehensive treatment).

4). The weakness of this result is that its generality makes it difficult to apply.[3] On the other hand, the introduction of the Henstock-Kurzweil integral plays a key role in the proofs of several other results and may be of some interest in itself.

The structure of the paper is as follows: Section 2 contains results on vector orders and vector lattices. Section 3 contains results on increasing differences. It begins with some notation and an overview and then splits into three subsections each of which can be read independently of the others. The first contains theorem 4 mentioned above, as well as the necessary prerequisites on Henstock-Kurzweil integration. The second looks at nearly everywhere differentiable functions. The third concerns functions which are locally Lipschitz (which is the assumption underlying Clarke calculus, see Clarke (1983)). Section 4 has but one purpose: To establish results for supermodularity which parallel those on increasing differences. The reader may jump directly to theorem 8 in this section, although the section does contain some other results which specialists in the field may find interesting. Finally, section 5 contains the two applications.

2 Vector Orders and Vector Lattices

Let $X = (\mathbb{R}^N, \succeq_X)$ and $T = (\mathbb{R}^M, \succeq_T)$ be ordered vector spaces with positive cones $X_+ \subset \mathbb{R}^N$ and $T_+ \subset \mathbb{R}^M$.[4] A function $f : X \to T$ is *order-preserving* if $f(x') \succeq_T f(x)$ whenever $x' \succeq_X x$. In the literature order-preserving functions are also called isotone, monotone, and increasing. A real-valued function defined on the product of X and T, $f : X \times T \to \mathbb{R}$ has *increasing differences in* (x, t) *on* $X \times T$ if for all $t' \succeq_T t$, the function $f(\cdot, t') - f(\cdot, t) : X \to (\mathbb{R}, \geq)$ is order-preserving. The set of order-preserving functions forms a convex cone

[3] The route to any characterization result for non-smooth functions is in principle straightforward: Use the fundamental theorem of calculus in one of its forms to recover the function from "what exists of its derivative". This leads to a statement such as our theorem 4, which, to be sure, does establish both necessary and sufficient conditions for a non-smooth function to have increasing differences. The problem is that any such result just lead to a new question which is often more difficult to answer: How does one actually check that the resulting conditions ? It is this second question which we are able to give a satisfactory answer to for nearly everywhere differentiable and Lipschitz continuous functions.

[4] The positive cone of X, say, is the set $X_+ = \{x \in \mathbb{R}^N : x \succeq_X 0\}$. That $\succeq_X$ is a (partial) vector order implies that X_+ is a proper cone, *i.e.*, a convex pointed cone (a cone is pointed if $-x, x \in X_+ \Rightarrow x = 0$). Conversely let $X_+ \subset \mathbb{R}^N$ be any proper cone. Then X_+ defines a vector order in $\mathbb{R}^N$ by virtue of "$x \succeq_X y \Leftrightarrow x - y \in X_+$". Thus in a vector space there is a one-to-one correspondence between the proper cones and the set of vector orders.

in the vector space, T^X, of all mappings from X to T. It is furthermore a *pointwise closed* cone provided that T_+ is (norm) closed in T:

Lemma 1 *Let $(f_n)_{n=1}^{\infty}$, $f_n : X \to T$, be a sequence of order-preserving functions which converges pointwise to a function $f : X \to T$. If T_+ is closed and f_n is order-preserving for every $n \in \mathbf{N}$, then the pointwise limit f is order-preserving.*

Proof: The result is immediate in light of the fact that $f_n(y) \succeq_T f_n(x) \Leftrightarrow f_n(y) - f_n(x) \in T_+$ since a pointwise convergent sequence must have its limit point in T_+ when T_+ is closed. $\square$

Remark 2.1 *In the previous lemma it is sufficient that X be an ordered set.*

A partially ordered set is *directed* if every two-element subset has an upper bound. For general vector spaces, X_+ directs X if and only if every $x \in X$ can be written $x = y - z$ where $y, z \in X_+$ (*i.e.*, X_+ is generating, cf. Schaefer (1999), chapter 15.1). When X is finite dimensional, it is directed under a vector order if and only if the positive cone of the order has non-empty interior (Birkhoff (1967), chapter 15, theorem 8). A partially ordered set is *conditionally complete* if every subset which has an upper bound has a supremum. A *vector lattice* is an ordered vector space in which every two element subset has a supremum.[5] It follows that a conditionally complete ordered vector space is a vector lattice if and only if it is directed. The next result, whose proof can be pieced together from results in Birkhoff (1967), chapter 15, shows when a vector lattice is conditionally complete.

Lemma 2 *If X is a vector lattice then it is conditionally complete if and only if the positive cone X_+ is closed.*

Let $(p_1, \ldots, p_P)$, $P \in \mathbf{N}$, be an finite sequence of vectors in X. The *conic hull* of such a sequence of vectors defines a closed, convex cone: $X_+ = \mathrm{cone}(p_1, \ldots, p_P) \equiv \{x \in X : x = \sum_{i=1}^{P} \lambda_i p_i,\ \lambda_i \geq 0,\ i = 1, \ldots, P\}$. Arranging the vectors in a matrix $\mathbf{Q} = [p_1, \ldots, p_P] \in \mathbb{R}^{N \times P}$ we say that $\mathbf{Q}$ *generates the cone* X_+, or that X_+ *has matrix representation* $\mathbf{Q}$, and write simply $X_+ = \mathrm{cone}(\mathbf{Q})$.

For example let $X = \mathbb{R}^2$ and $\mathbf{Q}$ an arbitrary full rank 2 by 2 matrix. Then $\mathbf{Q}$ generates a convex cone in X, $X_+ = \mathrm{cone}(\mathbf{Q})$, which is closed and directs

[5] Equivalently, that every two-element subset has an infimum, see Schaefer (1999), pp. 209.

$\mathbb{R}^2$. In two dimensions, it is clear that all directing cones must be of this form, and it is straight-forward to see that they turn $\mathbb{R}^2$ into a vector lattice (which, in fact, must be conditionally complete by lemma 2).

The observations in the previous paragraph remain valid also in higher dimensions. This is surprising because in three or more dimensions, a convex cone need not be a polyhedron and so it may not have a matrix representation. The result is an immediate consequence of the so-called *Choquet-Kendall theorem*. The Choquet-Kendall theorem says, when $X = \mathbb{R}^N$ and X_+ is closed and generating, that X will be a vector lattice if and only if it has a basis B which is a simplex of full dimension $N-1$ (see *e.g.* Peressini (1967), proposition 3.11.).[6]

Theorem 1 *Assume that X_+ is closed. Then $(\mathbb{R}^N, \succeq_X)$ is a vector lattice if and only if X_+ has a full rank N by N matrix representation $\mathbf{Q}$.*

Proof: "$\Rightarrow$": Since X_+ is closed, the vector lattice must be conditionally complete by lemma 2. So X_+ directs X or equivalently, it is generating. The statement now follows from the Croquet-Kendall theorem by taking as $\mathbf{Q}$'s column vectors the vertices of the simplex which is a basis for X_+. "$\Leftarrow$": A cone with a full rank matrix representation is of course both closed and generating, so again the conclusion follows from the Croquet-Kendall theorem. $\square$

In many applications, the choice set is not a vector lattice but an *interval* $\mathcal{I} = \{x \in \mathbb{R}^N : \underline{y} \leq x \leq \overline{y}\}$, where $\underline{y}^n = -\infty$ and $\overline{y}^m = +\infty$ are allowed for some or all coordinates. For example *smooth supermodular games* as defined in Milgrom and Roberts (1990) have joint strategy sets which are compact intervals. While theorem 1 provides conditions under which $(\mathbb{R}^N, \succeq_X)$ is a vector lattice, the following result further restricts the matrix representation such that such intervals become sublattices. While everything said so far has been standard, this result is new.

Theorem 2 *Assume that $(\mathbb{R}^N, \succeq_X)$ is a vector lattice with matrix representation $\mathbf{Q} \in \mathbb{R}^{N \times N}$ and let $\mathcal{I}$ be an interval as defined above. Then $\mathcal{I}$ will be a sublattice of $(\mathbb{R}^N, \succeq_X)$ if and only if $\mathbf{Q}$ has at most two non-zero entries in every row and the product of any two such non-zero row entries is non-positive.*

Proof: *Denote the inverse of the matrix representation by $\mathbf{Q}^{-1}$ (such an inverse exists since $\mathbf{Q}$ has full rank by theorem 1). The class of linear bijections $\mathbf{Q} : \mathbb{R}^N \to \mathbb{R}^N$, is the unique class that maps a conditionally complete vector lattice, lattice isomorphically into a Euclidean lattice. Using this*

[6] A subset $B \subseteq X$ is a *basis* for the cone $X_+ \subseteq X$ if every $x \in X_+ \backslash \{0\}$ has a unique representation of the form $x = \lambda b$, $\lambda > 0$, $b \in B$.

it is straight-forward to show that a subset $S \subseteq \mathbb{R}^N$ *will be a sublattice of* X *if and only if* $\mathbf{Q}^{-1}(S) = \{y \in \mathbb{R}^N : y = \mathbf{Q}^{-1}x, x \in S\}$ *is a sublattice of* $(\mathbb{R}^N, \geq_N)$ *(a Euclidean sublattice).*[7] *Let* S *be an interval and consider* $\mathbf{Q}^{-1}(S) = \{y \in \mathbb{R}^2 : \overline{y} \geq_N \mathbf{Q}y \geq_N \underline{y}\}$ *(if the interval is open, simply delete the corresponding inequality) which is seen to be a set of solutions to a system of linear inequalities, i.e., a polyhedron. Conveniently, a polyhedron in* $\mathbb{R}^N$ *is a Euclidean sublattice if and only if it is the solution set of a finite number of linear bimonotone inequalities (Veinott (1989)). A linear inequality* $(d_1, \ldots, d_N) \cdot (x_1, \ldots, x_N)^T \geq 0$ *is bimonotone if* $d = (d_1, \ldots, d_N)$ *has at most two non-zero entries, say* d_i *and* d_j*, and the product of these is non-positive,* $d_i d_j \leq 0$*. This implies the conclusion of the theorem.* □

The idea in the proof of theorem 2 works more generally and in section 5.1 an example is provided where the choice set is not an interval.

3 Increasing Differences

From the definition of increasing differences (section 2) it is seen that when T is an ordered vector space with positive cone T_+, $f : X \times T \to \mathbb{R}$ will have increasing differences if and only if,

$$\frac{f(x, t + hd) - f(x, t)}{h} \tag{1}$$

is order-preserving in x for all $h > 0$, $d \in T_+$, and $t \in T$. If for some *direction* in the positive cone $d \in T_+$,

$$f^+_{t,d}(x) := \lim_{h \downarrow 0} \frac{f(x, t + hd) - f(x, t)}{h} \tag{2}$$

exists *and* T_+ is closed, then this (one-sided) *directional derivative* must consequently be order-preserving in x by lemma 1 (and at the level of generality

[7] Let $f : \mathbb{R}^N \to \mathbb{R}$ and $X = (\mathbb{R}^N, \succeq_X)$ a partially ordered set. Consider a bijective mapping $h : \mathbb{R}^N \to \mathbb{R}^N$. If $f(h^{-1}(y))$ is twice differentiable in y with non-negative off-diagonal elements, we *cannot* conclude that f is supermodular in $x = h^{-1}(y)$ in some appropriate order $\succeq_X$, *unless* h is of the form $h(x) = \mathbf{Q}x$ (in which case we also know $\succeq_X$). This points to a serious pitfall when considering transformations: Increasing differences only implies supermodularity on a vector lattice (perhaps as model of a collection of chains), and - unless h is a linear bijection - X will not be a vector lattice in any order corresponding to a transformation. One should therefore make a clear separation between lattice isomorphic transformations and transformations of variables more generally.

marked by remark 2.1). If in (2), one can replaced $h \downarrow 0$ by $h \to 0$ to obtain the usual two-sided directional derivative, which throughout this paper is denoted $f'_{t,d}(x)$, this map must consequently also be order-preserving in x. If f is (Gâteaux) *differentiable* at the point (x,t), $f'_{t,d}(x) = f^+_{t,d}(x) = d^T D_t f(x,t)$ for *all* $d \in T_+$ where $D_t f(x,t)$ is the derivative.[8] Thus, if f is differentiable at (x,t) and has increasing differences, then - whenever it is defined - the evaluation $d^t D_t f(x,t)$ must be order-preserving in x for all $d \in T_+$, *i.e.*, $d^T D_t f(\cdot,t) : X \to (\mathbb{R}, \geq)$ must be order-preserving. Consider now T_+'s matrix representation $\mathbf{P} = [p_1, \ldots, p_P] \in \mathbb{R}^{M \times P}$. The requirement that $d^T D_t f(x,t)$ must be order-preserving in x for every direction $d \in T_+$ now translates into the very simple statement that $\mathbf{P}^T D_t f(\cdot,t) : X \to (\mathbb{R}^P, \geq_P)$ is an order-preserving mapping. If, in fact, $D_t f(x,t)$ exists everywhere (f is everywhere differentiable) then this is the same as saying that $\mathbf{P}^T[D_t f(x',t) - D_t f(x,t)] \geq_P 0$ for all $x' \succeq_X x$.

To sum up, whenever a function has one- or two-sided derivatives at a point, these must be order-preserving in x. This requirement is, however, not *sufficient* in general. To find sufficient conditions we must turn the previous argument around and recover the function f from its derivative (suitably generalized when f is not sufficiently smooth).

3.1 Increasing Differences with ACG Functions

The main result of this subsection (theorem 4) provides a result which characterizes increasing differences for a very general class of functions known as *ACG* (which is short for *generalized absolutely continuous in the restricted sense*). In the following two subsections several results will be presented which are either special cases or immediate consequences. Since the generality of theorem 4 will not be used outside the proofs, the non-mathematically minded reader may wish to skip this subsection entirely in a first reading.

Recall that $f : [a,b] \to \mathbb{R}$ is *absolutely continuous*, if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $\sum_{i=1}^I a_i - b_i < \delta \Rightarrow \sum_{i=1}^I |f(a_i) - f(b_i)| < \epsilon$ for any finite sequence of disjoint open intervals $]a_i, b_i[\subset [a,b]$, $i = 1, \ldots, I$. If under the same conditions one demands that: $\sum_{i=1}^I \sup_{u,v \in [a_i,b_i]} |f(u) - f(v)| < \epsilon$, f is ACG* (*absolutely continuous in the restricted sense*). Finally, f is ACG (*generalized absolutely continuous in the restricted sense*), if it is continuous and $[a,b]$ is equal to a countable union of subintervals upon each of which f is

[8] $D_t f(x,t)$ and the continuous linear mapping it defines: $d : D_t f(x,t) \mapsto d^T D_t f(x,t)$ are both referred to here as the *derivative* of f at (x,t). Although this terminology is abusive we shall extend its usage to include matrices which are identified both as a collection of numbers and as linear mappings.

ACG* (cf. Talvila (2001)). Any absolutely continuous function is ACG. Since any Lipschitz continuous function is absolutely continuous, and any concave (or convex) function is Lipschitz on the interior of its domain, the class of ACG function is very broad and accommodates in particular the non-smooth function classes studied by Clarke (1983) as we will be using in section 3.1. The class of ACG functions supplies what to the best of the author's knowledge, is the literature's most general formulation of the fundamental theorem of calculus. Before presenting that result we need to introduce the *Henstock-Kurzweil integral*. Let $I = [a, b] \subset \mathbb{R}$ be a compact interval. A *tagged partition* of I, $T = ([x_{i-1}, x_i], t_i)_{i=1}^n$ is a finite set of order pairs such that the intervals $I_i = [x_{i-1}, x_i]$, $i = 1, \dots, n$ forms a (disjoint) partition of I and $t_i \in I_i$ for all i. The *Riemann sum* of a function $f : I \to \mathbb{R}$ under the tagged partition T is $S(f, T) = \sum_{i=1}^n f(t_i)(x_i - x_{i-1})$. For any strictly positive function $\delta : I \to \mathbb{R}_{++}$, a tagged partition is said to be δ-*fine* if $0 < x_i - x_{i-1} \leq \delta(t_i)$ for all i.

Definition (Henstock-Kurzweil Integral) *Let $I = [a, b]$ be an interval and $f : I \to \mathbb{R}$ a function. f is said to be Henstock-Kurzweil integrable on I with Henstock-Kurzweil integral $\int_a^b f(x) \in \mathbb{R}$, if for every $\epsilon > 0$ there exists a function $\delta_\epsilon : I \to \mathbb{R}_{++}$ such that,*

$$|S(f, T) - \int_a^b f(x)| \leq \epsilon \tag{3}$$

for any tagged partition of I which is δ_ϵ-fine.

The simplicity of this definition compared to that of the *Lebesgue integral* is considerable.[9] Nonetheless, the Henstock-Kurzweil integral is actually *more* general than the Lebesgue integral (see remark 3.1 below). We are now ready to state the fundamental theorem of calculus (for the proof see Gordon (1994)).

Theorem 3 (Fundamental Theorem of Calculus for the Henstock-Kurzweil integral) *A function $f : [a, b] \to \mathbb{R}$ is* ACG *if and only if it is differentiable almost everywhere and there exists a Henstock-Kurzweil integrable function $g : [a, b] \to \mathbb{R}$ with $f'(x) = g(x)$, a.e., such that $f(x) = f(a) + \int_a^x g(s)$ for all $x \in [a, b]$.*

[9] Observe that if we replace the function $\delta : I \to \mathbb{R}_{++}$ in the definition with a constant $\delta > 0$ we get the Riemann integral. This is why the Henstock-Kurzweil integral is sometimes also called the *generalized Riemann integral*.

Remark 3.1 *The fundamental theorem of calculus for the Lebesgue-integral is a special case of theorem 3 since a function $f : [a, b] \to \mathbb{R}$ is Lebesgue integrable if and only if both f and $|f|$ are Henstock-Kurzweil integrable (cf. Bartle (1996)).*

For the next result, we recall from above that $f'_{t,d}(x)$ denotes the usual (two-sided) directional derivative of f taken with respect to t in the direction $d \in T$.

Theorem 4 *Assume that for all $(x, t) \in X \times T$ and $d \in T_+$, $f(x, t + hd)$ is* ACG *in h. Then $f : X \times T \to \mathbb{R}$ has increasing differences in (x, t) if and only if for all $d \in T_+$ and $t \in T$: $f'_{t,d}(y) - f'_{t,d}(x) \geq 0$ whenever $y \succeq_X x$ and both $f'_{t,d}(y)$ and $f'_{t,d}(x)$ exist.*

Proof: *"⇒": Proved at the beginning of section 3. "⇐" Let $f_{x,t,d}(h) := f(x, t + hd)$ and when this function is differentiable denote the derivative $f'_{x,t,d}(h)$. Note that $f'_{t,d}(x) = f'_{x,t,d}(0)$ in terms of the notation in the main body of the paper. f will have increasing differences if and only if:*

$$\frac{f_{x,t,d}(h) - f_{x,t,d}(0)}{h} \tag{4}$$

is order-preserving in x for all $(h, d, t) \in \mathbb{R}_{++} \times T_+ \times T$. Fix $(h, d, t) \in \mathbb{R}_{++} \times T_+ \times T$ and $x \in X$. By assumption $f_{x,t,d}(s)$ is ACG in s, in particular it is ACG in s on $[0, h]$. Hence by the fundamental theorem of calculus for the Henstock-Kurzweil integral there exists a (Henstock-Kurzweil) integrable function $g_{x,t,d} : [0, h] \to \mathbb{R}$ such that $f'_{x,t,d}(s) = g_{x,t,d}(s)$ for almost every $s \in [0, h]$, and $f_{x,t,d}(h) - f_{x,t,d}(0) = \int_0^h g_{x,t,d}(s)$. Since the previous argument is valid for arbitrary $x \in X$ and since the Henstock-Kurzweil integral is a linear operator, it follows that (4) will be order-preserving in x if and only if:

$$\frac{\int_0^h (g_{y,t,d}(s) - g_{x,t,d}(s))}{h} \geq 0 \tag{5}$$

whenever $y \succeq_x x$. By assumption $f'_{t+sd,d}(y) - f'_{t+sd,d}(x) \geq 0$ whenever these terms are well-defined. Equivalently $f'_{y,t,d}(s) - f'_{x,t,d}(s) \geq 0$ (the chain rule), which consequently holds for almost every $s \in [0, h]$ since an ACG function is differentiable almost everywhere. But then $g_{y,t,d}(s) - g_{x,t,d}(s) \geq 0$ for almost every $s \in [0, h]$ which implies (5) . Since $h > 0$, $d \in T_+$, and $t \in T$ were picked arbitrarily, this finishes the proof. □

Theorem 4 provides a general approach to characterizing increasing differences for non-smooth functions: Directionally ACG functions are differentiable almost everywhere in any given direction. Any sound notion of a *generalized derivative* for directionally ACG functions will consequently be single-valued almost everywhere, and by theorem 4 it is *only* at the differentiability points we need to check that the derivative is order preserving. This observation will be used to characterize increasing differences in the next two subsections.

3.2 The Nearly Everywhere Differentiable Case

A function is said to be differentiable *nearly everywhere* if it is differentiable in all except, perhaps, an at most countable number of points. Note that this statement makes no demand on part of the derivative such as continuity, boundedness, or agreement almost everywhere with a Lebesgue integrable function. On the other hand it is clearly a more restrictive assumption than differentiability *almost* everywhere with respect to the Lebesgue measure.

Our first characterization result is a direct generalization of the results concerning C^1 and C^2 functions on product spaces with the Euclidean order mentioned in the introduction (these results are normally attributed to Topkis (1978)).[10]

Theorem 5 *Let $X = (\mathbb{R}^N, \succeq_X)$ and $T = (\mathbb{R}^M, \succeq_T)$ be ordered vector spaces with closed positive cones generated by the matrices $\mathbf{Q} \in \mathbb{R}^{N\times Q}$ and $\mathbf{P} \in \mathbb{R}^{M\times P}$.*

1. *Assume that for all $x \in X$, $f : X \times T \to \mathbb{R}$ is continuous and differentiable in t nearly everywhere. Then $f : X \times T \to \mathbb{R}$ will have increasing differences in (x, t) on $X \times T$ if and only if $\mathbf{P}^T[D_t f(x', t) - D_t f(x, t)] \geq_P 0$ whenever $x' - x \in \mathrm{cone}(\mathbf{Q})$ and the concerned derivatives exist.*

2. *Assume that f is continuously differentiable and has well-defined second order cross-derivatives $D^2_{tx} f(x, t)$ nearly everywhere, i.e., for all $(x, t) \in X \times T$ except for an at most countable number of points. Then f has increasing differences in (x, t) on $X \times T$ if and only if the P by Q matrix,*

$$\mathbf{P}^T D^2_{tx} f(x, t)\mathbf{Q}$$

is non-negative for all $(x, t) \in \{(x, t) \in X \times T : D^2_{tx} f(x, t)$ exists$\}$.

[10] Topkis (1978) actually announces the results under the assumption that the function is merely differentiable (respectively, twice differentiable). Milgrom and Shannon (1994) refer to Topkis (*ibid.*) for the second-order result under the assumption of C^2-smoothness (theorem 6). Ironically, the results are in fact valid as stated by Topkis (1978) since a differentiable (not necessarily C^1!) function is nearly everywhere differentiable. If, in fact, Topkis (1978) refers to results which are part of the folklore, the results presented here obviously cannot claim originality in the everywhere differentiable case.

Proof: *Define $f_{x,t,d}(h) = f(x, t + hd)$, $f_{x,t,d} : \mathbb{R} \to \mathbb{R}$, and note that since f is differentiable nearly everywhere in t, $f_{x,t,d}$ is differentiable nearly everywhere in h for all $d \in T_+$. Therefore $f_{x,t,d}$ is generalized absolutely continuous in the restricted sense (cf. section 3.1), hence there exists a Henstock-Kurzweil integrable function $g_{x,t,d} : \mathbb{R} \to \mathbb{R}$, such that $f'_{x,t,d}(h) = g_{x,t,d}(h)$ for almost every $h \in [a, b]$, all $a, b \in \mathbb{R}$. The result now follows from theorem 4 below since* **P**'s *column vectors generate T_+ and the derivative is a linear mapping. The second part of theorem 5 is, of course, just the second order formulation that arises from changing the basis according to* **P** *and* **Q**. *Indeed, it is clear from the definition that if f is twice cross-differentiable at (x, t), then it will have increasing differences locally at that point if and only if,*

$$\frac{(f(x + \tilde{h}\tilde{d}, t + hd) - f(x + \tilde{h}\tilde{d}, t)) - (f(x, t + hd) - f(x, t))}{h\tilde{h}} \geq 0 \tag{6}$$

for all $h > 0$, $\tilde{h} > 0$, $\tilde{d} \in X_+$, $d \in T_+$, and $(x, t) \in X \times T$. If for some $(x, t) \in X \times T$, the limit exists as $h, \tilde{h} \to 0$ the resulting second order directional derivative, $f''_{d,\tilde{d}}(x, t)$ must consequently be non-negative. Again, the existence of a second order derivative, $D^2_{tx}f(x, t)$, adds sufficient structure for the situation to become a simple one for then $f''_{d,\tilde{d}}(x, t) = d^T D^2_{tx}f(x, t)\tilde{d}$, for all $(\tilde{d}, d) \in X_+ \times T_+$. Hence the statement of the theorem.[11] □

Topkis' results concerning increasing differences for C^1 and C^2 functions are clearly special cases of theorem 5. Indeed, if we take $\mathbf{P} = \mathbf{I}_M$ and $\mathbf{Q} = \mathbf{I}_N$ (the identity matrices) then $T_+ = \text{cone}(\mathbf{I}_M) = \mathbb{R}^M_+$ and $X_+ = \text{cone}(\mathbf{I}_N) = \mathbb{R}^N_+$. Since $\mathbb{R}^M_+$ and $\mathbb{R}^N_+$ are the positive cones of the Euclidean orders, and $\mathbf{P}^T D^2_{tx}f(x, t)\mathbf{Q} = D^2_{tx}f(x, t)$ in this case, Topkis' result arrives under weakened smoothness assumptions.

A straight-forward application of theorem 5 is to objectives which satisfy boundary conditions. In many situations (Cournot oligopoly, the Bertrand model, market games, etc.), agents are assumed to have choice set $\mathbb{R}_+$ or $[0, \overline{y}] \subset \mathbb{R}_+$. Often one wishes to rule out trivial equilibria and does this by imposing boundary conditions of Inada type (the first derivate approaches infinity as $x \to 0$ in the choice set). Obviously, a function satisfying such a boundary condition is not C^1 and so Topkis' characterization results do not apply. But theorem 5 does. For example take $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$; $f(x, t) = x^t$.

[11] A second order derivative is generally an operator-norm continuous bilinear mapping. As with $D_t f(x, t)$ no distinction will be made between the mapping and the evaluation defining it, however.

f is not differentiable in t nearly everywhere *for all* $x \in \mathbb{R}_+$ ($D_t f(x,t)$ exists for no $t \in \mathbb{R}_+$ if $x = 0$). However, reversing the roles of x and t we see that *for all* $t \in \mathbb{R}_+$, $f(x,t)$ is continuous and differentiable nearly everywhere in x ($D_x f(x,t)$ exists for all $x > 0$). Consequently theorem 5 applies. There exist exactly two proper cones which define vector orders upon $\mathbb{R}$, namely $X_+ = \mathbb{R}_+$ and $X_+ = \mathbb{R}_-(:= -\mathbb{R}_+)$. The former has matrix representation $\mathbf{P} = 1$, the latter $\mathbf{P} = -1$. By comparison with the condition of theorem 5, for f to have increasing differences either $t' x^{t'-1} - t x^{t-1} \geq 0$ all $t' > t$ ($\mathbf{P} = 1$) or $t' x^{t'-1} - t x^{t-1} \leq 0$ all $t' > t$ ($\mathbf{P} = -1$), $x > 0$. As may be checked the former will be the case if $(x,t) \in S = \{(a,b) \in \mathbb{R}_{++} \times \mathbb{R}_+ : b \ln a \geq -1\}$ while the latter will be the case otherwise.

3.3 The Locally Lipschitz-continuous Case

Convex or concave functions need of course not be differentiable, but they never the less possess very nice "pseudo-differentiability" properties which have been exploited thoroughly in economic analysis. First, convex functions admit one-sided directional derivatives everywhere, so $f^+_{t,d}(x)$ (cf. (2)) is always a well-defined finite quantity. Secondly, a bounded convex function is locally Lipschitz on the interior of its domain (Roberts and Varberg (1974)) and so is differentiable almost everywhere by Rademacher's theorem. To be more specific, let $f(x,t)$ be convex in t. $v \in \mathbb{R}^M$ is a *subgradient* of $f(x,t)$ with respect to t at a point (x_0, t_0) if:[12]

$$f(x_0, t) \geq f(x_0, t_0) + (t - t_0)^T v \text{ , for all } t \in \mathbb{R}^M \tag{7}$$

The set of subgradients at (x_0, t_0) is called the *subdifferential* and denoted $\partial_t f(x,t)$.

For locally Lipschitz continuous functions more generally, a slight weakening of the concept of directional derivative leads to a closely related construction. Moreover, it will allow us to exploit theorem 4 to obtain a very satisfactory result characterizing increasing differences. What we have in mind is the *generalized directional derivative* of Clarke (1983):

$$f^\circ_{t,d}(x) := \lim_{h \to 0+} \sup_{s \to t} \frac{f(x, s + hd) - f(x,s)}{h} \tag{8}$$

When f is locally Lipschitz of rank $k > 0$ at a point t, the difference quotient

[12] If f is instead concave in t, we reverse the inequality-sign in (7) and still refer to v as a subgradient.

whose supremum enters in this definition is (locally) bounded from above by $k\left\|d\right\|_T$. Hence $f^{\circ}_{t,d}(x)$ is well-defined. If f is convex in t, then $f^{\circ}_{t,d}(x) = f^{+}_{t,d}(x)$.

If T_+ is closed, increasing differences and lemma 1 imply that $f^{\circ}_{t,d}(x)$ must be order-preserving in x for every $t \in X$ and $d \in T_+$. The converse is true as an immediate consequence of theorem 4 for if f is locally Lipschitz in t then $f_{x,t,d}(h) := f(x, t + hd)$ is locally Lipschitz in h and therefore absolutely continuous in h. Never the less, this is not a very applicable result and in fact we can do much better than that. $f^{\circ}_{t,d}(x)$ will be upper semicontinuous in (t, d) for every fixed $x \in X$ and as a function of d alone, $f^{\circ}_{t,d}(x)$ will be positively homogenous and subadditive (Clarke (1983), proposition 2.1.1.). This allows one to define the *generalized derivative*:

$$(9) \qquad \partial_t f(t, x) := \{v \in \mathbb{R}^M : f^{\circ}_{t,d}(x) \geq d^T v \text{ for all } d \in T\} \subset \mathbb{R}^M$$

As it turns out (Clarke (1983), proposition 2.2.7), if f is convex and bounded, the subdifferential and the generalized derivative coincide - hence the present notation.

$\partial_t f(t, x)$ will be a convex, non-empty subset of $\mathbb{R}^M$; if it is a unique element it coincides with the usual derivate $D_t f(t, x)$ (which therefore, in particular, exists). Since economic analysis provides us with a firm intuitive grip of subgradients, and the Lipschitz case does not upset this understanding, asking what conditions on the convex-valued multifunction $\partial_t f(t, \cdot) : X \to 2^{\mathbb{R}^M}$ will lead to increasing differences is perhaps the most obvious question one can fathom within the agenda of this paper (at least as far as increasing differences is concerned). Once again, the answer turns out to be a simple one:

Theorem 6 *Assume that $f : X \times T \to \mathbb{R}$ is locally Lipschitz continuous in $t \in T$ and that T_+ is closed with matrix representation $\mathbf{P} \in \mathbb{R}^{M\times P}$. Define*

$$\mathbf{P}^T \partial_t f(x, t) = \{q \in \mathbb{R}^P : q = \mathbf{P}^T v, v \in \partial_t f(x, t)\} \subseteq \mathbb{R}^P$$

whose section at $t \in T$ is a well-defined convex-valued multifunction, $\mathbf{P}^T \partial_t f(\cdot, t)$: $x \mapsto 2^{\mathbb{R}^P}$. $f(x, t)$ then has increasing differences in (x, t) if and only if for every $t \in T$ there exists an order-preserving selection $\xi(x) \in \mathbf{P}^T \partial_t f(x, t)$, all $x \in X$, i.e., $\xi : X \to \mathbb{R}^P$ such that $\xi(x') - \xi(x) \geq_P 0$ whenever $x' \succeq_X x$.

Proof: "$\Rightarrow$": For every $d \in T_+$, $f^{\circ}_{t,d}(x) = \max_{v\in\partial_t f(x,t)} d^T v$ and so is a selection from $d^T \partial_t f(x, t) = \{q \in \mathbb{R} : d^T v, v \in \partial_t f(x, t)\}$. Since $f^{\circ}_{t,d}(x)$ is order-preserving in x (T_+ is closed), $(f^{\circ}_{t,p_1}(x), f^{\circ}_{t,m_2}(x), \ldots, f^{\circ}_{t,p_P}(x))$ is then an order-preserving selection from $\mathbf{P}^T \partial_t f(x, t)$. "$\Leftarrow$": Let $\xi(x, t)$ be an arbitrary order-preserving selection from $\mathbf{P}^T \partial_t f(x, t)$ and consider the i'th coordinate,

$\xi_i(x,t)$, which is obviously an order-preserving selection from $p_i^T \partial_t f(x,t)$. Since f is locally Lipschitz in t, it is locally Lipschitz in every direction $d \in T$. In particular, it is absolutely continuous hence ACG, in every direction. Again from local Lipschitz continuity follows that f is differentiable almost everywhere in t (Rademacher's theorem). Consequently for any $x \in X$: $\xi_i(x,t) = f'_{t,p_i}(x)$ for every $t \in T$ where $f'_{t,p_i}(x)$ exists. The result now follows from theorem 4 since $\mathbf{P}$ generates T_+ and the derivative is a linear mapping when it exists. $\square$

Recall theorem 5, which in the nearly everywhere differentiable case requires that $\mathbf{P}^T D_t f(x,t)$ is order-preserving in x between every two points where the derivatives exist. One might conjecture that this situation can be cast in the same way as theorem 6, but this is generally not true. The reason is that unless f is locally Lipschitz in t, $\partial_t f(x,t)$ may well be empty at points of non-differentiability. Economically put, there may be points which do not have any vector of shadow prices associated with them. Of course we cannot then seek an order-preserving selection from the generalized derivative for it is not well-defined. It is important to realize that theorem 4 - which states that if f is merely directionally ACG in t, then we can recover f from "what exists of its derivative"- is typically too abstract to be applicable. It is easy to simply *assume* that a function is ACG (for example, absolutely continuous), but from a practical perspective how do we verify it and how do we verify the resulting conditions on the derivatives expressed in theorem 4 ? These are the difficult questions. By the results so far, we now know two cases where this can be done: The nearly everywhere differentiable case and the locally Lipschitz continuous case (both of which contain the continuously differentiable case as a special case).

4 Supermodularity

We now turn to establish a result by which one can establish *supermodularity* of $f(x,t)$ in x on X when X is not a Euclidean vector lattice. We shall need this result for our applications.

Recall that a real-valued function f defined on a lattice X is *supermodular* provided that $f(x \vee y) + f(x \wedge y) \geq f(x) + f(y)$ for all $x, y \in X$. Here $x \vee y$ (the *join*) is the supremum of x and y and $x \wedge y$ (the *meet*) is the infimum as defined above. Topkis (1978) proves the equivalence of supermodularity and increasing differences for a function $f : X \to \mathbb{R}$, where X is a finite dimensional product set $X = \times_{\alpha \in A} X_\alpha$, each X_α a chain in its order $\succeq_\alpha$, and X ordered by the product order $x \succeq_x x' \Leftrightarrow [x_\alpha \succeq_\alpha x'_\alpha$ for all $\alpha \in A]$. Obviously this definition cannot be used here because we consider vector spaces ordered by an arbitrary proper cone X_+.[13]

Say that x and y are *disjoint*, written $x \perp y$, if the infimum of x and y is zero, *i.e.*, if $x \wedge y = 0$ (cf. Schaefer (1999), chapter 5, section 1). If, for example, $X_+ = \mathbb{R}^N_+$ in $\mathbb{R}^N$, according to this definition $x \perp y$ if and only if $(\min\{x_1, y_1\}, \ldots, \min\{x_N, y_N\}) = (0, \ldots, 0)$, *i.e.*, $x^T y = 0$. So this is just the usual inner product definition of orthogonality.

Definition *Let X be an ordered vector space. A function $f : X \to \mathbb{R}$ is said to have increasing differences in x on X provided that for all $d \in X_+$, $f(x+d) - f(x)$ is order-preserving in x in all directions $\tilde{d}$ which are disjoint from d, i.e., if:*

$$f(x + d + \tilde{d}) - f(x + d) \geq f(x + \tilde{d}) - f(x) \tag{10}$$

for all $x \in X$ and all $d, \tilde{d} \in X_+$ with $d \perp \tilde{d}$.

The next result shows that the previous definition of increasing differences is in fact equivalent to supermodularity.

Theorem 7 *Assume that X is a vector lattice and consider a function $f : X \to \mathbb{R}$. Then f has increasing differences in x on X if and only if f is supermodular in x on X.*

Proof: "$\Rightarrow$": Pick $x, y \in X$, set $d = x - (x \wedge y)$, $\tilde{d} = y - (x \wedge y)$ and note that $d, \tilde{d} \in X_+$ and, since X is a vector lattice, $d \wedge \tilde{d} = (x - (x \wedge y)) \wedge (y - (x \wedge y)) = (x \wedge y) - (x \wedge y) = 0$. Since X is a vector lattice, the following identity is valid for all $x, y \in X$: $x + y = x \wedge y + x \vee y$ (cf. Birkhoff (1967), chapter 15, theorem 1). Substitute for x with $d + (x \wedge y)$ and for y with $\tilde{d} + (x \wedge y)$ to get: (†) $x \vee y = (x \wedge y) + d + \tilde{d}$. By assumption $f(a + d + \tilde{d}) - f(a + \tilde{d}) \geq f(a + d) - f(a)$ for $d, \tilde{d} \in X_+$ as chosen and every $a \in X$. Now simply take $a = x \wedge y$ and use (†) to see that: $f(x \vee y) - f(y) \geq f(x) - f(x \wedge y)$. "$\Leftarrow$": Pick $a \in X$ and $d, \tilde{d} \in X_+$ such that $d \perp \tilde{d}$. By supermodularity $f(x \vee y) + f(x \wedge y) \geq f(x) + f(y)$. Now pick $x = a + d$ and $y = a + \tilde{d}$. Clearly, $x \wedge y = a + (d \wedge \tilde{d}) = a$. Hence, $x \vee y = a + (d \vee \tilde{d}) = a + d + \tilde{d} - (d \wedge \tilde{d}) = a + d + \tilde{d}$. Insert to get: $f(a + d + \tilde{d}) + f(a) \geq f(a + d) + f(a + \tilde{d})$. □

The theorem above is true for functions defined on arbitrary sublattices of vector lattices (in particular the dimension need not be finite). In the present

[13] This definition worked in the previous section exactly because there we considered the product set $X \times T$.

case where X is an N-dimensional vector lattice with a closed cone, the result can be interpreted as a suitable change of basis. Indeed, if $x, x' \in \mathbb{R}^N$ are elements from a vector lattice X with a positive cone represented by $\mathbf{Q} = [q_1, \ldots, q_N]$, then,

$$x \wedge x' = \sum_{n=1}^{N} \min\{\alpha_n, \alpha'_n\} q_n \text{ and } x \vee x' = \sum_{n=1}^{N} \max\{\alpha_n, \alpha'_n\} q_n \tag{11}$$

where $(\alpha_1, \ldots, \alpha_N)$ are the *coordinates* of x in the basis for $\mathbb{R}^N$ determined by $\mathbf{Q}$'s columns, *i.e.*, the unique vector such that $x = \sum_{n=1}^{N} \alpha_n q_n$ (and similarly for $x' = \sum_{n=1}^{N} \alpha'_n q_n$). From this is seen that if $X = (\mathbb{R}^N, \geq_N)$, then a function $f : X \to \mathbb{R}$ has increasing differences in the sense of the definition above if and only if it has increasing differences in the sense of Topkis (1978) (see also Topkis (1998)). However, the definition is *not* equivalent to Topkis' definition on more general product ordered sets.[14] From (11) is also seen that when $\mathbf{Q}$ is a full rank matrix representation and consequently has an inverse $\mathbf{Q}^{-1}$, then the requirement that x and x' are disjoint, *i.e.*, $x \wedge y = 0$, is equivalent to $x^T(\mathbf{Q}^{-1})^T\mathbf{Q}^{-1}x' = 0$. If in particular the order is $\geq_N$ (Euclidean product order, $\mathbf{Q} = \mathbf{I}_N$), disjointness reads simply $x^T x' = 0$, hence as special cases of the results below we find the characterization of supermodularity given in Topkis (1978).

The next result, the proof of which is straight-forward in light of the previous observations, describes under what conditions a function f defined on a vector lattice will be supermodular. Taking $\mathbf{Q}$ as the identity matrix (so $\succeq_X$ is the usual order), we recover the well-known characterization result of Topkis (1978) mentioned in the previous section.

Theorem 8 *Consider a function $f : X \to \mathbb{R}$ where X is a conditionally complete vector lattice whose positive cone's (full rank) matrix representation is $\mathbf{Q} \in \mathbb{R}^{N \times N}$.*

1. *Assume that f is continuous and differentiable nearly everywhere. Then f will be supermodular in x on X if and only if $\mathbf{Q}^T[Df(x') - Df(x)] \geq_N 0$ whenever $x' - x \in \operatorname{cone}(\mathbf{Q})$, $x^T(\mathbf{Q}^{-1})^T\mathbf{Q}^{-1}x' = 0$, and the concerned derivatives exist.*

[14]This observation is in fact obvious in light of an example in Topkis (1978) of a product ordered set in which a function has increasing differences but is not supermodular. This function can consequently not have increasing differences according to our definition since this would contradict theorem 7.

2. *Assume that f is continuously differentiable and twice differentiable nearly everywhere. Denote by $D^2 f(x)$ the Hesse matrix (when it is defined). Then f is supermodular if and only if the N by N matrix:*

$$\mathbf{Q}^T D^2 f(x)\mathbf{Q} \tag{12}$$

has non-negative off-diagonal elements for all $x \in \{x \in X : D^2 f(x)$ exists$\}$.

5 Two Applications

5.1 A Teamwork Game with Multiple Projects

Consider a teamwork game where two economists $i = 1, 2$ coauthor two papers.[15] Player i must choose levels of effort s_i^1 and s_i^2 to put into, respectively, the first and second paper. The higher the effort, the higher the probability of success=publication and the higher therefore, the expected gain $g(s_1, s_2)$ (here $s_i = (s_i^1, s_i^2)$). The cost to placing effort s_i into working on the papers is $C_i(s_i)$, and so the (expected) payoff function is $\pi_i(s_i, s_{-i}) = g(s_i, s_{-i}) - C_i(s_i)$, $i = 1, 2$. Each player's feasible effort is bounded above, say, $S_i = \{(s_i^1, s_i^2) \in \mathbb{R}_+^2 : s_i^1 + s_i^2 \leq 1\}$. Thus the players face a resource allocation decision.[16]

We wish to investigate when this game will be supermodular or submodular in a suitably chosen vector order. As explained in the proof of theorem 2, S_i will be a sublattice of $\mathbb{R}^2$ w.r.t. a vector order $\succeq_{S_i}$ with matrix representation $\mathbf{Q} \in \mathbb{R}^{2\times 2}$ if and only if $\mathbf{Q}^{-1}(S_i)$ is a Euclidean sublattice of $\mathbb{R}^2$. By an argument similar to the one used in the proof of that same theorem, one easily sees that this will be the case if and only if the 3 by 2 matrix, $\begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \mathbf{Q}$, has the property that the product of the two elements in each row is non-positive. There are many vector orders which have this property (but the usual order is not among them). We shall consider here the order $\succeq_i$ whose matrix representation is $\mathbf{Q} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}$.[17] The product order $\succeq_S = \succeq_{S_i} \times \succeq_{S_i}$ is then placed on the joint strategy set $S = S_1 \times S_2$.

[15] The following example is a variation of the teamwork game studied in Dubey et al. (2006).

[16] Of course the 1 is just a normalization. The point is that the players can only do research for a certain number of hours each day.

[17] This is not a random choice. This is the finest vector order (the one with the smallest positive cone) under which S_i becomes a sublattice.

Applying theorem 5 one finds that player i's objective function will exhibit increasing differences if and only if $\frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^1 \partial \bar{s}_{-i}^1} \geq 0$, $\frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^1 \partial \bar{s}_{-i}^2}, \frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^2 \partial \bar{s}_{-i}^1} \leq \frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^1 \partial \bar{s}_{-i}^1}$, and $\frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^1 \partial \bar{s}_{-i}^2} + \frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^2 \partial \bar{s}_{-i}^1} \leq \frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^1 \partial \bar{s}_{-i}^1} + \frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^2 \partial \bar{s}_{-i}^2}$. It will exhibit decreasing differences if all of the inequalities are reversed.[18]

These conditions are very intuitive: When $\frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^1 \partial \bar{s}_{-i}^1} \geq 0$ (strategic complementarity), the spill-over effects between the projects $\frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^1 \partial \bar{s}_{-i}^2}$ and $\frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^2 \partial \bar{s}_{-i}^1}$ must not be so strong that they dominate the direct complementarity effects. Applying theorem 8, one finds that the objective functions will be supermodular if and only if $\frac{\partial^2 \pi_i(s_i,\bar{s}_{-i})}{\partial^2 s_i^1} \geq \frac{\partial^2 \pi_i(s_i,\bar{s}_{-i})}{\partial s_i^1 \partial s_i^2}$. If, say, π_i is concave in s_i^1 (overall quasi-concavity in s_i again plays no role), this condition demands that there is a sufficiently strong substitution effect between the two strategies (the marginal payoff to s_i^1 must be decreasing in s_i^2).

5.2 Maximization of Value Functions

Applications of monotone methods are often motivated by lack of quasi-concavity of the relevant objective functions while smoothness is not a critical extra assumption. There is, however, a whole class of applications where this is not so. In any model where an indirect utility function is maximized, differentiability of the (direct) utility function is of no importance for whether the indirect utility function is differentiable. This depends entirely on whether the underlying maximizer is unique. To be specific, the class of problems in mind usually fits into the following very general welfare problem where $V^1, \ldots, V^I$ are indirect utility functions and W a social welfare function:

$$\text{(13)} \qquad \begin{array}{l} \max W(V^1(t;x), \ldots, V^1(t;x)) \\ \text{s.t.} \left\{ \; x \in \Gamma(t) \right. \end{array}$$

Here $t \in \mathbb{R}_+^N$ is taken as given and $x \in \mathbb{R}^M$ is the maximizer. The general interpretation is that a social planner must choose a policy variable in an environment where the individual agents' maximizing behavior depends on the policy pursued. In the *wealth distribution problem* (see Mas-Colell et al.

[18]Whether $\frac{\partial^2 g(s_i,\bar{s}_{-i})}{\partial s_i^1 \partial \bar{s}_{-i}^1}$ will be negative (strategic substitutes) or positive (strategic complements) depends on the specification of g. Substitutes arises in the case where only one author has to succeed for the paper to succeed; while complements arise when both must make a good contribution for the paper to get published. See Dubey et al. (2006) and also Jensen (2005).

(1995), chapter 4), the objective is to distribute aggregate income among the consumers such that the social welfare criterion W is maximized. Here $x = (w^1, \ldots, w^I)$, $t = (p, w)$ (prices and aggregate income), and $\Gamma(t) = \{x = (w^1, \ldots, w^I) \geq 0 : \sum_i w^i \leq w\}$. In *optimal taxation models* the objective is to maximize the indirect utility of (typically) a representative consumer subject to the requirement that some prespecified amount of revenue is raised (see *e.g.* Auerbach and Hines (2001)). Here $I = 1$, W the identity, t is a vector of producer prices, and x a vector of commodity tax rates.

One of the simplest applications of this framework is in *monopoly union* models of labor economics. There are many variations, but the one we consider here seems to include most of them (*e.g.* the one used in Soskice and Iversen (2000)). The union seeks to maximize a utilitarian social welfare function whose entries are the indirect utility functions of its members. The indirect utility of worker i is $V^i(p, w)$.[19]

$$\text{(14)} \qquad \begin{array}{l} \max \sum_i V^i(p,w) \\ \text{s.t.} \quad w \in \mathbb{R}_+ \end{array}$$

To relate the properties of utility functions to their value functions (the indirect utility functions), one uses an envelope theorem. As mentioned above, V^i will not be differentiable at a point unless the worker's optimal choice is unique. But because of the results in section 3, differentiability is not needed in the first place. The reader who is familiar with non-smooth analysis will know that there basically is an envelope theorem for each of the situations studied in that section. One instance is due to Clarke (Clarke (1975), theorem 2.1.) and produces a value function which is locally Lipschitz. Theorem 6 can then be applied.[20] Since a locally Lipschitz continuous function is differentiable almost everywhere the convex-valued multifunction (here we take the usual order $\mathbf{P} = 1$, and t in the theorem to be the wage w) $\partial_w \sum_i V^i(p, w)$ will be single-valued almost everywhere and equal to the shadow price of a marginal change in the wage. By theorem 6, the union's objective function

[19] What follows is only interesting when V^i is not homogenous of degree 0 in (p, w) which is the case, for example, when workers receive nominal income transfers (see also Soskice and Iversen (2000) for a model with monetary non-neutrality). The basic idea is that the union knows the firms' aggregate labor demand curve and work is distributed (perhaps randomly) across the members.

[20] In its simplest special case, Clarke's result requires that the problem can be recast as an unconstrained optimization problem with an objective function which is upper semicontinuous and continuously differentiable in the parameters (in fact, in this simple formulation the result appears to be originally due to Danskin (1967), though I am unsure whether Danskin also proved that the value function will be locally Lipschitz).

will exhibit increasing differences in w and some parameter, say p, if and only if $\partial_w \sum_i V^i(\cdot, w) : \mathbb{R}_+ \to 2^{\mathbb{R}}$ has an order-preserving selection for every fixed w. Since supermodularity is automatically satisfied, it follows immediately that if $V = \sum_i V^i$ exhibits increasing differences, then an increase in the general price level p will make the union increase the wage rate w in response. The intuition here is relatively simple: When $\partial_w \sum_i V^i(p, w)$ is single-valued the requirement is simply that the marginal utility of an increase in the wage must be increasing in the general price level p. When $\partial_w \sum_i V^i(p, w)$ is not single-valued, the concept of a marginal utility of the wage is not well-defined. But under the conditions above, the value function *does* have well-defined right and left partial (directional) derivatives and $\partial_w \sum_i V^i(\cdot, w)$ is then the entire interval between these. If both of the directional derivatives are increasing in p, some selection from $\partial_w \sum_i V^i(\cdot, w)$ is therefore certainly increasing. The right partial derivative is the marginal utility from an *increase* in the wage, the left derivative equals *minus* the marginal utility from a *decrease* in the wage. Bearing this in mind, the intuition from the single-valued case carries over directly: There may be many combinations of p and w where an increase and a decrease in the wage set by the union yield different marginal benefits, but as long as both of these are always increasing in p it will nevertheless be the case that the union raises w with p.

References

Auerbach, A. J. and J.R. Hines Jr. (2001): "Perfect Taxation with Imperfect Competition", *NBER Working Paper* W8138.

Bartle, R.G. (1996): "Return to the Riemann Integral", *American Mathematical Monthly* 103, 625-632.

Birkhoff, G., Lattice Theory, 3rd edition, American Mathematical Society, Providence, Rhode Island, 1967.

Clarke, F.H. (1975): "Generalized Gradients and Applications", *Transactions of the American Mathematical Society* 205, 247-262.

Clarke, F.H., Optimization and Nonsmooth Analysis, 1st Edition, Wiley and Sons, New York, 1983.

Danskin, J.M., The Theory of Max-Min and its Application to Weapons Allocation Problems, Springer-Verlag, New-York, 1967.

Dubey, P., O. Haimanko, and A. Zapechelnyuk (2006): "Strategic Complements and Substitutes, and Potential games", *Games and Economic Behavior* 54, 77-94.

Gordon, R.A., The integrals of Lebesgue, Denjoy, Perron, and Henstock, American Mathematical Society, Providence, 1994.

Jensen, M.K.: "Existence, Comparative Statics, and Stability in Games with Strategic Substitutes", University of Birmingham, 2005.

Mas-Collel, A., M.D. Whinston, and J.R. Green, Microeconomic Theory, Oxford University Press, New York, 1995.

Milgrom, P. and C. Shannon (1994): "Monotone Comparative Statics", *Econometrica* 62, 157-180.

Milgrom, P. and J. Roberts (1990): "Rationalizability and Learning in Games with Strategic Complementarities", *Econometrica* 58, 1255-1278.

Peressini, A.L., Ordered Topological Vector Spaces, 1st edition, Harper and Row, New York, 1967.

Roberts, A.W. and D.E. Varberg (1974): "Another Proof that Convex Functions are Locally Lipschitz", *The American Mathematical Monthly* 81, 1014-1016.

Schaefer, H.H., with M.P. Wolff, Topological Vector Spaces, 2nd edition, Springer-Verlag, New York, 1999.

Soskice, D. and T. Iversen (2000): "The Nonneutrality of Monetary Policy with Large Price or Wage Setters", *Quarterly Journal of Economics* 115, 265-284.

Talvila, E. (2001): "Necessary and Sufficient Conditions for Differentiating Under the Integral Sign" *American Mathematical Monthly* 108, 544-548.

Topkis, D.M. (1978): "Minimizing a Supermodular Function on a Lattice", *Operations Research*, 26, 305-321.

Topkis, D.M., Supermodularity and Complementarity, Princeton University Press, Princeton, New Jersey, 1998.

Veinott, A.F. (1989): “Representation of General and Polyhedral Subsemilattices and Sublattices of Product Spaces” *Linear Algebra and Its Applications* 114/115, 681-704.