ENVELOPE THEOREMS FOR ARBITRARY CHOICE SETS

By Paul Milgrom and Ilya Segal[1]

The standard envelope theorems apply to choice sets with convex and topological structure, providing sufficient conditions for the value function to be differentiable in a parameter and characterizing its derivative. This paper studies optimization with arbitrary choice sets and shows that the traditional envelope formula holds at any differentiability point of the value function. We also provide conditions for the value function to be, variously, absolutely continuous, left- and right-differentiable, or fully differentiable. These results are applied to mechanism design, convex programming, continuous optimization problems, saddle-point problems, problems with parameterized constraints, and optimal stopping problems.

KEYWORDS: Envelope theorem, differentiable value function, sensitivity analysis, math programming, mechanism design.

1. INTRODUCTION

TRADITIONAL "ENVELOPE THEOREMS" do two things: describe sufficient conditions for the value of a parameterized optimization problem to be differentiable in the parameter and provide a formula for the derivative. Economists initially used envelope theorems for concave optimization problems in demand theory. The theorems were used to analyze the effects of changing prices, incomes, and technology on the welfare and profits of consumers and firms. With households and firms choosing quantities of consumer goods and inputs, the choice sets had both the convex and topological structure required by the early envelope theorems.

In recent years, results that may be regarded as extensions of envelope theorems have frequently been used to study incentive constraints in contract theory and game theory,[2] to examine nonconvex production problems,[3] and to develop the theory of "monotone" or "robust" comparative statics.[4] The choice sets and objective functions in these applications generally lack the topological and convexity properties required by the traditional envelope theorems. At the same time, the analysis of these applications does not always require full differentiability of the value function everywhere. For example, contract theory considers

[2] There are many such examples, beginning with Mirrlees (1971).

[3] For example, see Milgrom and Roberts (1988).

[4] See Milgrom and Shannon (1994) and Athey, Milgrom, and Roberts (2000).

incentive mechanisms with arbitrary message spaces and arbitrary outcome functions. While an agent's value function in such a mechanism need not be a differentiable function of his type, it can nevertheless be represented as an integral of the partial derivative of the agent's payoff function with respect to his type. This representation constitutes an important step in the analysis of optimal contracts. While some progress has been made in extending traditional envelope theorems to be useful in such modern applications, none has been general enough to encompass them all.[5]

The core contributions of this paper are envelope theorems for maximization problems with arbitrary choice sets, in which such properties of the objective function as differentiability, concavity, or continuity in the choice variable cannot be utilized. First we show that the traditional envelope *formula* holds at any differentiability point of the value function. Then we provide a sufficient condition for the value function to be absolutely continuous. This condition ensures that the value function is differentiable almost everywhere and can be represented as an integral of its derivative. We also provide a sufficient condition for the value function to have right- and left-hand directional derivatives *everywhere* and characterize those derivatives. When the two directional derivatives are equal, the function is differentiable.

Associated with the new envelope theorems is a new intuition, distinct from the one offered in leading graduate economics textbooks.[6] In our approach, the choice set has no structure and is used merely as a set of indices to identify elements of a family of functions on the set $[0, 1]$ of possible parameter values. Figure 1 illustrates this approach for the case of a finite choice set $X = \{x_1, x_2, x_3\}$.

The value function $V(t) = \max_{x \in X} f(x, t)$ is the "upper envelope" of the functions $f(x, t)$. The figure illustrates several of its general properties when the choice set is finite and the objective function f is continuously differentiable in the parameter t. First, the value function is differentiable almost everywhere and has directional derivatives everywhere. Its right-hand derivative at parameter value t is everywhere equal to the largest of the partial derivatives $f_t(x, t)$ on the set of optimal choices at t, while the left-hand derivative is everywhere equal to the smallest of the partial derivatives. Consequently, V is differentiable at t if and only if the derivative is constant on the set of optimal choices. This occurs wherever the maximum is unique but, as the Figure shows, it can also happen at other points.

Our general envelope theorems, stated and proved in Section 2, expand upon this example. In Section 3, we explore several applications, utilizing the additional structure available in these applications. The first application is to problems of mechanism design. The second is to maximization problems that are concave in

---

[5] The mathematical literature on "sensitivity analysis" has formulated several generalized Envelope Theorems—see Bonnans and Shapiro (2000, Section 4.3) for a recent survey. These results by and large rely on topological assumptions on the choice set and continuity of the objective function in the choice variable. We compare these results to ours in Section 3.

[6] See, for example, Mas-Colell, Whinston, and Green (1995) and Simon and Blume (1994).
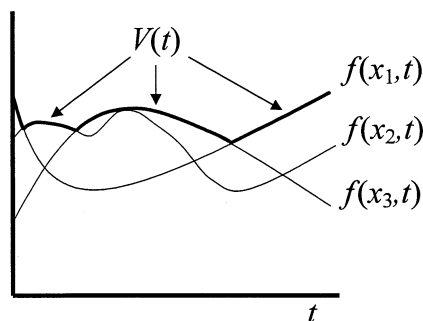
FIGURE 1

both the choice variable and the parameter, generalizing the envelope theorem formulated by Benveniste and Scheinkman (1979). The third is to the case where the choice set is compact and both the objective function and its derivative are continuous with respect to the parameter. The fourth is to saddle-point problems on compact sets. The fifth applies the saddle-point envelope theorem to constrained maximization problems with a parameterized constraint, using the characterization of solutions as saddle points of the Lagrangian. The sixth application derives the smooth pasting condition in optimal stopping problems. Section 4 concludes.

2. GENERAL RESULTS

Let X denote the choice set and let the relevant parameter be $t \in [0, 1]$.[7] Letting $f : X \times [0, 1] \to \mathbb{R}$ denote the parameterized objective function, the value function V and the optimal choice correspondence (set-valued function) X^* are given by:[8]

$$(1) \qquad V(t) = \sup_{x \in X} f(x, t),$$

$$(2) \qquad X^*(t) = \{x \in X : f(x, t) = V(t)\}.$$

Our first result relates the derivatives of the value function to the partial derivative $f_t(x, t)$ of the objective function with respect to the parameter.

THEOREM 1: *Take $t \in [0, 1]$ and $x^* \in X^*(t)$, and suppose that $f_t(x^*, t)$ exists. If $t > 0$ and V is left-hand differentiable at t, then $V'(t-) \le f_t(x^*, t)$. If $t < 1$ and*

---

[7] More generally, when the parameter lies in a normed vector space, this treatment applies to directional derivatives and path derivatives in that space.

[8] In this section we will assume nonemptiness of $X^*(t)$ at various points t as needed. In Section 3 we demonstrate how this nonemptiness is ensured by additional structure available in various economic applications.

V is right-hand differentiable at t, then $V'(t+) \geq f_t(x^*, t)$. If $t \in (0, 1)$ and V is differentiable at t, then $V'(t) = f_t(x^*, t)$.

PROOF: Using (1) and (2), we see that for any $t' \in [0, 1]$,

$$f(x^*, t') - f(x^*, t) \leq V(t') - V(t).$$

Taking $t' \in (t, 1)$, dividing both sides by $t' - t > 0$, and taking their limits as $t' \to t+$ yields $f_t(x^*, t) \leq V'(t+)$ if the latter derivative exists. Taking instead $t' \in (0, t)$, dividing both sides by $t - t' > 0$, and taking their limits as $t' \to t-$ yields $f_t(x^*, t) \geq V'(t-)$ if the latter derivative exists. When V is differentiable at $t \in (0, 1)$, we must have $V'(t) = V'(t-) = V'(t+) = f_t(x^*, t)$.   Q.E.D.

Theorem 1 is only useful when the value function V is sufficiently well-behaved—for example, differentiable, directionally differentiable, or absolutely continuous. In the remainder of this section, we identify sufficient conditions for the value function to have these properties. These conditions do not exploit any structure of the choice set X, but treat it as merely a set of indices identifying elements of the family of functions $\{f(x, \cdot)\}_{x \in X}$ on the set $[0, 1]$ of possible parameter values. The conditions for the value function to be well behaved will involve certain properties that the functions $\{f(x, \cdot)\}_{x \in X}$ must satisfy uniformly.[9]

In particular, the following result offers a sufficient condition for the value function to be absolutely continuous. In this case, the value function is differentiable almost everywhere and can be represented as an integral of its derivative:

THEOREM 2: *Suppose that $f(x, \cdot)$ is absolutely continuous for all $x \in X$. Suppose also that there exists an integrable function $b : [0, 1] \to \mathbb{R}_+$ such that $|f_t(x, t)| \leq b(t)$ for all $x \in X$ and almost all $t \in [0, 1]$. Then V is absolutely continuous. Suppose, in addition, that $f(x, \cdot)$ is differentiable for all $x \in X$, and that $X^*(t) \neq \varnothing$ almost everywhere on $[0, 1]$. Then for any selection $x^*(t) \in X^*(t)$,*

$$(3) \qquad V(t) = V(0) + \int_0^t f_t(x^*(s), s)\, ds.$$

PROOF: Using (1), observe that for any $t', t'' \in [0, 1]$ with $t' < t''$,

$$|V(t'') - V(t')| \leq \sup_{x \in X} |f(x, t'') - f(x, t')|$$

$$= \sup_{x \in X} \left| \int_{t'}^{t''} f_t(x, t)\, dt \right| \leq \int_{t'}^{t''} \sup_{x \in X} |f_t(x, t)|\, dt \leq \int_{t'}^{t''} b(t)\, dt.$$

This implies that V is absolutely continuous. Therefore, V is differentiable almost everywhere, and $V(t) = V(0) + \int_0^t V'(s)\, ds$. If $f(x, t)$ is differentiable in t, then $V'(s)$ is given by Theorem 1 wherever it exists, and we obtain (3).   Q.E.D.

---

[9] Mathematical concepts and results used in this paper can be found in Aliprantis and Border (1994), Royden (1988), Rockafellar (1970), and Apostol (1969).

The integral representation (3) plays a key role in mechanism design (see Section 3). The role of the integrable bound in Theorem 2 is illustrated with the following example:

EXAMPLE 1: Let $X = (0, +\infty)$ and $f(x, t) = g(t/x)$, where $g(z)$ is a differentiable function that achieves a unique maximum at $z = 1$, and $\alpha \equiv \sup_{z \in (0, +\infty)} |zg'(z)| < +\infty$. (For example, $g(z) = ze^{-z}$ satisfies these conditions.) Observe that $\sup_{x \in X} |f_t(x, t)| = \sup_{x \in X} |\frac{1}{t}(\frac{t}{x}g'(t/x))| = \alpha/t$, which is not integrable on $[0, 1]$. By inspection, for all $t > 0$, $X^*(t) = \{t\}$, and $V(t) = g(1) > V(0) = g(0)$. Note that for any $t \in (0, 1)$, $f_t(x^*(t), t) = g'(1)/t = 0 = V'(t)$, illustrating Theorem 1. However, the conclusion of Theorem 2 does not hold, for V is discontinuous at $t = 0$. It follows that the integrable bound assumed in Theorem 2 is not dispensable.

The assumptions of Theorem 2 do not ensure that the value function is differentiable everywhere, as the example depicted in Figure 1 makes clear. However, in the example the value function is right- and left-differentiable everywhere. This observation can be extended from finite to arbitrary choice sets, provided that the family of objective functions satisfies the following property:

DEFINITION: The family of functions $\{f(x, \cdot)\}_{x \in X}$ is *equidifferentiable* at $t \in [0, 1]$ if $(f(x, t') - f(x, t))/(t' - t)$ converges uniformly as $t' \to t$.

When the set X is infinite, uniform convergence on X is stronger than pointwise convergence, hence equidifferentiability is stronger than differentiability. A simple sufficient condition for the equidifferentiability of $\{f(x, \cdot)\}_{x \in X}$ is provided by the equicontinuity of $\{f_t(x, \cdot)\}_{x \in X}$ everywhere. Indeed, in this case the Mean Value Theorem allows us to write $(f(x, t') - f(x, t))/(t' - t) = f_t(x, s)$ for some s between t and t', and the equicontinuity condition implies that this expression converges uniformly to $f_t(x, t)$ as $t' \to t$.

THEOREM 3: *Suppose that the family of functions* $\{f_t(x, \cdot)\}_{x \in X}$ *is equidifferentiable at* $t_0 \in [0, 1]$, *that* $\sup_{x \in X} |f_t(x, t_0)| < +\infty$, *and that* $X^*(t) \neq \varnothing$ *for all* t. *Then* V *is left- and right-hand differentiable at* t_0. *For any selection* $x^*(t) \in X^*(t)$, *the directional derivatives are*

(4)
$$V'(t_0+) = \lim_{t \to t_0+} f_t(x^*(t), t_0) \quad \text{for} \quad t_0 < 1,$$
$$V'(t_0-) = \lim_{t \to t_0-} f_t(x^*(t), t_0) \quad \text{for} \quad t_0 > 0.$$

V *is differentiable at* $t_0 \in (0, 1)$ *if and only if* $f_t(x^*(t), t_0)$ *is continuous in* t *at* $t = t_0$.

PROOF: Using (1) and the assumption that $\sup_{x \in X} |f_t(x, t_0)| < +\infty$, equidifferentiability implies

$$|V(t) - V(t_0)| \leq \sup_{x \in X} |f(x, t) - f(x, t_0)|$$

$$\leq \sup_{x \in X} |f_t(x, t_0)| \cdot |t - t_0| + o(t - t_0) \to 0 \quad \text{as} \quad t \to t_0.$$

Therefore, $\{f(x, \cdot)\}_{x \in X}$ is equicontinuous at t_0 and the value function V is continuous at t_0.

Take $t_0 < t' < t''$. Using (1), we can write

$$\frac{f(x^*(t'), t'') - f(x^*(t'), t')}{t'' - t'} \leq \frac{V(t'') - V(t')}{t'' - t'}$$

$$\leq \frac{f(x^*(t''), t'') - f(x^*(t''), t')}{t'' - t'}.$$

Taking the limit superior as $t' \to t_0+$, and using the equicontinuity of $\{f(x, \cdot)\}_{x \in X}$ and continuity of V at t_0, this yields

$$\overline{\lim_{t' \to t_0+}} \frac{f(x^*(t'), t'') - f(x^*(t'), t_0)}{t'' - t_0} \leq \frac{V(t'') - V(t_0)}{t'' - t_0}$$

$$\leq \frac{f(x^*(t''), t'') - f(x^*(t''), t_0)}{t'' - t_0}.$$

Using equidifferentiability, this implies

$$\overline{\lim_{t' \to t_0+}} f_t(x^*(t'), t_0) + \frac{o(t'' - t_0)}{t'' - t_0} \leq \frac{V(t'') - V(t_0)}{t'' - t_0}$$

$$\leq f_t(x^*(t''), t_0) + \frac{o(t'' - t_0)}{t'' - t_0}.$$

Taking the limit inferior of the two bounds as $t'' \to t_0+$, we see that $\overline{\lim}_{t \to t_0+} f_t(x^*(t), t_0) \leq \underline{\lim}_{t \to t_0+} f_t(x^*(t), t_0)$, and therefore $\lim_{t \to t_0+} f_t(x^*(t), t_0)$ exists. Since this is the limit of both bounds in the above double inequality as $t'' \to t_0+$, we obtain the first line in (4). The second line is established similarly.

V is differentiable at $t_0 \in (0, 1)$ if and only if $V'(t_0+) = V'(t_0-) = f_t(x^*(t_0), t_0)$, where the second equality is by Theorem 1. By (4), this double equality means that $f_t(x^*(t), t_0)$ is continuous in t at $t = t_0$.   Q.E.D.

The following example demonstrates that simple differentiability of $f(x, t)$ in t for all x does not suffice for the conclusion of Theorem 3:

EXAMPLE 2: Let $X = \{1, 2, \ldots\}$ and

$$f(x, t) = \begin{cases} t \sin \log t & \text{if } t > t(x), \\ -t & \text{if } t \leq t(x), \end{cases} \quad \text{where } t(x) = \exp\{-\pi/2 - 2\pi x\}.$$

It is easy to see that $V(t) = t \sin \log t$. Observe that $f(x, t)$ is differentiable in t for all x, with $|f_t(x, t)| \leq 2$ for all (x, t). (In particular, the assumptions of Theorem 2 are satisfied.) However, $\{f(x, \cdot)\}_{x \in X}$ is not equidifferentiable at $t_0 = 0$:

$$\sup_{x \in X} \left| \frac{f(x, t) - f(x, 0)}{t - 0} - f_t(x, 0) \right| = \sin \log t + 1 \not\rightarrow 0 \quad \text{as} \quad t \rightarrow 0.$$

Observe that V does not have a right-hand derivative at $t_0 = 0$, since

$$\overline{\lim_{t \rightarrow 0+}} [V(t)/t] = 1 \neq \underline{\lim_{t \rightarrow 0+}} [V(t)/t] = -1.$$

Therefore, we cannot dispense with the assumption of equidifferentiability in Theorem 3.

In conclusion of this section, observe that Theorems 2 and 3 can be applied when their assumptions hold only on the reduced choice set $X^*([0, 1]) = \bigcup_{s \in [0,1]} X^*(s)$. Indeed, replacement of X with $X^*([0, 1])$ will not affect the value function V or the optimal choice correspondence X^*.

3. APPLICATIONS

In this section we demonstrate how the general results outlined above can be applied to several important economic settings. The additional structure available in these settings can be utilized to verify the assumptions of Theorems 1–3, as well as to strengthen their conclusions.

3.1. *Mechanism Design*

Consider an agent whose utility function $f(x, t)$ over outcomes $x \in Y$ depends on his type $t \in [0, 1]$. The agent is offered a mechanism, described by a message set M and an outcome function $h: M \rightarrow Y$. The mechanism induces the *menu* $X = \{h(m) : m \in M\} \subset Y$, i.e., the set of outcomes that are accessible to the agent. The agent's equilibrium utility $V(t)$ in the mechanism is then given by (1), and the set $X^*(t)$ of the mechanism's equilibrium outcomes is given by (2). Any selection $x^*(t) \in X^*(t)$ is a *choice rule* implemented by the mechanism.

For this setting, Theorem 2 immediately implies the following corollary.

COROLLARY 1: *Suppose that the agent's utility function $f(x, t)$ is differentiable and absolutely continuous in t for all $x \in Y$, and that $\sup_{x \in Y} |f_t(x, t)|$ is integrable on $[0, 1]$.*[10] *Then the agent's equilibrium utility V in any mechanism implementing a given choice rule x^* must satisfy the integral condition* (3).

---

[10] The last assumption can be relaxed in some commonly studied mechanism design settings. For example, suppose that an outcome can be described as $x = (z, w)$, where $w \in \mathbb{R}$ is the monetary transfer to the agent and $z \in Z \subset \mathbb{R}$ is a nonmonetary decision. Suppose furthermore that the agent's utility function takes the quasilinear form $f(z, w, t) = g(z, t) + w$, and that g has strictly increasing differences in (z, t) (equivalently, f has the Spence-Mirrlees single-crossing property). Then

Deducing condition (3) is a key step in the analysis of mechanism design problems with continuous type spaces. Mirrlees (1971), Laffont and Maskin (1980), Fudenberg and Tirole (1991), and Williams (1999) derived and exploited this condition by restricting attention to (piecewise) continuously differentiable choice rules. This is not fully satisfactory, because a mechanism designer may find it optimal to implement a choice rule that is not piecewise continuously differentiable. For example, in the trade setting with linear utility (see, e.g., Myerson (1991, Section 6.5)), both the profit-maximizing and total surplus-maximizing choice rules are usually discontinuous.[11] At the same time, the integral condition (3) still holds in this setting and implies such important results as the Revenue Equivalence Theorem for auctions and the Myerson-Satterthwaite inefficiency theorem.

It should be noted that Corollary 1 can be applied to multidimensional type spaces as well. For example, suppose that the agent's type space is $\Theta \subset \mathbb{R}^k$ and his utility function is $g : X \times \Theta \to \mathbb{R}$. Suppose that Θ is smoothly connected, that is, any two points $a, b \in \Theta$ are connected by a path described by a continuously differentiable function $\tau : [0, 1] \to \Theta$ such that $\tau(0) = a$ and $\tau(1) = b$. If g is differentiable in $\theta \in \Theta$ and the gradient $g_\theta(x, \theta)$ is bounded on $X \times \Theta$, then the function $f(x, t) = g(x, \tau(t))$ satisfies the assumptions of Corollary 1. The Corollary then implies that if $V : \Theta \to \mathbb{R}$ is the agent's value function in a mechanism implementing the choice rule $x^* : \Theta \to X$, then $V(b) - V(a)$ equals the path integral of the gradient $g_\theta(x^*(\theta), \theta)$ along the path connecting a and b. Since this result holds for any smooth path in Θ, V is a potential function for the vector field $g_\theta(x^*(\theta), \theta)$, and is therefore determined by this field up to a constant (see, e.g., Apostol (1969)).[12]

In addition to the integral representation (3), it is sometimes of interest to know that the agent's equilibrium utility V is differentiable. For example, suppose that, as in Segal and Whinston (2002), the agent chooses his type t, interpreted as investment, before participating in the mechanism.[13] Suppose the agent maxi-

---

the Monotone Selection Theorem (Milgrom and Shannon (1994)) implies that for any selection $x^*(t) = (z^*(t), w^*(t)) \in X^*(t)$, $z^*(t)$ is nondecreasing in t. Furthermore, under strictly increasing differences, $g_t(z, t)$ is nondecreasing in z, and therefore $f_t(x^*(s), t) = g_t(z^*(s), t) \in [g_t(z^*(0), t), g_t(z^*(1), t)]$ for all s. Therefore, $f_t(x, t)$ is uniformly bounded on $(x, t) \in X^*([0, 1]) \times [0, 1]$. This allows us to apply Theorem 2 on the reduced choice set $X^*([0, 1])$ and obtain the integral representation (3).

[11] Myerson (1981) proves condition (3) utilizing the special structure of the linear setting. However, his proof does not readily generalize to other settings. While monotonicity of implementable decision rules is typically used to show that the value function is differentiable almost everywhere, this by itself does nto imply that it equals the integral of the derivative. For example, it does not rule out the possibility that the value function is discontinuous. Even establishing continuity of the value function would not suffice: a counterexample is provided by the Cantor ternary function (see, e.g., Royden (1988)). Thus, establishing *absolute* continuity of the value function is an indispensable step for deriving (3).

[12] Krishna and Maenner (2001) derive this result independently, but under unnecessary restrictions on the agent's payoffs or the mechanism itself (their Hypotheses I and II).

[13] Any cost of this investment is included in f.

mizes his equilibrium utility in the mechanism by choosing investment $t_0 \in (0, 1)$. Then Theorem 3 implies the following result:

COROLLARY 2: *Suppose that a mechanism implements a choice rule x^* and gives rise to the agent's equilibrium utility V, and that $t_0 \in \arg\max_{t \in (0,1)} V(t)$. If $\{f(x, \cdot)\}_{x \in Y}$ is equidifferentiable and $\sup_{x \in Y} |f_t(x, t_0)| < +\infty$,[14] then V is differentiable at t_0, and $V'(t_0) = f_t(x^*(t_0), t_0) = 0$.*

PROOF: Since the menu X induced by the mechanism is a subset of Y, the assumptions of Theorem 3 hold. Therefore, V is directionally differentiable at t_0. Since $t_0 \in \arg\max_{t \in (0,1)} V(t)$, the directional derivatives must satisfy $V'(t_0-) \geq 0 \geq V'(t_0+)$. On the other hand, by Theorem 1, $V'(t_0-) \leq f_t(x^*(t_0), t_0) \leq V'(t_0+)$.
                                                                        Q.E.D.

Corollary 2 implies that any mechanism sustaining an interior investment t_0 can be replaced with a fixed outcome $x^*(t_0)$ sustaining the same investment, provided that the function $f(x^*(t_0), t)$ is concave in t. This parallels a key finding of Segal and Whinston (2002).[15]

The results of this subsection apply to multi-agent mechanism design settings as well. Consider such a setting from the viewpoint of one agent, where the implemented choice rule and the agent's equilibrium utility in general depend on other agents' messages. In dominant-strategy implementation, our analysis applies for any given profile of other agents' messages. In Bayesian-Nash implementation, the outcome set Y can be defined as the set of probability distributions over a set Z of primitive outcomes. In a Bayesian-Nash equilibrium of a mechanism, an agent chooses from a set $X \subset Y$ of probability distributions that are accessible to him given equilibrium behavior by other agents. If the agent's underlying Bernoulli utility function over primitive outcomes from Z satisfies the integrable bound and equidifferentiability conditions in Corollaries 1 and 2, then his von Neumann-Morgenstern expected utility over distributions from Y also satisfies these conditions, and our analysis applies.

3.2. *Convex Programming with Convex Parameterization*

We can use Theorem 1 to generalize the well-known envelope theorem of Benveniste and Scheinkman (1979), by incorporating a requirement that the objective be concave in both the choice variable *and the parameter*.

COROLLARY 3: *Suppose that X is a convex set in a linear space and $f : X \times [0, 1] \rightarrow \mathbb{R}$ is a concave function. Also suppose that $t_0 \in (0, 1)$, and that there*

---

[14] These conditions are in turn ensured by the compactness of Y and the continuity of $f_t(x, t)$ in (x, t), as shown in the proof of Corollary 4 below.

[15] Segal and Whinston (2002) consider *two* agents, who choose investments and then participate in a mechanism. The special case of their model in which only one agent invests satisfies the assumptions of Corollary 2.

is some $x^* \in X^*(t_0)$ such that $f_t(x^*, t_0)$ exists. Then V is differentiable at t_0 and $V'(t_0) = f_t(x^*, t_0)$.

PROOF: Take $t', t'', \lambda \in [0, 1]$. By the convexity of X and the concavity of f, for any $x', x'' \in X$ we can write

$$f(\lambda x' + (1 - \lambda)x'', \lambda t' + (1 - \lambda)t'') \geq \lambda f(x', t') + (1 - \lambda)f(x'', t'').$$

Taking the supremum of both sides over $x', x'' \in X$, and using the convexity of X, we obtain $V(\lambda t' + (1 - \lambda)t'') \geq \lambda V(t') + (1 - \lambda)V(t'')$, and therefore V is concave. This implies that V is directionally differentiable at each $t \in (0, 1)$ and $V'(t-) \geq V'(t+)$ (see, e.g., Rockafellar (1970)). On the other hand, by Theorem 1, $V'(t_0-) \leq f_t(x^*, t_0) \leq V'(t_0+)$.   Q.E.D.

The Benveniste and Scheinkman theorem established the differentiability of the value function in a class of infinite-horizon consumption problems with a parameterized initial endowment. In their setting, X is the set of technologically feasible consumption paths, and the objective function is the consumer's intertemporal utility, e.g., $f(x, t) = u(x_0 + t) + \sum_{s=1}^{\infty} \delta^s u(x_s)$.[16]

3.3. *Continuous Objective Functions on Compact Choice Sets*

If X is a nonempty compact space and $f(x, t)$ is upper semicontinuous in x, then $X^*(t)$ is nonempty. If, in addition, $f_t(x, t)$ is continuous in (x, t), then all the assumptions of Theorems 2 and 3 are satisfied. Furthermore, in this case we can simplify the expressions for the directional derivatives of V and the characterization of the differentiability points of V. These results can be summarized as follows:

COROLLARY 4: *Suppose that X is a nonempty compact space, $f(x, t)$ is upper semicontinuous in x, and $f_t(x, t)$ is continuous in (x, t). Then*
  (i) *V is absolutely continuous and the integral representation* (3) *holds.*
  (ii) *$V'(t+) = \max_{x \in X^*(t)} f_t(x, t)$ for any $t \in [0, 1)$ and $V'(t-) = \min_{x \in X^*(t)} f_t(x, t)$ for any $t \in (0, 1]$.*
  (iii) *V is differentiable at a given $t \in (0, 1)$ if and only if $\{f_t(x, t) | x \in X^*(t)\}$ is a singleton, and in that case $V'(t) = f_t(x, t)$ for all $x \in X^*(t)$.*

PROOF: The continuous function $|f_t(x, t)|$ is bounded on $X \times [0, 1]$, so the "integrable bound" condition of Theorem 2 is satisfied. Furthermore, since $f(x, t)$ is upper semicontinuous in x, $X^*(t)$ is a nonempty compact set for all t. Also, the absolute continuity of $f(x, t)$ in t is implied by its continuous differentiability in t. Therefore, all assumptions of Theorem 2 are satisfied, which establishes part (i).

---

[16] If, in addition to the technological constraints embodied in X, there is a constraint on feasible consumption $x_0 + t$ in the first period (e.g., $x_0 + t \geq 0$), then the present analysis applies on neighborhoods in the parameter set where the consumption constraint is nonbinding.

Next, the continuity of f_t and the compactness of X imply that the family of functions $\{f_t(x, \cdot)\}_{x \in X}$ is equicontinuous. As noted in Section 2, this implies that $\{f(x, \cdot)\}_{x \in X}$ is equidifferentiable at any t. Since f_t is also bounded on $X \times [0, 1]$, all assumptions of Theorem 3 are satisfied. Therefore, V has directional derivatives, which are given by (4).

Take $t_0 \in [0, 1)$. Berge's Maximum Theorem (see, e.g., Aliprantis and Border (1994)) and the continuity of f_t imply that for any selection $x^*(t) \in X^*(t)$, $\overline{\lim}_{t \to t_0+} f_t(x^*(t), t_0) \leq \max_{x \in X^*(t_0)} f_t(x, t_0)$. Combining with (4), we see that $V'(t_0+) \leq \max_{x \in X^*(t_0)} f_t(x, t_0)$. Since Theorem 1 implies the reverse inequality, this establishes the first part of (ii). The second part is established similarly. Part (iii) follows immediately.   *Q.E.D.*

A version of this result was first obtained by Danskin (1967). In the economic literature, the result was rediscovered by Kim (1993) and Sah and Zhao (1998).

Corollary 4 makes it clear that, contrary to the conventional wisdom in the economic literature, good behavior of the value function does not rely on good behavior of maximizers. For example, consider a bounded linear programming problem in a Euclidean space. At a parameter value at which there are multiple maximizers, any selection of maximizers is typically discontinuous in the parameter. Nevertheless, Corollary 4(i) establishes that the value function is absolutely continuous.

As another example, suppose that X is a convex compact set in a Euclidean space described by a collection of inequality constraints, and that the objective function is strictly concave in x. Then the optimal choice is unique, and therefore by Corollary 4(iii) the value function is differentiable everywhere, even at parameter values where the maximizer is not differentiable (e.g., where the set of binding constraints changes). While the traditional envelope theorem derived from first-order conditions (see, e.g., Simon and Blume (1994)) cannot be used at such points, Corollary 4(iii) establishes that the envelope formula must still hold.

To understand the role of compactness in parts (ii) and (iii) of Corollary 4, consider the following example:

EXAMPLE 3: Let $X = \{0\} \cup (\frac{1}{2}, 1]$, and

$$f(x, t) = \begin{cases} -(t - x)^2 & \text{for } x \in (\frac{1}{2}, 1], \\ \frac{1}{2} - t & \text{for } x = 0. \end{cases}$$

With the Euclidean topology on X, the example satisfies all the assumptions of Corollary 4 except for compactness of X.[17] Note that $X^*(t)$ is a singleton for all t: in particular, for $t \leq \frac{1}{2}$, $X^*(t) = \{0\}$ and $V(t) = \frac{1}{2} - t$, while for $t > \frac{1}{2}$, $X^*(t) = \{t\}$ and $V(t) = 0$. Nevertheless, V is not differentiable at $t = \frac{1}{2}$, and its right-hand derivative at this point does not satisfy the formula in Corollary 4(ii).

----

[17] By changing the topology on X, the same example can be construed as one in which X is compact but the continuity assumptions of Theorem 3 are violated.

3.4. *Saddle-Point Problems*

In this subsection we extend our previous analysis to obtain envelope theorems for saddle-point problems. The theorems will tell us, for example, how the players' Nash equilibrium payoffs in a two-player zero-sum game depend on a parameter. In mechanism design, such zero-sum games emerge when the outcome prescribed by a mechanism is renegotiated towards an ex post efficient outcome in all states of the world, as in Segal and Whinston (2002). The analysis of saddle-point problems is also useful for the study of parameterized constraints (see the next subsection.)

Let X and Y be nonempty sets, and let $f : X \times Y \times [0, 1] \to \mathbb{R}$. $(x^*, y^*) \in X \times Y$ is a *saddle point* of f at parameter value t if

$$f(x, y^*, t) \leq f(x^*, y^*, t) \leq f(x^*, y, t) \quad \text{for all} \quad x \in X, y \in Y.$$

One interpretation of a saddle point is as an equilibrium of the zero-sum game in which player 1 chooses $x \in X$, player 2 chooses $y \in Y$, and their payoffs are $f(x, y, t)$ and $-f(x, y, t)$ respectively.

It is well known (see, e.g., Rockafellar (1970)) that whenever the set of saddle points (the *saddle set*) is nonempty, it is a product set $X^*(t) \times Y^*(t) \subset X \times Y$, where

$$X^*(t) = \operatorname*{Arg\,max}_{x \in X} \inf_{y \in Y} f(x, y, t), \quad Y^*(t) = \operatorname*{Arg\,min}_{y \in Y} \sup_{x \in X} f(x, y, t).$$

In this case, for all saddle points $(x^*, y^*) \in X^*(t) \times Y^*(t)$ we must have

$$f(x^*, y^*, t) = \sup_{x \in X} \inf_{y \in Y} f(x, y, t) = \inf_{y \in Y} \sup_{x \in X} f(x, y, t) \equiv V(t),$$

where $V(t)$ is called the *saddle value* of f at t.

First we extend Theorem 2's integral representation of the value function to saddle-point problems:

THEOREM 4: *Suppose that $f(x, y, \cdot)$ is absolutely continuous for all $(x, y) \in X \times Y$, that $X^*(t) \times Y^*(t) \neq \varnothing$ for almost all $t \in [0, 1]$, and that there exists an integrable function $b : [0, 1] \to \mathbb{R}_+$ such that $|f_t(x, y, t)| \leq b(t)$ for all $(x, y) \in X \times Y$ and almost every $t \in [0, 1]$. Then V is absolutely continuous.*

Suppose, in addition, that X and Y are topological spaces satisfying the second axiom of countability,[18] that $f_t(x, y, t)$ is continuous in each of $x \in X$ and $y \in Y$, and that the family $\{f(x, y, \cdot)\}_{(x,y) \in X \times Y}$ is equidifferentiable. Then for any selection $(x^*(t), y^*(t)) \in X^*(t) \times Y^*(t)$,

$$(5) \qquad V(t) = V(0) + \int_0^t f_t(x^*(s), y^*(s), s) ds.$$

---

[18] That is, having countable bases (see, e.g., Royden (1988)). In particular, X and Y could be separable metric spaces.

PROOF: The absolute continuity of $V(t) = \sup_{x \in X} \inf_{y \in Y} f(x, y, t)$ obtains by double application of the absolute continuity result of Theorem 2. Therefore, V is differentiable almost everywhere and $V(t) = V(0) + \int_0^t V'(s)\,ds$.

Now, consider the graph of the saddle-point selection: $G \equiv \{(t, x^*(t), y^*(t)) : t \in [0, 1]\} \subset [0, 1] \times X \times Y$. Since the product topological space $[0, 1] \times X \times Y$ satisfies the second axiom of countability by our assumptions, the set of isolated points of G is at most countable. Therefore, the set S of points $t \in [0, 1]$ such that $(t, x^*(t), y^*(t))$ is not isolated in G and $V'(t)$ exists has full measure on $[0, 1]$.

Take any point $t_0 \in S$ and let $(x_0, y_0) = (x^*(t_0), y^*(t_0))$. Since (t_0, x_0, y_0) is not isolated in G, there exists a sequence $(t_k, x_k, y_k)_{k=1}^\infty \subset G$ such that $(t_k, x_k, y_k) \to (t_0, x_0, y_0)$ as $k \to \infty$ and $t_k \neq t_0$ for all k. Furthermore, the sequence can be chosen so that $t_k - t_0$ has a constant sign, and for definiteness let it be positive. By the definition of a saddle point, we can write

$$\frac{f(x_0, y_k, t_k) - f(x_0, y_k, t_0)}{t_k - t_0} \leq \frac{V(t_k) - V(t_0)}{t_k - t_0}$$

$$\leq \frac{f(x_k, y_0, t_k) - f(x_k, y_0, t_0)}{t_k - t_0}.$$

Using equidifferentiability of $\{f(x, y, \cdot)\}_{(x,y) \in X \times Y}$, this implies

$$f_t(x_0, y_k, t_0) + \frac{o(t_k - t_0)}{t_k - t_0} \leq \frac{V(t_k) - V(t_0)}{t_k - t_0}$$

$$\leq f_t(x_k, y_0, t_0) + \frac{o(t_k - t_0)}{t_k - t_0}.$$

As $k \to \infty$, by the continuity of $f_t(x, y, t)$ in x and in y, both bounds converge to $f_t(x_0, y_0, t_0)$. Therefore, we must have $V'(t_0) = f_t(x^*(t_0), y^*(t_0), t_0)$. Since this formula holds for each t_0 in the set S, which has full measure in $[0, 1]$, we obtain the result. Q.E.D.

Note that in contrast to Theorem 2 for maximization programs, Theorem 4 utilizes topologies on the choice sets X, Y and the continuity of $f_t(x, y, t)$ in these topologies. The following example demonstrates that these extra assumptions are indispensable:

EXAMPLE 4: Let $X = Y = [0, 1]$, and

$$f(x, y, t) = \begin{cases} t - x & \text{if } x \geq y, \\ y - t & \text{otherwise.} \end{cases}$$

It can be verified that for each t, the function has a unique saddle point $(x^*(t), y^*(t)) = (t, t)$, and $V(t) = 0$. Note that $V'(t) = 0$, while $f_t(x^*(t), y^*(t), t) = 1$, for all t. Thus, the integral representation (5) does not hold. Note that all the assumptions of Theorem 4 but for those involving topologies on X and Y

are satisfied. Observe that $f_t(x, y, t)$ is not continuous in x or y in the standard topology on $\mathbb{R}$. The function is trivially continuous in the discrete topology on X and Y (in which all points are isolated), but this topology does not satisfy the second countability axiom.

Under appropriate continuity assumptions, a saddle-point extension of Corollary 4 can also be obtained.[19]

THEOREM 5: *Let X and Y be compact spaces and suppose that $f : X \times Y \times [0, 1] \to \mathbb{R}$ and $f_t : X \times Y \times [0, 1] \to \mathbb{R}$ are continuous functions. Suppose also that $X^*(t) \times Y^*(t) \neq \varnothing$ for all $t \in [0, 1]$. Then V is directionally differentiable, and the directional derivatives are*

$$V'(t+) = \max_{x \in X^*(t)} \min_{y \in Y^*(t)} f_t(x, y, t) = \min_{y \in Y^*(t)} \max_{x \in X^*(t)} f_t(x, y, t) \quad \text{for} \quad t < 1,$$

$$V'(t-) = \min_{x \in X^*(t)} \max_{y \in Y^*(t)} f_t(x, y, t) = \max_{y \in Y^*(t)} \min_{x \in X^*(t)} f_t(x, y, t) \quad \text{for} \quad t > 0.$$

PROOF: Take $t_0 \in [0, 1)$, and a selection $(x^*(t), y^*(t)) \in X^*(t) \times Y^*(t)$. For any $t > t_0$ we can write

$$\frac{f(x^*(t_0), y^*(t), t) - f(x^*(t_0), y^*(t), t_0)}{t - t_0}$$
$$\leq \frac{V(t) - V(t_0)}{t - t_0} \leq \frac{f(x^*(t), y^*(t_0), t) - f(x^*(t), y^*(t_0), t_0)}{t - t_0}.$$

Therefore, by the Mean Value Theorem,

$$f_t(x^*(t_0), y^*(t), s'(t)) \leq \frac{V(t) - V(t_0)}{t - t_0} \leq f_t(x^*(t), y^*(t_0), s''(t))$$

for some $s'(t), s''(t) \in [t_0, t]$. This implies that

$$(6) \qquad \max_{x \in X^*(t_0)} f_t(x, y^*(t), s'(t)) \leq \frac{V(t) - V(t_0)}{t - t_0} \leq \min_{y \in Y^*(t_0)} f_t(x^*(t), y, s''(t)).$$

Berge's Maximum Theorem implies that $\max_{x \in X^*(t_0)} f_t(x, y, t)$ is continuous in (y, t) and $\min_{y \in Y^*(t_0)} f_t(x, y, t)$ is continuous in (x, t). The theorem also implies that the saddle set correspondence, being the Nash equilibrium correspondence of a zero-sum game, is upper hemicontinuous (see, e.g., Fudenberg and Tirole (1991)). These two observations imply that

$$\varliminf_{t \to t_0+} \max_{x \in X^*(t_0)} f_t(x, y^*(t), s'(t)) \geq \min_{y \in Y^*(t_0)} \max_{x \in X^*(t_0)} f_t(x, y, t_0),$$

$$\varlimsup_{t \to t_0+} \min_{y \in Y^*(t_0)} f_t(x^*(t), y, s''(t)) \leq \max_{x \in X^*(t_0)} \min_{y \in Y^*(t_0)} f_t(x, y, t_0).$$

[19] For the particular case where X and Y are unit simplexes representing the two players' mixed strategies in a finite zero-sum game, and hence the payoff $f(x, y, t)$ is bilinear in (x, y), the result has been obtained by Mills (1956).

Therefore, taking the limits inferior and superior in (6), we obtain

$$\min_{y \in Y^*(t_0)} \max_{x \in X^*(t_0)} f_t(x, y, t_0) \leq \varliminf_{t \to t_0+} \frac{V(t) - V(t_0)}{t - t_0} \leq \varlimsup_{t \to t_0+} \frac{V(t) - V(t_0)}{t - t_0}$$

$$\leq \max_{x \in X^*(t_0)} \min_{y \in Y^*(t_0)} f_t(x, y, t_0).$$

Since we also know that

$$\max_{x \in X^*(t_0)} \min_{y \in Y^*(t_0)} f_t(x, y, t_0) \leq \min_{y \in Y^*(t_0)} \max_{x \in X^*(t_0)} f_t(x, y, t_0)$$

(see, e.g., Rockafellar (1970)), the first result follows. The second result is established similarly. *Q.E.D.*

3.5. *Problems with Parameterized Constraints*

Consider the following maximization program with k parameterized inequality constraints:

$$V(t) = \sup_{x \in X: g(x,t) \geq 0} f(x, t), \quad \text{where} \quad g: X \times [0, 1] \to \mathbb{R}^k,$$

$$X^*(t) = \{x \in X : g(x, t) \geq 0, f(x, t) = V(t)\}.$$

It is well known (see, e.g., Luenberger (1969) and Rockafellar (1970)) that if X is a convex set, f and g are concave in x, and $g(\hat{x}, t) \gg 0$ for some $\hat{x} \in X$,[20] then the constrained maximization problem can be represented as a saddle-point problem for the associated Lagrangian. Specifically, letting $y \in \mathbb{R}^k_+$ be the vector of Lagrange multipliers corresponding to the k constraints, the Lagrangian can be written as

$$L(x, y, t) = f(x, t) + \sum_{i=1}^{k} y_i g_i(x, t).$$

The set of saddle points of the Lagrangian over $(x, y) \in X \times \mathbb{R}^k_+$ at parameter value t takes the form $X^*(t) \times Y^*(t)$, where $X^*(t)$ is the set of solutions to the above constrained maximization program, and $Y^*(t)$ is the set of solutions to the dual program:

$$Y^*(t) = \operatorname*{Arg\,min}_{y \in \mathbb{R}^k_+} \left(\sup_{x \in X} L(x, y, t) \right).$$

The value $V(t)$ of the constrained maximization problem equals the saddle value of the Lagrangian with parameter t. Application of Theorems 4 and 5 to this saddle-point problem yields the following corollary.

COROLLARY 5: *Suppose that X is a convex compact set in a normed linear space, f and g are continuous and concave in x, $f_t(x, t)$ and $g_t(x, t)$ are continuous*

---

[20] This means that all components of $g(\hat{x}, t)$ are strictly positive.

in (x, t), *and there exists* $\hat{x} \in X$ *such that* $g(\hat{x}, t) \gg 0$ *for all* $t \in [0, 1]$. *Then*:

(i) V *is absolutely continuous, and for any selection* $(x^*(t), y^*(t)) \in X^*(t) \times Y^*(t)$,

$$V(t) = V(0) + \int_0^t L_t(x^*(s), y^*(s), s)ds.$$

(ii) V *is directionally differentiable, and its directional derivatives equal*:

$$V'(t+) = \max_{x \in X^*(t)} \min_{y \in Y^*(t)} L_t(x, y, t) = \min_{y \in Y^*(t)} \max_{x \in X^*(t)} L_t(x, y, t) \quad \text{for} \quad t < 1,$$

$$V'(t-) = \min_{x \in X^*(t)} \max_{y \in Y^*(t)} L_t(x, y, t) = \max_{y \in Y^*(t)} \min_{x \in X^*(t)} L_t(x, y, t) \quad \text{for} \quad t > 0.$$

PROOF: For all $t \in [0, 1]$, all $y^* \in Y^*(t)$, and each $i = 1, 2, \ldots, k$ we can write

$$V(t) \geq L(\hat{x}, y^*, t) \geq f(\hat{x}, t) + y_i^* g_i(\hat{x}, t),$$

where the first inequality is by the definition of the saddle value, and the second by nonnegativity of Lagrange multipliers. This implies that

$$y_i^* \leq \bar{y}_i \equiv \sup_{t \in [0,1]} \frac{V(t) - f(\hat{x}, t)}{g_i(\hat{x}, t)}.$$

Observe that $\bar{y}_i < +\infty$, since the numerator of the above fraction is bounded, and the denominator is bounded away from zero by the definition of $\hat{x}$ and continuity of $g(\hat{x}, \cdot)$. Therefore, the set $Y = \prod_{i=1}^k [0, \bar{y}_i] \subset \mathbb{R}_+^k$ is compact. Since we have shown that $Y^*(t) \subset Y$ for all t, $X^*(t) \times Y^*(t)$ is the saddle set of the Lagrangian on $X \times Y$. The assumptions of Theorems 4 and 5 can now be verified, and the theorems yield the results.   *Q.E.D.*

A version of result (ii) was first obtained by Gol'shtein (1972). Also, note that in the particular case where $k = 1$, $g(x, t) = h(x) + t$, and $f(x, t) = f(x)$, it yields $V'(t+) = \min Y^*(t)$ and $V'(t-) = \max Y^*(t)$. This special case of Corollary 5(ii), which allows the interpretation of a Lagrange multiplier as the "price" of the constraint, is stated in Rockafellar (1970).

3.6. *Smooth Pasting in Optimal Stopping Problems*

Optimal stopping theory has become a standard tool in economics to model decisions involving "real" or financial options, such as when and whether to exercise an option to buy securities, convert a bond, harvest a crop, adopt a new technology, or terminate a research project (see, e.g., Dixit and Pindyck (1994)). In the usual formulation, the decision maker chooses a stopping time of a continuous time Markov process $\{z(\tau)\}$ with state space Σ and paths that are right-continuous. The decision maker's flow payoff at any time τ in state z is $\pi(z)$. If at any time the process is stopped in state z, the decision maker receives a termination payoff of $\Omega(z)$.

Suppose the decision maker adopts the Markovian policy of terminating whenever the state lies in the closed set S. Define $T_S = \inf\{\tau \mid z(\tau) \in S\}$ to be the first time that the process enters the set S. The decision maker's payoff is

$$\Pi(S) \equiv \int_0^{T_S} e^{-\rho s} \pi(z(s)) ds + e^{-\rho T_S} \Omega(z(T_S)),$$

with expected payoff beginning in state z_0 of $f(S, z_0) = E[\Pi(S)|z(0) = z_0]$. The optimal value function is $V(z_0) \equiv \sup_S f(S, z_0)$ and a policy S^* is *Markov optimal* if for all z_0, $V(z_0) \equiv f(S^*, z_0)$.

COROLLARY 6: *Suppose that Ω and V are differentiable and that a Markov optimal strategy S^* exists. Then, for all $z_0 \in S^*$, $V(z_0) = \Omega(z_0)$ and $V_z(z_0) = \Omega_z(z_0)$.*[21]

PROOF: For any $z_0 \in S^*$, $S = \Sigma$ ("always stop immediately") is an optimal policy beginning in z_0 and its value is $f(\Sigma, z_0) \equiv \Omega(z_0)$. Since Ω and V are differentiable, the conclusion follows from Theorem 1.   *Q.E.D.*

This conclusion is known as "smooth pasting," because it asserts that V melds smoothly into Ω. Economic models exploiting smooth pasting frequently assume that $\{z(\tau)\}$ is a Markov diffusion process satisfying the assumptions of Corollary 6.

The conditions that imply differentiability of the function $f(S, z_0)$ in z_0 are subtle (see Fleming and Soner (1993)) and frequently depend on properties of both the stochastic process and the payoff functions, but not on the optimality of the stopping set S. Given this technical structure, the advantage of the present treatment of smooth pasting is that it separates the issue of the differentiability of the value of Markov policies from the issue of the equality of two derivatives, which under such differentiability follows simply from the optimality of the stopping rule S^*.

4. CONCLUSION

It is common for economic optimization models to include a variety of mathematical assumptions to ease the analysis. These include the assumptions of convexity, differentiability of certain functions, and sign restrictions on the derivatives that are used for comparative statics analysis.

It has long been understood that one class of conclusions—those about the existence of prices supporting the optimum—depend only on the assumptions that are invariant to linear transformations of the choice variables, such as convexity. Similarly, as emphasized by Milgrom and Shannon (1994), directional comparative statics conclusions depend only on assumptions that are invariant to order-preserving transformations of the choice variables and the parameter,

---

[21] If the process is nonstationary, the corollary still applies with time as a component of the state variable.

such as supermodularity, quasisupermodularity, and single crossing. In that same spirit, the present paper is rooted in the observation that the absolute continuity or differentiability of the value function can depend only on assumptions that are invariant to any relabeling of the choice variables, even one that does not preserve convex, topological, or order structures of the choice set. Our general envelope theorems utilize only such assumptions. Taken together with the other cited results about the relation between economic hypotheses and conclusions, the new envelope theorems contribute to a deeper understanding of the overall structure of economic optimization models.

Dept. of Economics, Stanford University, Stanford, CA 94305, U.S.A.; paul@milgrom.net; http://www.milgrom.net

and

Dept. of Economics, Stanford University, Stanford, CA 94305, U.S.A.; ilya.segal@stanford.edu; http://www.stanford.edu/~isegal

REFERENCES

ALIPRANTIS, C. D., AND K. C. BORDER (1994): *Infinite Dimensional Analysis*. New York: Springer-Verlag.

APOSTOL, T. M. (1969): *Calculus*, Vol. II, Second Edition. New York: John Wiley & Sons.

ATHEY, S., P. MILGROM, AND J. ROBERTS (2000): *Robust Comparative Statics* (in preparation). Princeton: Princeton University Press.

BENVENISTE, L. M., AND J. A. SCHEINKMAN (1979): "On the Differentiability of the Value Function in Dynamic Models of Economics," *Econometrica*, 47, 727–732.

BONNANS, J. F., AND A. SHAPIRO (2000): *Perturbation Analysis of Optimization Problems*. New York: Springer-Verlag.

DANSKIN, J. M. (1967): *The Theory of Max-Min and Its Applications to Weapons Allocation Problems*. New York: Springer-Verlag.

DIXIT, A. K., AND R. S. PINDYCK (1994): *Investment Under Uncertainty*. Princeton: Princeton University Press.

FLEMING, W. H., AND H. M. SONER (1993): *Controlled Markov Processes and Viscosity Solutions*. New York: Springer-Verlag.

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. Cambridge: MIT Press.

GOL'SHTEIN, E. G. (1972): *Theory of Convex Programming*, Volume 36 of *Translations of Mathematical Monographs*. Providence: American Mathematical Society.

KIM, T. (1993): "Differentiability of the Value Function: A New Characterization," *Seoul Journal of Economics*, 6, 257–265.

KRISHNA, V., AND E. MAENNER (2001): "Convex Potentials with an Application to Mechanism Design," *Econometrica*, 69, 1113–1119.

LAFFONT, J.-J., AND E. MASKIN (1980): "A Differential Approach to Dominant-Strategy Mechanisms," *Econometrica*, 48, 1507–1520.

LUENBERGER, D. G. (1969): *Optimization by Vector Space Methods*. New York: John Wiley & Sons.

MAS-COLELL, A., M. WHINSTON, AND J. GREEN (1995): *Microeconomic Theory*. New York: Oxford University Press.

MILGROM, P., AND J. ROBERTS (1988): "Communication and Inventories as Substitutes in Organizing Production," *Scandinavian Journal of Economics*, 90, 275–289.

MILGROM, P., AND C. SHANNON (1994): "Monotone Comparative Statics," *Econometrica*, 62, 157–180.

MILLS, H. D. (1956): "Marginal Values of Matrix Games and Linear Programs," in *Linear Inequalities and Related Systems, Annals of Mathematical Studies*, 38, ed. by H. W. Kuhn and A. W. Tucker. Princeton: Princeton University Press, 183–193.

MIRRLEES, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *The Review of Economic Studies*, 38, 175–208.

MYERSON, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

——— (1991): *Game Theory*. Cambridge: Harvard University Press.

ROCKAFELLAR, R. T. (1970): *Convex Analysis*. Princeton: Princeton University Press.

ROYDEN, H. L. (1988): *Real Analysis*, Third Edition. Englewood Cliffs: Prentice-Hall.

SAH, R., AND J. ZHAO (1998): "Some Envelope Theorems for Integer and Discrete Choice Variables," *International Economic Review*, 39, 623–634.

SEGAL, I., AND M. WHINSTON (2002): "The Mirrlees Approach to Mechanism Design with Renegotiation: Theory and Application to Hold-Up and Risk Sharing," *Econometrica*, 70, 1–45.

SIMON, C. P., AND L. BLUME (1994): *Mathematics for Economists*. New York: W. W. Norton & Co.

WILLIAMS, S. R. (1999): "A Characterization of Efficient, Bayesian Incentive-Compatible Mechanisms," *Economic Theory*, 14, 155–180.