



**COLLANA DEL
DIPARTIMENTO DI ECONOMIA**

A NEW APPROACH TO THE ENVELOPE THEOREM

Francesco Ruscitti

Working Paper n° 137, 2011

I Working Papers del Dipartimento di Economia svolgono la funzione di divulgare tempestivamente, in forma definitiva o provvisoria, i risultati di ricerche scientifiche originali. La loro pubblicazione è soggetta all'approvazione del Comitato Scientifico.

Per ciascuna pubblicazione vengono soddisfatti gli obblighi previsti dall'art. 1 del D.L.L. 31.8.1945, n. 660 e successive modifiche.

Copie della presente pubblicazione possono essere richieste alla Redazione.

REDAZIONE:

Dipartimento di Economia
Università degli Studi Roma Tre
Via Silvio D'Amico, 77 - 00145 Roma
Tel. 0039-06-57335655 fax 0039-06-57335771
E-mail: dip_eco@uniroma3.it



DIPARTIMENTO DI ECONOMIA

A NEW APPROACH TO THE ENVELOPE THEOREM

Francesco Ruscitti

Comitato Scientifico:

F. De Filippis

A. Giunta

P. Lazzara

L. Mastroeni

S. Terzi

A NEW APPROACH TO THE ENVELOPE THEOREM*

FRANCESCO RUSCITTI[†]

Abstract

We study the differentiability of the value function of a constrained optimization problem. We consider the envelope-theorem framework of Milgrom and Segal (2002), and we accomplish two goals. We show how one can relax Milgrom and Segal's assumption that the choice set does not vary with parameters. More importantly, we develop a new approach to proving the differentiability of the value function. The key idea and main mathematical tool we employ in our approach are a novel feature in the literature dealing with the differentiability of the value function.

JEL classification: C60, C61, C65.

Keywords: value function, uniform convergence, differentiability, correspondences.

1 Introduction

This paper deals with envelope theorems for constrained maximization problems. That is, we study a parameterized maximization problem and we address the differentiability of the value function associated with the optimization problem. The standard textbook approach to the envelope theorem¹ posits that the objective function (and the functions defining the constraints) is twice continuously differentiable in the choice variables. This allows for the use of the implicit function theorem to obtain a local differentiable selection of maximizers. These assumptions are rather strong, and thus it would be desirable to relax them. To this end, Milgrom and Segal (2002) develop an approach to the envelope theorem under minimal assumptions on the objective function and the choice set. They also discuss several economic applications of their theorems. Indeed, Milgrom and Segal assume, in the first part of their paper, that the

*We thank David Khaski for reading the manuscript and drawing the figures.

[†]Department of Economics, University of Rome III, 77 Via Silvio D'Amico, 00145, Roma (Italy). E-mail: fruscitti@uniroma3.it

¹See, e.g., A. Mas-Colell et al. (1995), and H. Varian (1992).

choice set lacks any topological and algebraic structure, which is remarkable. In particular, Milgrom and Segal do not assume that the objective function is differentiable in the choice variables. On the other hand, the authors assume that the choice set *does not* vary with the parameter that appears as argument of the objective function. The latter assumption is quite stringent. Therefore, our first objective in this paper is to show that it is still possible to easily characterize the derivative of the value function even though we dispense with Milgrom and Segal’s assumption that the choice set is constant, and we do not assume that the objective function is differentiable in the choice variables. Specifically, unlike Milgrom et al. (2002), we do not require that the choice set be independent of the parameter. In our framework the feasible set is a ‘function’ of the given parameter, and we formalize this relation by means of a correspondence that maps the parameter space into the control-variables space. It turns out that we only need to assume that this correspondence satisfies fairly mild assumptions, such as lower hemicontinuity and convex-valuedness. We remark that this way of relating parameters to choice variables is very standard, especially in dynamic programming models. In contrast to Milgrom and Segal (2002), though, we need to put some structure on the choice set. Specifically, in the first part of section 3 we assume that the choice set is a convex subset of a finite-dimensional linear space. As in Milgrom and Segal (2002), we show that the derivative of the value function reduces to the well-known envelope theorem formula.

The second and main objective we pursue in this paper is as follows: we retain the framework of Milgrom and Segal (2002), and in the final part of section 3 and in section 4 we introduce another approach to the envelope theorem. Our approach is novel, to the best of our knowledge. It hinges on a simple theorem on uniform convergence, from real analysis, which has been overlooked in the literature about the differentiability of the value function. We assume that the choice set is an arbitrary compact metric space. Hence, notice that the choice set may be infinite-dimensional. The assumption that the choice set be a compact metric space ensures that the choice set is separable. In turn, separability plays a crucial role in our approach. Recall that an arbitrary compact topological space need not be separable. This is why we need the choice set to be a metric space. Also, we show that our approach can be nicely combined with the approach put forward in the seminal paper by Benveniste and Scheinkman (1979).

The paper is structured as follows: in section 2 we set up the background and notation, and we state three theorems that will be used to prove our results. In Section 3 we lay out the assumptions and illustrate how our approach works. In Section 4 we drop a key assumption and we replace it with a weaker one. We then show that combining our approach with Benveniste and Scheinkman’s still enables us to prove the differentiability of the value function under the milder assumption. Finally, in section 5 we point out some drawbacks of our approach, and we outline directions for future research.

2 Background

Following Milgrom and Segal (2002), the problem envelope theorems are concerned with is the following:

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

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

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

Note that in Milgrom and Segal (2002) the choice set X is independent of the parameter t .

In the sequel we shall invoke the following theorem from real analysis (see Sohrab, 2003, Theorem 8.3.4):

Theorem 1.2 *Let $(f_n)_{n=1}^{\infty}$ be a sequence of differentiable functions defined on $[a, b]$ such that $(f_n(x_0))_{n=1}^{\infty}$ converges for some $x_0 \in [a, b]$. If the sequence $(f'_n)_{n=1}^{\infty}$ of derivatives converges to a function g uniformly on $[a, b]$, then the sequence $(f_n)_{n=1}^{\infty}$ converges uniformly on $[a, b]$ to a differentiable function f and we have*

$$f'(x) = \lim_{n \rightarrow \infty} f'_n(x) = g(x) \text{ for all } x \in [a, b].$$

We shall also employ the following theorem from convex analysis (see Rockafellar, 1970). It was first used by Benveniste and Scheinkman (1979) to prove the differentiability of the value function arising in dynamic economic models.

Theorem 2.2 *Let f be a real valued concave function defined on a convex set $D \subseteq \mathbb{R}^n$. If W is a concave differentiable function in a neighborhood N of x_0 in D with the property that $W(x_0) = f(x_0)$ and $W(x) \leq f(x)$ for all x in N , then f is differentiable at x_0 . Moreover, the differential of f at x_0 coincides with the differential of W at x_0 .*

In what follows, $\text{int}(\cdot)$ denotes the interior in the range space of the correspondences at hand. To allow for the choice set to depend on the parameter t , and characterize the derivative of V , we shall make use of the following theorem:

Theorem 3.2 *Let $\Gamma : M \rightarrow \mathbb{R}^n$ be a lower hemi-continuous correspondence, where M is an arbitrary metric space. Suppose that Γ is convex-valued, and that for some $\bar{x} \in X$, $\text{int}\Gamma(\bar{x})$ is non-empty. Then, for any compact set $K \subset \text{int}\Gamma(\bar{x})$, there exists a neighborhood $V(\bar{x})$ of \bar{x} such that $K \subset \Gamma(x')$ for all $x' \in V(\bar{x})$.*

The above theorem was first proven by Zhou (1995). An alternative proof based on a simple separation argument can be found in Bagh et al. (2010). In the

special case in which M is finite-dimensional, a proof of Theorem 3.2 can be derived from Proposition 4.15 and Theorem 5.9 in Rockafellar and Wets (2009)². Note that when the compact set K is a singleton, Theorem 3.2 implies the following property of Γ . If $\bar{y} \in \text{int}\Gamma(\bar{x})$, then there exists a neighborhood $V(\bar{x})$ of \bar{x} such that $\bar{y} \in \Gamma(x')$ for all $x' \in V(\bar{x})$. This is the property we will need to use later on in the proof of Proposition 1.3.

3 The mathematical framework

Refer to the envelope-theorem framework set-forth above in the background paragraph. Next, we shall relax the assumption that the choice set is independent of the parameter, and we shall characterize the derivative of the value function. We shall think of X as a given control-variables space, but we shall let the feasible set, as a subset of X , vary with the parameter t . Theorem 3.2 will play a key role in accomplishing this. Here is how we formalize the correspondence relating the parameter to the admissible control variables.

Let $t \in [0, 1]$, and let $\Gamma : [0, 1] \rightarrow X$, where $X \subseteq \mathbb{R}^n$. Define the parameterized objective function by $f : G_\Gamma \rightarrow \mathbb{R}$, where

$$G_\Gamma := \{(t, x) \in [0, 1] \times X : x \in \Gamma(t)\}$$

is the graph of Γ .

We can now re-define the value function V and the (possibly empty-valued) optimal choice correspondence X^* as follows:

$$\begin{aligned} V(t) &= \sup_{x \in \Gamma(t)} f(t, x) \\ X^*(t) &= \{x \in \Gamma(t) : f(t, x) = V(t)\} \end{aligned}$$

Assumption 1.3 $\Gamma : [0, 1] \rightarrow X$ is lower hemicontinuous with convex values. Furthermore, there exists a $t \in [0, 1]$ such that $\text{int}\Gamma(t) \neq \emptyset$.

The following proposition should be compared to Theorem 1 in Milgrom and Segal (2002). By using Theorem 3.2 one can easily prove:

Proposition 1.3 Suppose that Assumption 1.3 holds. Assume that $x^* \in X^*(t)$, with $x^* \in \text{Int}\Gamma(t)$, and suppose that $\frac{\partial f}{\partial t}(t, x^*)$ exists. If $t > 0$ and V is left-hand differentiable at t , then $V'(t-) \leq \frac{\partial f}{\partial t}(t, x^*)$. If $t < 1$ and V is right-hand differentiable at t , then $V'(t+) \geq \frac{\partial f}{\partial t}(t, x^*)$. If $t \in (0, 1)$ and V is differentiable at t , then $V'(t) = \frac{\partial f}{\partial t}(t, x^*)$.

Proof. By Theorem 3.2, there exists a neighborhood of t in $[0, 1]$, say $\mathcal{N}(t)$, such that $x^* \in \Gamma(t')$ for each $t' \in \mathcal{N}(t)$. Clearly, for any $t' \in \mathcal{N}(t)$ we have

²We thank A. Bagh for pointing this out to us.

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

From this point onward we can simply follow the original proof of Theorem 1 in Milgrom and Segal (2002): for $t < 1$, we can just perform their argument on $(t, 1) \cap \mathcal{N}(t)$. For $t > 0$, we can just perform their argument on $(0, t) \cap \mathcal{N}(t)$. ■

Assumption 2.3 $\Gamma : [0, 1] \rightarrow X$ is lower hemicontinuous with convex graph. Furthermore, there exists a $t_0 \in (0, 1)$ such that $\text{int}\Gamma(t_0) \neq \emptyset$.

The following proposition should be compared to Corollary 3 in Milgrom and Segal (2002). In Corollary 3, Milgrom and Segal place additional structure on the choice set which is assumed to be convex. Also, the authors assume that the objective function is jointly concave and the choice set does not vary with the parameter. By exploiting Proposition 1.3 one can easily prove:

Proposition 2.3 Suppose that Assumption 2.3 holds, and let $f : G_\Gamma \rightarrow \mathbb{R}$ be concave. Assume that there exists a $x^* \in X^*(t_0)$, with $x^* \in \text{Int}\Gamma(t_0)$, such that $\frac{\partial f}{\partial t}(t_0, x^*)$ exists. Then V is differentiable at t_0 and $V'(t_0) = \frac{\partial f}{\partial t}(t_0, x^*)$.

Proof. Pick two arbitrary (t', x') , $(t'', x'') \in G_\Gamma$, and any $\lambda \in [0, 1]$. By convexity of G_Γ , concavity of f , and by definition of V , we have that:

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

Taking the supremum of the right-hand side over $x' \in \Gamma(t')$ and $x'' \in \Gamma(t'')$, yields

$$V(\lambda t' + (1 - \lambda)t'') \geq \lambda V(t') + (1 - \lambda)V(t'').$$

Hence, V is concave, and therefore V is directionally differentiable at t_0 with $V'(t_0+) \leq V'(t_0-)$ (see Rockafellar, 1970). On the other hand, by Proposition 1.3 we know that $V'(t_0-) \leq \frac{\partial f}{\partial t}(t_0, x^*) \leq V'(t_0+)$. Thus, V is differentiable at t_0 and $V'(t_0) = \frac{\partial f}{\partial t}(t_0, x^*)$. ■

Remark 1.3 Note that $V'(t_0) = \frac{\partial f}{\partial t}(t_0, x^*)$ is the standard envelope-theorem formula one obtains under the assumptions that x^* is an interior strict maximizer and f is twice differentiable with respect to x .

Let us now turn back to the original Milgrom and Segal's framework in which X is fixed (see the first part of the background section above). In what follows we shall present a new approach to the envelope theorem. For simplicity, and to illustrate the key idea underlying our approach, we begin with a set of quite strong assumptions on f . In section 4 below we shall relax these assumptions, and we shall use them in conjunction with some convexity assumptions to guarantee that the value function is still well-defined, continuous, and differentiable

at a specific point in the parameter space. We stress that, in what follows, X can well be infinite-dimensional.

Assumption 3.3 X is a compact metric space.

It follows from Assumption 3.3 that X is separable. Hence, there exists a dense subset of X , say T , which is countable. Put $T = \{x_1, x_2, \dots, x_n, \dots\}$.

Assumption 4.3 $f : [0, 1] \times X \rightarrow \mathbb{R}$ is continuous in the product topology (jointly continuous).

Given Assumptions 3.3 and 4.3, by Berge maximum theorem $V(t)$ is continuous. Now define

$$\Psi(t) = \sup_{x \in T} f(t, x).$$

The following result will come in handy:

Lemma 1.3 Under Assumption 3.3, if $V : [0, 1] \rightarrow \mathbb{R}$, and $f(t, \cdot) : X \rightarrow \mathbb{R}$ is lower semicontinuous for each $t \in [0, 1]$, then $\Psi = V$.

Proof: Fix an arbitrary $t \in [0, 1]$, and pick any $x_i \in T$. Clearly $f(t, x_i) \leq V(t)$. Thus, $\Psi : [0, 1] \rightarrow \mathbb{R}$ is well defined and $\Psi(t) \leq V(t)$. We claim that $\Psi(t) = V(t)$. For, if $\Psi(t) < V(t)$, then we can pick $\bar{x} \in X$ such that $\Psi(t) < f(t, \bar{x}) \leq V(t)$. Because T is dense in X , there exists a sequence in T , say (x_n) , that converges to \bar{x} . Therefore, by virtue of lower semicontinuity we can find a N such that $f(t, x_n) > \Psi(t)$ for each $n \geq N$, which is a contradiction. ■

Remark 2.3 Clearly the above result holds under Assumptions 3.3 and 4.3. Note that in view of Lemma 1.3 we can restrict attention to Ψ .

Next, define:

$$f_n : [0, 1] \rightarrow \mathbb{R} \text{ by } f_n(t) = f(t, x_n) \text{ for each } x_n \in T,$$

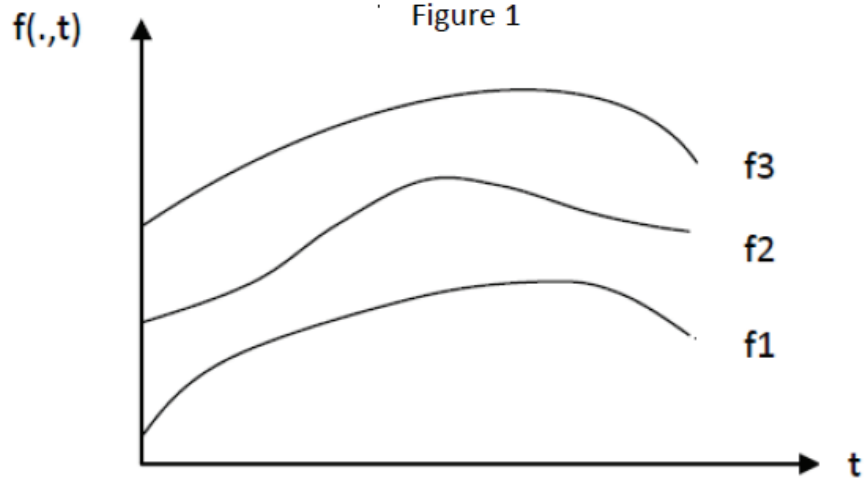
and note that $f_n \in C([0, 1])$ for each $n \in \mathbb{N}$ (this follows from Assumption 4.3).

Assumption 5.3 The sequence (f_n) is monotonically increasing, i.e., $f_1(t) \leq f_2(t) \dots$ for all $t \in [0, 1]$.

Figure 1 on the next page illustrates Assumption 5.3.

Lemma 2.3 If Assumptions 3.3, 4.3, and 5.3 hold, then $f_n \rightarrow \Psi$ uniformly.

Proof: From Lemma 1.3 we know that $\Psi = V$, and therefore $\Psi \in C([0, 1])$. By Assumption 5.3 it will suffice to prove that $f_n \rightarrow \Psi$ pointwise. Dini's theorem will then yield the desired result (see Aliprantis and Border, 2006, Theorem



2.66). To this end, fix an arbitrary $t \in [0, 1]$ and pick any $\varepsilon > 0$. By definition of $\Psi(t)$ there exists a N such that $-\varepsilon < \Psi(t) - f_N(t) < \varepsilon$. But the sequence (f_n) is increasing (Assumption 5.3), hence $-\varepsilon < \Psi(t) - f_n(t) < \varepsilon$ for each $n \geq N$. To finish the proof invoke Dini's theorem. ■

Hereafter we shall assume that $\frac{\partial f}{\partial t}(t, x_n) \in C([0, 1])$ for each n . Note that we do not require the objective function to be differentiable in the choice variables.

Assumption 6.3 *The sequence $\left(\frac{\partial f}{\partial t}(t, x_n)\right)$ is equicontinuous and uniformly bounded.*

Proposition 3.3 *If Assumptions 3.3, 4.3, 5.3, and 6.3. hold, then the value function V is continuously differentiable on $(0, 1)$.*

Proof: By Assumption 6.3 the sequence $\left(\frac{\partial f}{\partial t}(t, x_n)\right)$ has a subsequence, say $\left(\frac{\partial f}{\partial t}(t, x_{n_k})\right)$, that converges uniformly (this follows from Arzela'-Ascoli theorem). Clearly, (f_{n_k}) converges uniformly to Ψ (see Assumption 5.3 and Lemma 2.3). Since $\Psi = V$ (see Lemma 1.3), it follows from Theorem 1.2 that V is differentiable on $(0, 1)$, and $\frac{dV}{dt}(t)$ is equal to the uniform limit of $\left(\frac{\partial f}{\partial t}(t, x_{n_k})\right)$. Finally, recall that $C([0, 1])$ equipped with the uniform metric is complete. This implies that the uniform limit of $\left(\frac{\partial f}{\partial t}(t, x_{n_k})\right)$ belongs to $C([0, 1])$. Hence, $\frac{dV}{dt}(t)$ is continuous. ■

4 Combining our approach with Benveniste and Scheinkman's

In this section we relax Assumption 5.3 which, arguably, is quite stringent. We replace it with a milder assumption, and we show how one can use Theorem 1.2 in conjunction with Theorem 2.2 to prove the differentiability, at a certain point, of the value function of the parameterized optimization problem at hand. In what follows, X need not be a linear space in its own right. Hence, again we stress that X can be infinite-dimensional.

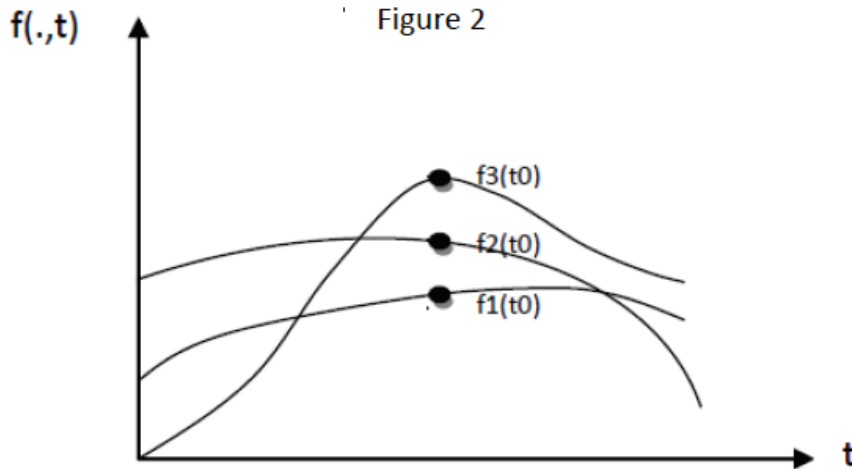
Assumption 1.4 X is a convex and compact subset of a metric linear space.

Assumption 2.4 $f : [0, 1] \times X \rightarrow \mathbb{R}$ is jointly concave and continuous in the product topology.

Clearly, Lemma 1.3 still holds, so we can restrict attention to Ψ . We retain the assumption $\frac{\partial f}{\partial t}(t, x_n) \in C([0, 1])$ for each n , and Assumption 6.3. Instead of Assumption 5.3, we posit:

Assumption 3.4 There exists a $t_0 \in (0, 1)$ such that the sequence $(f_n(t_0))_{n=1}^{\infty}$ is monotonically increasing, i.e., $f_1(t_0) \leq f_2(t_0) \dots$.

Figure 2 below illustrates Assumption 3.4.



Proposition 1.4 If Assumptions 1.4, 2.4, 3.4, and 6.3 hold, then V is differentiable at t_0 .

Proof: By assumptions 1.4 and 2.4, $\Psi = V$ is concave on $[0, 1]$ (see Corollary 3 in Milgrom and Segal, 2002). Given Assumption 3.4, it should be clear that $(f_n(t_0))_{n=1}^{\infty}$ converges to $\Psi(t_0)$ (see the proof of Lemma 2.3). By virtue of Assumption 6.3, the sequence $\left(\frac{\partial f}{\partial t}(t, x_{n_k})\right)_{k=1}^{\infty}$ converges uniformly. Clearly,

$(f_{n_k}(t_0))_{k=1}^\infty$ converges to $\Psi(t_0)$. From Theorem 1.2 it follows that the sequence $(f_{n_k})_{k=1}^\infty$ converges uniformly on $[0, 1]$ to a differentiable function f , thus $(f_{n_k})_{k=1}^\infty$ converges pointwise to f . In particular, $(f_{n_k}(t_0))_{k=1}^\infty$ converges to $f(t_0)$. Therefore we must have $\Psi(t_0) = f(t_0)$. By Assumption 2.4, $f_{n_k}(t)$ is concave on $[0, 1]$ for each k . Since $(f_{n_k})_{k=1}^\infty$ converges uniformly to f , then f is concave on $[0, 1]$ as well³. Now we claim that $f(t) \leq \Psi(t)$ for each $t \in [0, 1]$. Indeed, assume, by way of contradiction, that there is a $\bar{t} \in [0, 1]$ such that $f(\bar{t}) > \Psi(\bar{t})$. But since $(f_{n_k}(\bar{t}))_{k=1}^\infty$ converges to $f(\bar{t})$, for k large enough we get $f_{n_k}(\bar{t}) > \Psi(\bar{t})$, which is impossible. Finally, by theorems 1.2 and 2.2, V is differentiable at t_0 , and $\frac{dV}{dt}(t_0)$ is equal to $\lim_{k \rightarrow \infty} \frac{\partial f}{\partial t}(t_0, x_{n_k})$. ■

5 Concluding remarks and future research

This paper ought to be viewed as a work in progress that needs to be improved. It still exhibits a few weaknesses which we summarize briefly.

First and foremost, it remains to verify that our construction is well-posed, that is that the results herein obtained are invariant to the choice of the dense and countable subset of the feasible set (given the assumptions, there exists at least one such a subset).

Secondly, the paper lacks examples that illustrate the scope of the assumptions and the applicability of the model. For instance, it would be interesting to construct economic models (presumably dynamic programming models) for which Milgrom and Segal's approach does not work but our approach does work.

Finally, one might want to relax the assumption that the optimal solution is interior, and allow for boundary solutions. This issue has been already studied by Rincon-Zapatero et al. (2009) in a dynamic programming context.

References

- Aliprantis, C. D., and Border K. C. (2006): "Infinite Dimensional Analysis. A Hitchhiker's Guide", third Edition. New York: Springer-Verlag.
- Bagh A., Impicciatore G., and Ruscitti F. (2010): "The open graph theorem for correspondences revisited: applications and counter-examples", *mimeo*.
- Benveniste, L. M., Scheinkman J.A. (1979): "On the Differentiability of the Value Function in Dynamic Models of Economics", *Econometrica* 47, 727-732.
- Mas-Colell, Whinston M.D., and Green J.R. (1995): "Microeconomic Theory", Oxford University Press, New York-Oxford.
- Milgrom, P., and Segal I. (2002): "Envelope Theorems for Arbitrary Choice Sets", *Econometrica* 70, 583-601.

³Recall that uniform convergence preserves concavity. See, e.g., Stokey et al. (1989, Theorem 4.8).

- Rincon-Zapatero, J.P., Santos, M.S. (2009), “Differentiability of the value function without interiority assumptions”, *Journal of Economic Theory* 144, 1948-1964.
- Rockafellar, R. T. (1970): “Convex Analysis”, Princeton, NJ: Princeton University Press.
- Rockafellar R.T., and Wets R. (2009): “Variational Analysis”, third edition, Vol. 317, Springer.
- Sohrab H. (2003): “Basic Real Analysis”, Birkhauser, Boston, Basel, Berlin: Birkhauser.
- Stokey, N., Lucas R.E., and Prescott E (1989): “Recursive Methods in Economic Dynamics”, Cambridge, MA: Harvard University Press.
- Varian, Hal R. (1992): “Microeconomic Analysis”, third edition, Norton International Student Edition.
- Zhou, J., (1995), “On the existence of equilibrium for abstract economies”, *Journal of Mathematical Analysis and Applications* 193, 839-858.