



CLARKE DIFFERENTIALS AND THE ENVELOPE THEOREM IN DYNAMIC PROGRAMMING

YUHKI HOSOYA

Faculty of Economics, Chuo University, Tokyo, Japan

Abstract. In this paper, we consider a deterministic dynamic programming model, and derive the envelope theorem by using the Clarke differential. Comparing with previous research, we do not require differentiability, convexity, or boundedness.

Keywords. Clarke differential; Dynamic programming; Envelope theorem; Value function.

2020 Mathematics Subject Classification. 90C39, 91B55.

1. INTRODUCTION

Although the term ‘envelope theorem’ covers many different results in various fields of applied mathematics, one of the simplest is the following. First, consider the one-dimensional parametrized optimization problem of $f(x, p)$ with respect to x , and let $x(p)$ be the solution function. Define the value function $V(p) = f(x(p), p)$. Suppose that f is differentiable at $(x(p), p)$, and that x is also differentiable at p . Then,

$$V'(p) = \frac{\partial f}{\partial x}(x(p), p)x'(p) + \frac{\partial f}{\partial p}(x(p), p).$$

However, by the first-order condition, we have that $\frac{\partial f}{\partial x}(x(p), p) = 0$, which leads to the following result:

$$V'(p) = \frac{\partial f}{\partial p}(x(p), p).$$

This is the envelope theorem. Although this result can be extended to many optimization problems, its proof is usually heavily dependent on the differentiability of f and x . Hence, when f is not differentiable, this result may be inapplicable.

On the other hand, in dynamic programming models, the value function $\bar{V}(x)$ satisfies the following **Bellman equation**:

$$\bar{V}(x) = \max\{w(x, y) + \delta \bar{V}(y) | y \in \Gamma(x)\}.$$

E-mail address: ukki@gs.econ.keio.ac.jp

Received June 3, 2025; Accepted November 27, 2025.

If the envelope theorem can be applied, then we have that

$$\bar{V}'(x) = \frac{\partial w}{\partial x}(x, y^*(x)),$$

where $y^*(x)$ is an optimizer for the maximization problem on the right-hand side of the Bellman equation. However, there is a problem. To prove this result, the objective function $w(x, y) + \delta \bar{V}(y)$ must be differentiable. Because w is the exogenously determined objective function, we can assume the differentiability of w . However, $\bar{V}$ is a function derived endogenously from the model, and thus it is difficult to exogenously assume its differentiability. Furthermore, even if $\bar{V}$ is differentiable, the above proof of the envelope theorem requires the differentiability of y^* . If this problem is not resolved, the envelope theorem cannot be applied.

Theorem 1 of Benveniste and Scheinkman [1] is the first result to overcome the above difficulty. They introduced some convex structure in the model and showed the envelope theorem without assuming the differentiability of $\bar{V}$ and y^* using the techniques of convex analysis. This result is referred to in many macroeconomic textbooks as the ‘Benveniste–Scheinkman envelope theorem’.

Although this result is good, several of the assumptions used are somewhat strong. Hence, several relaxed results of this ‘envelope theorem’ have been proved. One such result is Theorem 5.2 of Mordukhovich and Sagara [4]. They used the Clarke differential introduced in Clarke [2] instead of convex analysis. Their result can be applied to stochastic dynamic programming models in which the variable is in some separable Banach space, and thus the scope of applicable models became much wider.

However, a problem remains. Although Mordukhovich–Sagara included stochastic models in their application, stochastic dynamic programming models include several technical problems that typically do not arise in deterministic models. For example, the value function may not be measurable in some models. To address this issue, they have made some strong assumptions. For example, the domain of the objective function must be the entire space, and the value of this function must be bounded. These assumptions are not valid in many economic applications.¹

In this paper, we restrict our discussion to deterministic models and reproduce the results of Mordukhovich and Sagara [4] under weaker assumptions. The main result is Theorem 3.3, which proves the envelope theorem under very weak assumptions that are satisfied by many economic models.

This paper makes extensive use of the Clarke differential. To make the paper self-contained, we introduce several fundamental results related to the Clarke differential with proofs. Of course, most of these results have already been shown by Clarke [2]. For the sake of readability, the proofs of these results are included in the appendix.

The structure of this paper is as follows. In Section 2, we define several terms, including the Clarke differential, and discuss their properties. In Section 3, we discuss the main results. A brief conclusion is presented in Section 4.

¹As will be discussed in detail in subsection 3.2, there exists a one-parameter family of functions frequently appearing in economic theory named CRRA, which includes the natural logarithm function. Since the natural logarithm function neither has the entire real line as its domain nor is bounded, it is highly incompatible with Mordukhovich–Sagara’s assumption.

2. DEFINITIONS AND PRELIMINARY RESULTS

Throughout this section, the symbols X, Y basically denote some Banach spaces and X' denotes the dual space of X . For $x \in X$ and $r > 0$, define $B_r(x) = \{y \in X \mid \|y - x\| < r\}$. This set $B_r(x)$ is called the **open ball centered at x with radius r** . Similarly, define $\bar{B}_r(x) = \{y \in X \mid \|y - x\| \leq r\}$. This set $\bar{B}_r(x)$ is called the **closed ball centered at x with radius r** .

2.1. Locally Lipschitz Functions. Let $U \subset X$ be a nonempty set and suppose that $f : U \rightarrow Y$ is given. This function f is said to be **L -Lipschitz** if $\|f(x) - f(y)\| \leq L\|x - y\|$ holds for any $x, y \in U$. A function $f : U \rightarrow Y$ is said to be **Lipschitz** if and only if there exists $L > 0$ such that f is L -Lipschitz. We call such an L a **Lipschitz constant** for f . If $C \subset U$ is also nonempty, then f is said to be **Lipschitz on C** if and only if the restriction of f on C is Lipschitz. A natural extension to this concept is as follows. $f : U \rightarrow Y$ is said to be **locally Lipschitz** if for each $x \in U$, there exists $r > 0$ such that f is Lipschitz on $B_r(x) \cap U$. We need the following result.

Proposition 2.1. *1) Every locally Lipschitz function $f : U \rightarrow Y$ is Lipschitz on any compact subset $C \subset U$.
2) If $X = \mathbb{R}^n$ and $U \subset \mathbb{R}^n$ is open, then $f : U \rightarrow Y$ is locally Lipschitz if and only if f is Lipschitz on any compact subset $C \subset U$.*

Next, recall the definition of several notions for derivatives. A function $T : X \rightarrow Y$ is said to be **linear** if and only if for each $x, y \in X$ and $a, b \in \mathbb{R}$, $T(ax + by) = aT(x) + bT(y)$. Note that, when T is linear, it is customary to write Tx instead of $T(x)$. If Y is $\mathbb{R}$, then Tx is also written as $\langle T, x \rangle$. It is well-known that if $T : X \rightarrow Y$ is linear, then T is continuous if and only if T is Lipschitz, and if we define

$$\|T\| = \inf\{L > 0 \mid L \text{ is a Lipschitz constant of } T\} = \sup\{\|Tx\| \mid \|x\| \leq 1\},$$

then $\|\cdot\|$ satisfies all axioms of the norm, and the set of all continuous linear functions $L(X, Y)$ becomes a Banach space with this norm $\|\cdot\|$. This norm is called the **operator norm** of T . In particular, $L(X, \mathbb{R})$ is called the **dual space** of X , denoted by X' .

Suppose that $U \subset X$ is a nonempty open set, and that $f : U \rightarrow Y$ and $x \in U$ are given. Suppose that there exists $T \in L(X, Y)$ such that, for each $v \in X$,

$$\lim_{t \downarrow 0} \frac{f(x + tv) - f(x)}{t} = Tv.$$

Then, T is called the **Gâteaux derivative** of f at x , denoted by $D_G f(x)$. Moreover, if the above convergence is uniform in v on every compact set C , then T is called the **Hadamard derivative** of f at x , denoted by $D_H f(x)$. Finally, if the above convergence is uniform in v on every bounded set B , then T is called the **Fréchet derivative** of f at x , denoted by $Df(x)$. If the mapping $x \mapsto D_G f(x)$ can be defined and is continuous, then we can show that $D_G f(x) = Df(x)$.² In this case, f is said to be **continuously differentiable**, or C^1 . Using the mean value theorem, we can easily show that every C^1 function is locally Lipschitz. If $X = \mathbb{R}^n, Y = \mathbb{R}^m$, then the ‘almost everywhere’ converse relationship holds. That is, if $U \subset X$ is nonempty and open, then every locally Lipschitz function $f : U \rightarrow Y$ is Fréchet differentiable almost everywhere (Rademacher’s theorem). For a detailed argument, see Heinonen [3].

²See Proposition 2.7.

A set $C \subset X$ is said to be **convex** if and only if $x, y \in C$ and $t \in [0, 1]$ implies $(1-t)x + ty \in C$. Suppose that $U \subset X$ is a nonempty and convex subset. A real-valued function $f : U \rightarrow \mathbb{R}$ is said to be **convex** if and only if the epigraph $E = \{(x, a) \in U \times \mathbb{R} | f(x) \leq a\}$ is convex, and, f is said to be **concave** if and only if $-f$ is convex. If $X = \mathbb{R}^n$, then it is well-known that any convex (or, concave) function defined on an open set is locally Lipschitz. See Theorem 10.4 of Rockafellar [5].

2.2. Clarke Derivatives. We start from the definition of generalized directional derivatives. Suppose that $U \subset X$ is a nonempty open subset, a real-valued function $f : U \rightarrow \mathbb{R}$ is given, and $x \in U, v \in X$. Define

$$f^\circ(x; v) = \limsup_{y \rightarrow x, t \downarrow 0} \frac{f(y + tv) - f(y)}{t}.$$

We call this value $f^\circ(x; v)$ the **generalized directional derivative of f at x with direction v** . The following result is useful.³

Proposition 2.2. *Suppose that $U \subset X$ is a nonempty open set and $f : U \rightarrow \mathbb{R}$ is locally Lipschitz. For a fixed $x \in U$, $f^\circ(x; v)$ is well-defined and a real number. If f is L -Lipschitz around x , then $|f^\circ(x; v)| \leq L\|v\|$. Furthermore, the mapping $v \mapsto f^\circ(x; v)$ is sublinear.*

Now, suppose that $U \subset X$ is a nonempty open set, and $f : U \rightarrow \mathbb{R}$ is locally Lipschitz. Define a set-valued function $\partial^\circ f : X \rightrightarrows X'$ as $\partial^\circ f(x) = \{p \in X' | \forall v \in X, \langle p, v \rangle \leq f^\circ(x; v)\}$. We call this function $\partial^\circ f$ the **Clarke differential** of f . Using Proposition 2.2, we can easily check that $\partial^\circ f(x)$ is nonempty, convex, and weak* compact. First, because $v \mapsto f^\circ(x; v)$ is sublinear, the Hahn-Banach theorem can be applied, and thus there exists a linear functional $p : X \rightarrow \mathbb{R}$ such that $\langle p, v \rangle \leq f^\circ(x; v)$ for any v . Now, because f is assumed to be locally Lipschitz, there exists $L > 0$ such that f is L -Lipschitz around x . By Proposition 2.2, $|f^\circ(x; v)| \leq L\|v\|$, and thus p is L -Lipschitz. This implies that $p \in X'$, and thus $\partial^\circ f(x)$ is nonempty. It is clear that $\partial^\circ f(x)$ is convex and weak* closed. Because $p \in \partial^\circ f(x)$ implies that p is L -Lipschitz, $\partial^\circ f(x)$ is bounded. By Alaoglu's theorem, $\partial^\circ f(x)$ is weak* compact.

The following proposition is very useful.

Proposition 2.3. *Suppose that $U \subset X$ is a nonempty open subset of a Banach space X , and $f : U \rightarrow \mathbb{R}$ is locally Lipschitz. Then, $-\partial^\circ f(x) = \partial^\circ(-f)(x)$.*

Using this proposition, we can show the following fundamental result for the optimization theory.

Corollary 2.4. *Suppose that U is a nonempty open subset of a Banach space X , and $f : U \rightarrow \mathbb{R}$ is locally Lipschitz. Choose $x^* \in U$. If x^* is either a local maximum point or a local minimum point, then $0 \in \partial^\circ f(x^*)$.*

Proof. First, suppose that x^* is a local minimum point. Then, $f^\circ(x^*; v) \geq 0$ by definition, and thus, $0 \in \partial^\circ f(x)$. Next, suppose that x^* is a local maximum point. Then, x^* is a local minimum point of $-f$. Therefore, $0 \in \partial^\circ(-f)(x^*)$. By Proposition 2.3,

$$0 = -0 \in -\partial^\circ(-f)(x^*) = \partial^\circ f(x^*).$$

This completes the proof. □

³A function $\varphi : X \rightarrow \mathbb{R}$ is **sublinear** if 1) for any $a \geq 0$, $\varphi(av) = a\varphi(v)$, and 2) $\varphi(v+w) \leq \varphi(v) + \varphi(w)$.

NOTE. In several textbooks, $\partial^\circ f(x)$ is called the Clarke **subdifferential**. Actually, there is another notion of generalized differentials that is called the Clarke **superdifferential**, which was introduced in Sagara [6]. Define

$$f^*(x; v) = \liminf_{y \rightarrow x, t \downarrow 0} \frac{f(y + tv) - f(y)}{t}$$

and

$$\partial^* f(x) = \{p \in X' \mid \langle p, v \rangle \geq f^*(x; v) \text{ for all } v \in X\}.$$

This $\partial^* f(x)$ is called the Clarke superdifferential.

However, this ‘Clarke superdifferential’ is actually the same as the ‘Clarke subdifferential’. That is, $\partial^* f(x) = \partial^\circ f(x)$. To show this, we first note the following equation:

$$\begin{aligned} -f^*(x; v) &= -\liminf_{y \rightarrow x, t \downarrow 0} \frac{f(y + tv) - f(y)}{t} \\ &= \limsup_{y \rightarrow x, t \downarrow 0} \frac{(-f)(y + tv) - (-f)(y)}{t} = (-f)^\circ(x; v). \end{aligned}$$

Therefore, for any $v \in X$,

$$\begin{aligned} \langle p, v \rangle \geq f^*(x; v) &\Leftrightarrow \langle p, v \rangle \geq -(-f)^\circ(x; v) \\ &\Leftrightarrow \langle -p, v \rangle \leq (-f)^\circ(x; v). \end{aligned}$$

In particular,

$$p \in \partial^* f(x) \Leftrightarrow -p \in \partial^\circ (-f)(x) = -\partial^\circ f(x),$$

which implies that $\partial^* f(x) = \partial^\circ f(x)$. Therefore, the Clarke ‘subdifferential’ coincides with the ‘superdifferential’, and hence we simply call $\partial^\circ f(x)$ the Clarke ‘differential’.

We additionally note the relationship between the usual ‘subdifferential’ and ‘Clarke differential’. Suppose that U is an open and convex subset of a Banach space X . By definition, $f : U \rightarrow \mathbb{R}$ is said to be **convex** if and only if, for any $x, y \in U$ and $t \in [0, 1]$,

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

If f is convex, then the set

$$\partial f(x) = \{p \in X' \mid \langle p, v \rangle \leq f(x+v) - f(x) \text{ for all } v \in X\}$$

is called the **subdifferential** for the function f at x . Then, the following proposition holds.⁴

Proposition 2.5. *Suppose that U is a nonempty, open, and convex subset of a Banach space X , and $f : U \rightarrow \mathbb{R}$ is convex and locally Lipschitz. Then, $\partial f(x) = \partial^\circ f(x)$.*

2.3. Strong Differentiability. For a subdifferential of the concave function $f : I \rightarrow \mathbb{R}$ defined on an open interval I , $\partial f(x)$ is a singleton if and only if f is differentiable at x . For the Clarke differential, however, this result does not hold. For example, consider the following function

$$f(x) = \begin{cases} x^2 \sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

⁴Note that, if $X = \mathbb{R}^n$, then any convex function $f : U \rightarrow \mathbb{R}$ is locally Lipschitz.

If $x \neq 0$, then $f'(x) = 2x \sin(1/x) - \cos(1/x)$, and thus, $|f'(x)| \leq 2x + 1$. Next,

$$f'(0) = \lim_{h \rightarrow 0} \frac{f(h) - f(0)}{h} = \lim_{h \rightarrow 0} h \sin(1/h) = 0.$$

By the mean value theorem, we have that f is $(2M + 1)$ -Lipschitz on $[-M, M]$, and thus it is locally Lipschitz. However, we can show that $\partial^\circ f(0) = [-1, 1]$, and thus, $\partial^\circ f(0) \neq \{f'(0)\}$ in this case.

To solve this problem, we introduce the notion of **strong differentiability**. Suppose that U is a nonempty open set of a Banach space X , and $f : U \rightarrow \mathbb{R}$ is locally Lipschitz. Suppose that there exists $p \in X'$ such that

$$\lim_{y \rightarrow x, t \downarrow 0} \frac{f(y + tv) - f(y)}{t} = \langle p, v \rangle$$

for all $v \in X$. Then, f is said to be **strongly differentiable** at x , and p is called the **strong derivative** of f at x , denoted by $D_s f(x)$.

We additionally introduce the **regularity** of f at x . Suppose that U is a nonempty open subset of a Banach space X , and $f : U \rightarrow \mathbb{R}$ is locally Lipschitz. We define the classical directional derivative $f'(x; v)$ as

$$f'(x; v) = \lim_{t \downarrow 0} \frac{f(x + tv) - f(x)}{t}.$$

The function f is said to be regular at x if $f^\circ(x; v) = f'(x; v)$ for any $v \in X$.

The following proposition clarifies the relationship between Clarke derivative, usual derivatives, the strong derivative, and the regularity.

Proposition 2.6. *Suppose that U is a nonempty open subset of a Banach space X , $f : U \rightarrow \mathbb{R}$ is locally Lipschitz, and $x \in U$. Then, the following results hold.*

- (i) *If f is Gâteaux differentiable at x , then $D_G f(x) \in \partial^\circ f(x)$.*
- (ii) *$\partial^\circ f(x)$ is a singleton if and only if f is strongly differentiable. In this case, $\partial^\circ f(x) = \{D_s f(x)\}$.*
- (iii) *If f is strongly differentiable at x , then it is Hadamard differentiable at x , and $D_s f(x) = D_H f(x)$.⁵*
- (iv) *If f is strongly differentiable at x , then f is regular at x .*

The above different notions on differentiability can be simplified by assuming continuous differentiability. That is, the following proposition holds.

Proposition 2.7. *Suppose that U is a nonempty open subset of a Banach space X , and $f : U \rightarrow \mathbb{R}$ is Gâteaux differentiable at any point. Moreover, suppose that $D_G f : U \rightarrow X'$ is continuous at $x \in U$. Then, f is both Fréchet differentiable and strongly differentiable at x , and $Df(x) = D_s f(x) = D_G f(x)$.*

Finally, we present a proposition that ensures the regularity of f at x .

Proposition 2.8. *Suppose that U is a nonempty open and convex subset of a Banach space X , and $f : U \rightarrow \mathbb{R}$ is locally Lipschitz and either concave or convex. Then, f is regular at any point.*

⁵If $X = \mathbb{R}^n$, then $D_H f(x) = Df(x)$, and thus, $D_s f(x) = Df(x)$.

3. MAIN RESULTS

3.1. **The Model.** We consider the following model:

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \delta^t w(x_t, x_{t+1}) \\ & \text{subject to } x_{t+1} \in \Gamma(x_t), \\ & \quad x_0 = \bar{x} \in W, \end{aligned} \tag{3.1}$$

where W is a subset of a Banach space X , $\bar{W}$ is the closure of W , the function $w : T \rightarrow \mathbb{R} \cup \{-\infty\}$ and the set-valued function $\Gamma : \bar{W} \rightrightarrows \bar{W}$ are given (where T denotes the graph of Γ), and $\delta \in]0, 1[$.

This type of model is referred to as a reduced form model, and many economic dynamic models can be transformed into this form in some way. Therefore, the reduced form model is more general than each specified model, which means that the conclusions obtained from it are more powerful. However, to examine the natural assumptions for w , Γ , and W , we must consider what w, Γ, W represent in economic models. Hence, we present a specific example of representative dynamic economic model and examine the properties of w , Γ , and W in this model.

We introduce the discretized version of the Ramsey–Cass–Koopmans type capital accumulation model (hereafter, RCK model). This model is represented by the following optimization problem.

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \delta^t u(c_t) \\ & \text{subject to } k_t \geq 0, c_t \geq 0, \\ & \quad c_t + i_t = f(k_t), \\ & \quad k_{t+1} = (1 - d)k_t + i_t, \\ & \quad k_0 = \bar{k} > 0. \end{aligned} \tag{3.2}$$

Here, c_t represents consumption, i_t represents investment, and k_t represents the capital stock. The proportion d represents the capital depreciation rate. At the end of t -th period, some of the capital stock is broken, but some is newly created. dk_t represents the amount of the capital stock that is broken, and i_t represents the amount of the capital stock that is newly created. Hence, k_{t+1} is equal to $(1 - d)k_t + i_t$. The assumptions $k_t \geq 0, c_t \geq 0$ are natural. Although $i_t \geq 0$ is also natural, we ignore this assumption for simplicity. The function f is the production function, and the gross product $f(k_t)$ is distributed by consumption c_t and investment i_t . The function u is called the **utility function**, which evaluates the wellness of the consumption level. Hence, the RCK model represents the optimization problem for the plan of capital accumulation using the evaluation function u .

In this model, $i_t = f(k_t) - c_t$, and thus it can be omitted. Hence, $c_t = f(k_t) + (1 - d)k_t - k_{t+1}$, and thus c_t can also be omitted. Therefore, this model can be transformed into the following model

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \delta^t u(f(k_t) + (1 - d)k_t - k_{t+1}) \\ & \text{subject to } k_t \geq 0, f(k_t) + (1 - d)k_t - k_{t+1} \geq 0, k_0 = \bar{k} > 0. \end{aligned}$$

Thus, if we define

$$\begin{aligned} w(x, y) &= u(f(x) + (1 - d)x - y), \\ \Gamma(x) &= \{y \in \mathbb{R} | 0 \leq y \leq f(x) + (1 - d)x\}, \\ W = \mathbb{R}_{++} &\equiv \{x \in \mathbb{R} | x > 0\}, \bar{W} = \mathbb{R}_+ \equiv \{x \in \mathbb{R} | x \geq 0\}, \end{aligned}$$

then this model coincides with (3.1).

In (3.2), the typical examples of u, f are as follows:

$$\begin{aligned} u(c) = u_\theta(c) &\equiv \begin{cases} \log c & \text{if } \theta = 1, \\ \frac{c^{1-\theta}-1}{1-\theta} & \text{if } \theta \neq 1, \end{cases} \\ f(k) &= k^a \text{ or } f(k) = Ak, \end{aligned}$$

where $\theta > 0$, $0 < a < 1$, and $A > 0$. u_θ is the solution to the following differential equation:

$$-\frac{xu''(x)}{u'(x)} = \theta, \quad u'(1) = 1, \quad u(1) = 0.$$

Because the left-hand side is called the **relative risk aversion** of u , this function is called the **constant relative risk aversion** (CRRA) function. The function $f(k) = k^a$ is called the **Cobb–Douglas technology**, and $f(k) = Ak$ is called the **AK technology**. These are very common in macroeconomic dynamics. Hence, we must make suitable assumptions for w, Γ, W that do not prohibit these examples. This means that w is not necessarily bounded, and the value may be $-\infty$ (because $u(0)$ may be $-\infty$).

3.2. The Result. We consider the following functional equation on the unknown function V :

$$V(x) = \max\{w(x, y) + \delta V(y) | y \in \Gamma(x)\}. \quad (3.3)$$

This equation is called the **Bellman equation**.

Let

$$\bar{V}(x) = \sup \left\{ \sum_{t=0}^{\infty} \delta^t w(x_t, x_{t+1}) \mid x_{t+1} \in \Gamma(x_t) \text{ for all } t \geq 0 \right\}.$$

This function $\bar{V} : W \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$ is called the **value function** of (3.1). In many cases, $\bar{V}$ solves the Bellman equation (3.3). If so, define

$$G(x) = \arg \max\{w(x, y) + \delta \bar{V}(y) | y \in \Gamma(x)\}.$$

The multi-valued function G is called the **policy function** of (3.1). It is well-known that (x_t) is a solution to (3.1) if and only if $x_{t+1} \in G(x_t)$ for all $t \geq 0$. See ch.4 of Stokey and Lucas [7] for detailed arguments.

The following theorem was proved by Benveniste and Scheinkman [1], and is so-called the Benveniste–Scheinkman envelope theorem.

Theorem 3.1. *Suppose that $X = \mathbb{R}^n$. Let T be the graph of Γ and $\text{int}.T$ be the interior of T . Suppose that the following requirements hold.*

- i) T is convex and $\text{int}.T$ is nonempty.
- ii) w is continuous and concave on T , and real-valued and C^1 on $\text{int}.T$.
- iii) For any $\bar{x} \in W$, there exists $\bar{y} \in G(\bar{x}) \cap \text{int}.T$.

Then, $\bar{V}$ is C^1 on W , and for $\bar{y} \in G(\bar{x}) \cap \text{int}.T$,

$$\bar{V}'(\bar{x}) = \frac{\partial w}{\partial x}(\bar{x}, \bar{y}).$$

However, the requirements in Theorem 3.1 are sometimes strong, and thus it is desirable to relax the assumptions. In this respect, Mordukhovich and Sagara [4] derived the following theorem. First, fix $y \in \bar{W}$ and define

$$w_y(x) = w(x, y), \quad \partial_x^\circ w(x, y) = \partial^\circ w_y(x).$$

Theorem 3.2. *Let X be a separable Banach space, $W = X$, $w : W^2 \rightarrow \mathbb{R}$ be continuous and bounded, and $\Gamma : W \rightrightarrows W$ be nonempty-valued, compact-valued, and continuous. Choose any $\bar{x} \in W$, and suppose that the following requirements hold.*

- i) *There exists an open neighborhood W' of $\bar{x}$ such that if $x, x' \in W'$, then $G(x) \neq \emptyset$ and $G(x) \subset \Gamma(x')$.*
- ii) *w is Lipschitz around (x, y) for all $x \in W'$ and $y \in G(x)$.*
- iii) *For every $\bar{y} \in G(\bar{x})$, w is regular at $(\bar{x}, \bar{y})$.⁶*

Then, $\bar{V}$ is locally Lipschitz around $\bar{x}$, and for $\bar{y} \in G(\bar{x})$, $\partial^\circ \bar{V}(\bar{x}) \subset \partial_x^\circ w(\bar{x}, \bar{y})$.

In fact, Mordukhovich–Sagara’s result is applied to a **stochastic reduced form model**. In such models, $\bar{V}$ is not necessarily measurable unless we do not introduce some strong assumptions.⁷ Therefore, the boundedness of w on the entire space X^2 is necessary for their result. However, in deterministic models, such assumptions are too strong. In fact, the boundedness assumption of w excludes the typical u of the RCK model. Therefore, we relax such assumptions.

Recall the definition of the regularity. A locally Lipschitz real-valued function f is said to be regular at x iff $f'(x; v) = f^\circ(x; v)$ for all $v \in X$. The following is our main result.

Theorem 3.3. *Let X be a Banach space, $W \subset X$ be nonempty and open, and $\Gamma : \bar{W} \rightrightarrows \bar{W}$ be nonempty-valued. Suppose that $w : T \rightarrow \mathbb{R} \cup \{-\infty\}$ is defined and continuous (where, $T = \{(x, y) \in \bar{W}^2 | y \in \Gamma(x)\}$), and that $\bar{V} : \bar{W} \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$ is real-valued on W . Fix $\bar{x} \in W$, and suppose that the following requirements hold.*

- i) *There exists an open neighborhood $W' \subset W$ of $\bar{x}$ such that Γ is compact-valued and continuous on W' , and if $x, x' \in W'$, then $G(x) \cap \Gamma(x') \neq \emptyset$.*
- ii) *w is real-valued and locally Lipschitz around (x, y) for all $x \in W'$ and $y \in \text{cl}.G(x)$.*
- iii) *For every $\bar{y} \in \text{cl}.G(\bar{x}) \cap W$, w is regular at $(\bar{x}, \bar{y})$.*
- iv) *$G(\bar{x}) \subset W$ and G is upper hemi-continuous at $\bar{x}$.⁸*

Then, $\bar{V}$ is Lipschitz around $\bar{x}$, and for some $\bar{y} \in \text{cl}.G(\bar{x})$, $\partial^\circ \bar{V}(\bar{x}) \subset \partial_x^\circ w(\bar{x}, \bar{y})$.

⁶Mordukhovich and Sagara [4] stated that this theorem holds if i), ii) holds and w is regular at $(\bar{x}, \bar{y})$ for **some** $\bar{y} \in G(\bar{x})$. However, we believe that under this condition, their proof of this theorem includes a gap. To be specific, since we cannot choose $\bar{y}$ appearing in the fifth paragraph of the proof of Theorem 3.3, the ‘for some’ assumption cannot ensure that w is regular at $(\bar{x}, \bar{y})$ for this $\bar{y}$.

⁷See, for example, chapter 9 of Stokey and Lucas [7].

⁸This is an additional requirement in this theorem. Because w is not bounded, the Bellman operator is not a contraction in this theorem, and thus there may be a solution to (3.3) other than $\bar{V}$. Hence, the usual arguments for verifying the continuity of $\bar{V}$ using the fixed point theorem cannot be applied in this theorem, and this additional requirement is needed. Note that, if $\bar{V}$ is continuous on some neighborhood of $G(\bar{x})$, then this requirement automatically holds.

Proof. Define

$$\Pi(x) = \{(x_t) \in \bar{W}^{\mathbb{N}} | x_0 = x, x_{t+1} \in \Gamma(x_t) \text{ for all } t\}.$$

Because Γ is nonempty-valued, $\Pi(x) \neq \emptyset$. Let W' satisfy all requirements in i). First, choose any $x \in W'$ and $y \in \Gamma(x)$. Choose $(y_t) \in \Pi(y)$, and let $x_0 = x, x_1 = y$, and $x_t = y_{t-1}$ for $t \geq 2$. Then, $(x_t) \in \Pi(x)$ and $x_1 = y$. Therefore,

$$\bar{V}(x) \geq \sum_{t=0}^{\infty} \delta^t w(x_t, x_{t+1}) = w(x, y) + \delta \sum_{t=0}^{\infty} \delta^t w(y_t, y_{t+1}).$$

Taking the supremum of the right-hand side, we have $\bar{V}(x) \geq w(x, y) + \delta \bar{V}(y)$. Because $\bar{V}(x) \in \mathbb{R}$ by assumption, for any $\varepsilon > 0$, there exists $(x_t) \in \Pi(x)$ such that $\sum_{t=0}^{\infty} \delta^t w(x_t, x_{t+1}) \geq \bar{V}(x) - \varepsilon$. Let $x_1 = y$ and $y_t = x_{t+1}$. Then, $(y_t) \in \Pi(y)$, and thus,

$$\bar{V}(x) - \varepsilon \leq w(x, y) + \delta \sum_{t=0}^{\infty} \delta^t w(y_t, y_{t+1}) \leq w(x, y) + \delta \bar{V}(y).$$

In conclusion, we obtain $\bar{V}(x) \geq \sup\{w(x, y) + \delta \bar{V}(y) | y \in \Gamma(x)\} \geq \bar{V}(x) - \varepsilon$, and hence $\bar{V}(x) = \sup\{w(x, y) + \delta \bar{V}(y) | y \in \Gamma(x)\}$. By i), $G(x)$ is nonempty, and if $y^* \in G(x)$, then

$$\bar{V}(x) = w(x, y^*) + \delta \bar{V}(y^*) = \max\{w(x, y) + \delta \bar{V}(y) | y \in \Gamma(x)\},$$

which implies that $\bar{V}$ satisfies (3.3) on W' .

Second, because $\Gamma(\bar{x})$ is compact, we have that $\text{cl}.G(\bar{x})$ is also compact. For each $y \in \text{cl}.G(\bar{x})$, there exists an open neighborhood U_y of $(\bar{x}, y)$ and $L_y > 0$ such that w is L_y -Lipschitz on U_y . By the compactness of $\text{cl}.G(\bar{x})$, there exists $y_1, \dots, y_k \in \text{cl}.G(\bar{x})$ such that $\cup_{i=1}^k U_{y_i}$ is an open neighborhood of $\{\bar{x}\} \times \text{cl}.G(\bar{x})$. To shrink W' if necessary, there exists an open neighborhood U of $\text{cl}.G(\bar{x})$, $W' \times U \subset \cup_{i=1}^k U_{y_i}$. Because G is upper hemi-continuous at $\bar{x}$, we can assume without loss of generality that $G(x) \subset U$ for all $x \in W'$. Define $L = \max_i L_{y_i}$.

Third, let $x, x' \in W'$, and choose $y \in G(x) \cap \Gamma(x')$. Then, $\bar{V}(x) = w(x, y) + \delta \bar{V}(y)$ and $\bar{V}(x') \geq w(x', y) + \delta \bar{V}(y)$. Hence, $\bar{V}(x) - \bar{V}(x') \leq w(x, y) - w(x', y) \leq L\|x - x'\|$. By the symmetrical argument, we have $\bar{V}(x') - \bar{V}(x) \leq L\|x - x'\|$. Thus $|\bar{V}(x) - \bar{V}(x')| \leq L\|x - x'\|$. Therefore, $\bar{V}$ is L -Lipschitz on W' . Choose any $v \in X$, and suppose that (x_n) is a sequence on W' that converges to $\bar{x}$, (t_n) is a positive monotone sequence that converges to 0, and

$$\lim_{n \rightarrow \infty} \frac{\bar{V}(x_n + t_n v) - \bar{V}(x_n)}{t_n} = \bar{V}^\circ(\bar{x}; v).$$

Let $z_n = x_n + t_n v$. Because $z_n \rightarrow \bar{x}$ as $n \rightarrow \infty$, we can assume that $z_n \in W'$ for all n . Therefore, there exists $y_n \in G(z_n) \cap \Gamma(x_n)$. Because G is upper hemi-continuous, we can assume without loss of generality that (y_n) converges to $\bar{y} \in \text{cl}.G(\bar{x})$. Then, $\bar{V}(z_n) = w(z_n, y_n) + \delta \bar{V}(y_n)$ and

$$\bar{V}(z_n - t_n v) \geq w(z_n - t_n v, y_n) + \delta \bar{V}(y_n).$$

	Domain	Concavity	C^1	Boundedness	RCK
Theorem 3.1(B-S)	Convex domain	yes	yes	no	CRRA
Theorem 3.2(M-S)	Banach space	no	no	yes	no
Theorem 3.3(ours)	Open domain	no	no	no	CRRA+ α

TABLE 1. Comparison of Theorems

Thus

$$\begin{aligned}
\bar{V}^\circ(\bar{x}; v) &= \lim_{n \rightarrow \infty} \frac{\bar{V}(z_n) - \bar{V}(z_n - t_n v)}{t_n} \\
&\leq \limsup_{n \rightarrow \infty} \frac{w(z_n, y_n) - w(z_n - t_n v, y_n)}{t_n} \\
&= \limsup_{n \rightarrow \infty} \frac{w(x_n + t_n v, y_n) - w(x_n, y_n)}{t_n} \\
&\leq w^\circ(\bar{x}, \bar{y}; v, 0).
\end{aligned}$$

On the other hand, because w is regular at $(\bar{x}, \bar{y})$,

$$\begin{aligned}
w_{\bar{y}}^\circ(\bar{x}; v) &= \limsup_{x \rightarrow \bar{x}, t \downarrow 0} \frac{w(x + tv, \bar{y}) - w(x, \bar{y})}{t} \leq \limsup_{(x, y) \rightarrow (\bar{x}, \bar{y}), t \downarrow 0} \frac{w(x + tv, y) - w(x, y)}{t} \\
&= w^\circ(\bar{x}, \bar{y}; v, 0) = w'(\bar{x}, \bar{y}; v, 0) = \lim_{t \downarrow 0} \frac{w(\bar{x} + tv, \bar{y}) - w(\bar{x}, \bar{y})}{t} \\
&\leq \limsup_{x \rightarrow \bar{x}, t \downarrow 0} \frac{w(x + tv, \bar{y}) - w(x, \bar{y})}{t} = w_{\bar{y}}^\circ(\bar{x}; v).
\end{aligned}$$

Therefore, $w^\circ(\bar{x}, \bar{y}; v, 0) = w_{\bar{y}}^\circ(\bar{x}; v)$, and if $p \in \partial^\circ \bar{V}(\bar{x})$, then $\langle p, v \rangle \leq w_{\bar{y}}^\circ(\bar{x}; v)$, which implies that $p \in \partial_x^\circ w(\bar{x}, \bar{y})$. This completes the proof. $\square$

Table 1 compares the main assumptions of Theorems 1-3. Although it does not compare every minor assumption, it should be useful for a rough comparison. The ‘RCK’ column indicates which types of RCK models can be handled. In RCK, neither the domain of u nor the domain of f extends to the entire real line. Therefore, the domain requirement in Theorem 2 is highly incompatible with RCK. Even if this could be remedied, CRRA remains incompatible because this function is unbounded under any $\theta > 0$. Meanwhile, the assumptions in Theorem 1 are compatible with CRRA. However, Theorem 3 can handle a wider range of economic situations. Later, we will explain the difference between these two theorems in subsection 3.3.

The following additional result is also useful.

Corollary 3.4. *In addition to the assumptions of Theorem 3.3, suppose that $w_{\bar{y}}$ is strongly differentiable at $\bar{x}$, where $\bar{y}$ is given in Theorem 3.3. Then, $\bar{V}$ is also strongly differentiable and Hadamard differentiable at $\bar{x}$. Moreover, if $G(\bar{x}) \subset \Gamma(x)$ for all $x \in W'$, then, for any $y \in G(\bar{x}) \cap W$, $D_H \bar{V}(\bar{x}) \in \partial_x^\circ w(\bar{x}, y)$.*

Proof. By Proposition 2.6, $\partial_x^\circ w(\bar{x}, \bar{y})$ is a singleton $\{D_s w_{\bar{y}}(\bar{x})\}$. Therefore, $\partial^\circ \bar{V}(\bar{x}) = \{D_s w_{\bar{y}}(\bar{x})\}$, which implies that $\bar{V}$ is both strong differentiable and Hadamard differentiable at $\bar{x}$. Next, suppose that $G(\bar{x}) \subset \Gamma(x)$ for all $x \in W'$. Choose any $y \in G(\bar{x})$ and $v \in X$. Let (t_n) be a positive

monotone sequence that converges to 0. We can assume without loss of generality that $\bar{x} + t_n v \in W'$ for all n . Thus, $\bar{V}(\bar{x}) = w(\bar{x}, y) + \delta \bar{V}(y)$ and $\bar{V}(\bar{x} + t_n v) \geq w(\bar{x} + t_n v, y) + \delta \bar{V}(y)$. Hence,

$$\begin{aligned} \langle D_H \bar{V}(\bar{x}), v \rangle &= - \lim_{n \rightarrow \infty} \frac{\bar{V}(\bar{x}) - \bar{V}(\bar{x} + t_n v)}{t_n} \\ &\leq - \limsup_{n \rightarrow \infty} \frac{w(\bar{x}, y) - w(\bar{x} + t_n v, y)}{t_n} \\ &= \liminf_{n \rightarrow \infty} \frac{w(\bar{x} + t_n v, y) - w(\bar{x}, y)}{t_n} \\ &= w'(\bar{x}, y; v, 0) = w^\circ(\bar{x}, y; v, 0). \end{aligned}$$

By the same arguments as in the proof of Theorem 3.3, we have $w^\circ(\bar{x}, y; v, 0) = w_y^\circ(\bar{x}; v)$. Therefore, $D_H \bar{V}(\bar{x}) \in \partial_x^\circ w(\bar{x}, y)$, as desired. This completes the proof. $\square$

Property i) of Theorem 3.3 may be a little difficult to understand. Hence, we provide a sufficient condition for ensuring this property.

Lemma 3.5. *Let X be a topological space, $W \subset X$ be nonempty and open, $\Gamma : \bar{W} \rightrightarrows \bar{W}$ be nonempty-valued and compact-valued on W , $T = \{(x, y) | x \in \bar{W}, y \in \Gamma(x)\}$, $w : T \rightarrow \mathbb{R} \cup \{-\infty\}$ be continuous, and $\bar{V} : \bar{W} \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$ be a continuous solution to (3.3) on $\bar{W}$ and real-valued on W . Moreover, suppose that there exists a continuous real-valued function $H : \bar{W}^2 \rightarrow \mathbb{R}$ such that if $H(x, y) \leq 0$, then $y \in \Gamma(x)$. Fix $\bar{x} \in W$, and suppose that $H(\bar{x}, \bar{y}) < 0$ for all $\bar{y} \in G(\bar{x})$. Then, there exists an open neighborhood $W' \subset W$ of $\bar{x}$ such that if $x, x' \in W'$, then $G(x) \neq \emptyset$ and $G(x) \subset \Gamma(x')$.⁹*

Proof. First, because $\bar{V}$ and w are continuous, by Berge's maximum theorem, G is nonempty-valued, compact-valued, and upper hemi-continuous on W . Because $G(\bar{x})$ is compact, there exists $\varepsilon > 0$ such that $\max_{y \in G(\bar{x})} H(\bar{x}, y) < -\varepsilon$. Because $\{\bar{x}\} \times G(\bar{x})$ is compact in X^2 , there exist open sets W_1, U such that $\{\bar{x}\} \times G(\bar{x}) \subset W_1 \times U$ and $H(x, y) < 0$ for all $x \in W_1, y \in U$. Because G is upper hemi-continuous, there exists an open neighborhood W_2 of $\bar{x}$ such that if $x \in W_2$, then $G(x) \subset U$. Let $W' = W_1 \cap W_2$. Then, W' is an open neighborhood of $\bar{x}$. Suppose that $x, x' \in W'$. Because G is nonempty-valued, $G(x) \neq \emptyset$. Because $x \in W_2$, $G(x) \subset U$, and because $x' \in W_1$, $H(x', y) < 0$ for all $y \in U$, which implies that $U \subset \Gamma(x')$. Therefore, $G(x) \subset \Gamma(x')$, as desired. This completes the proof. $\square$

To understand this lemma, consider a model in which $W = \mathbb{R}_+$ and $\Gamma = [0, F(x)]$ for some continuous and increasing function F such that $F(0) = 0$. Suppose that for each x sufficiently near to $\bar{x}$, G is single-valued and continuous at x , and $G(x) < F(x)$ for all $y \in G(x)$. Define $H(x, y) = y - F(x)$. Then, all the requirements in Lemma 3.5 hold. In this case, if $\varepsilon > 0$ is sufficiently small, then for all $x, x' \in W' \equiv [\bar{x} - \varepsilon, \bar{x} + \varepsilon]$, $G(x) < F(x')$, and thus $G(x) \subset \Gamma(x')$. In particular, the RCK model with a non-boundary optimal path must satisfy this requirement, and thus i) of Theorem 3.3 is naturally met.

⁹If X is an infinite-dimensional Banach space, then any compact subset of this space has an empty interior. On the other hand, under this theorem, $\Gamma(\bar{x})$ is a compact set and $G(\bar{x})$ is included in its interior. Therefore, for the norm topology, this result can only be applied to finite dimensional cases. However, because this result only requires that X has a topology, this result can also be applicable for an infinite dimensional Banach space with a weaker topology than the norm topology.

3.3. Application to Economic Models. In many economic models, Theorem 3.1 is applicable. Theorem 3.3 can be seen as a variant of Theorem 3.1. We now investigate the relationship between these theorems.

First, consider the RCK model (3.2). Recall that $W = \mathbb{R}_{++} \equiv \{k \in \mathbb{R} | k > 0\}$ and $\bar{W} = \mathbb{R}_+ \equiv \{k \in \mathbb{R} | k \geq 0\}$. To simplify the expression, we define $F(k) = f(k) + (1-d)k$. As we have seen, this model can be transformed into the following reduced form model:

$$\begin{aligned} & \max \sum_{t=0}^{\infty} \delta^t u(F(k_t) - k_{t+1}), \\ & \text{subject to. } 0 \leq k_{t+1} \leq F(k_t) \text{ for all } t \geq 0, \\ & k_0 = \bar{k} > 0. \end{aligned} \tag{3.4}$$

The following result is fundamental.

Proposition 3.6. *Let $u : \mathbb{R}_+ \rightarrow \mathbb{R} \cup \{-\infty\}$ be continuous and increasing on $\mathbb{R}_+$.¹⁰ Suppose also that $F : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is continuous and increasing, and satisfies $F(0) = 0$ and $\limsup_{k \rightarrow +\infty} \frac{F(k)}{k} < 1$. Then, there exists a solution (k_t^*) to (3.4) and $\bar{V}(\bar{k}) < +\infty$, and the value function $\bar{V}$ is continuous on $\mathbb{R}_+$. Furthermore, if we assume that there exists $\varepsilon^* > 0$ such that $F(k) > k$ whenever $0 < k \leq \varepsilon^*$, then $\bar{V}(\bar{k}) \in \mathbb{R}$ for all $\bar{k} > 0$.*

To ensure readability, we put the proof of this proposition into the appendix. Here, we briefly outline the idea of the proof. First, under the assumption $\limsup_{k \rightarrow +\infty} \frac{F(k)}{k} < 1$, if $K^* > 0$ is sufficiently large, the set of sequences that satisfies the constraints in (3.4) must belong to $[0, K^*]^{\mathbb{N}}$. By Tychonoff's theorem, this set is compact with respect to the product topology. Therefore, if the objective function is continuous with respect to this topology, then the existence of a solution follows, and $\bar{V}(\bar{k}) < +\infty$. By a standard argument using Berge's theorem, we can show the continuity of $\bar{V}$. If $F(k) > k$ for sufficiently small $k > 0$, then a sequence (k_t) such that $k_t \equiv k$ for $t \geq 1$ is feasible in (3.4). Thus, the value of the objective function is real-valued, and thus $\bar{V}(\bar{k}) > -\infty$, as desired.

Hence, under very weak assumptions, (3.4) has a solution, $\bar{V}$ is continuous, and $\bar{V}(\bar{k}) \in \mathbb{R}$ for all $\bar{k} > 0$. In this case, G is nonempty-valued, compact-valued, and upper hemi-continuous.

Now, suppose that u is continuous, increasing, and strictly concave on $\bar{W}$, C^1 on W , and $\lim_{c \rightarrow 0} u'(c) = +\infty$. Moreover, suppose that f is continuous, increasing, and concave on $\bar{W}$, C^1 on W , and satisfies 1) $f(0) = 0$, 2) $\limsup_{k \rightarrow \infty} \frac{f(k)}{k} < d$, and 3) there exists $\varepsilon^* > 0$ such that if $0 < k \leq \varepsilon^*$, then $f(k) > dk$. Then, $F(k) = f(k) + (1-d)k$ satisfies that 1) $F(0) = 0$, 2) $\limsup_{k \rightarrow \infty} \frac{F(k)}{k} < 1$, and 3) there exists $\varepsilon^* > 0$ such that if $0 < k \leq \varepsilon^*$, then $F(k) > k$. By Proposition 3.6, there exists a solution to (3.4), and $\bar{V}$ is finite and continuous on W . Moreover,

$$\Gamma(k) = [0, F(k)], \quad T = \{(k, k') | k \geq 0, k' \geq 0, F(k) \geq k'\},$$

and thus T is convex and its interior is nonempty. If $(k, k') \in \text{int.} T$, then $k' > 0$ and $F(k) - k' > 0$. Therefore, $w(k, k') = u(F(k) - k')$ is continuous and concave on T , and C^1 on $\text{int.} T$. Next, choose any $\bar{k} > 0$. Then, there exists $\varepsilon > 0$ such that $\delta u'(F(\varepsilon)) > u'(F(\bar{k}) - \varepsilon)$. Note that, because $F(0) = 0$, $k_t \equiv 0$ is optimal if $k_0 = 0$. Hence, if $u(0) = -\infty$, then $\bar{V}(0) = -\infty < \bar{V}(\varepsilon)$, and thus

$$u(F(\bar{k})) + \delta \bar{V}(0) = -\infty < u(F(\bar{k}) - \varepsilon) + \delta \bar{V}(\varepsilon),$$

¹⁰By this assumption, we have that $u(c) > -\infty$ whenever $c > 0$.

which implies that $0 \notin G(\bar{k})$. If $u(0) > -\infty$, then

$$\begin{aligned} u(F(\bar{k})) + \delta \bar{V}(0) &= u(F(\bar{k})) + \sum_{t=1}^{\infty} \delta^t u(0) \\ &< u(F(\bar{k}) - \varepsilon) + \delta u(F(\varepsilon)) + \sum_{t=2}^{\infty} \delta^t u(0) \\ &\leq u(F(\bar{k}) - \varepsilon) + \delta \bar{V}(\varepsilon), \end{aligned}$$

which implies that $0 \notin G(\bar{k})$. Next, we can easily check that $\bar{V}$ is concave, and thus there exists $L > 0$ such that $\bar{V}$ is L -Lipschitz on $]F(\bar{k}) - \varepsilon, F(\bar{k}) + \varepsilon[$. Choose $\varepsilon' > 0$ so small that $\varepsilon' < \varepsilon$ and $u'(\varepsilon') > L$. Then,

$$u(0) + \delta \bar{V}(F(\bar{k})) < u(\varepsilon') + \delta \bar{V}(F(\bar{k}) - \varepsilon'),$$

which implies that $F(\bar{k}) \notin G(\bar{k})$. Therefore, this model satisfies all requirements of Theorem 3.1.¹¹

On the other hand, let $H(\bar{k}, k) = -k(F(\bar{k}) - k)$. Then, $H(\bar{k}, k) \leq 0$ implies that $k \in \Gamma(\bar{k})$, and $H(\bar{k}, k) < 0$ for all $k \in G(\bar{k})$. Hence, Lemma 3.5 is applicable and requirement i) of Theorem 3.3 holds. By proposition 2.7, $w(k, k') = u(F(k) - k')$ is regular at $(\bar{k}, k)$ for all $k \in G(\bar{k})$, and thus requirement iii) holds. Requirements ii) and iv) can easily be verified, and thus this model also satisfies all requirements of Theorem 3.3. That is, the usual model of the form (3.4) admits all requirements of both theorems.

Next, we drop the C^1 assumption for u and F . In this case, u and F are still concave, and any concave function on an open and convex set of $\mathbb{R}^n$ is locally Lipschitz. Therefore, $w(k, k') = u(F(k) - k')$ is continuous and concave on T , and locally Lipschitz on $\text{int}.T$. It is easy to show that $\bar{V}$ is concave, and is thus locally Lipschitz. By Berge's maximum theorem, we have that $G(k)$ is nonempty-valued and upper semi-continuous. If, in addition, we assume that $\lim_{c \rightarrow 0} \min \partial u(c) = +\infty$, then by almost the same arguments as in the previous paragraph, we can show that $0, F(\bar{k}) \notin G(\bar{k})$. Define $H(\bar{k}, k) = -k(F(\bar{k}) - k)$. Then, $H(\bar{k}, k) \leq 0$ implies that $k \in \Gamma(\bar{k})$, and $H(\bar{k}, k) < 0$ for any $k \in G(\bar{k})$. By Lemma 3.5, requirement i) of Theorem 3.3 holds. It is easy to check that requirements ii) and iv) are satisfied, and because of Proposition 2.8, iii) holds. Hence, Theorem 3.3 is applicable.

Third, we consider the following alternative situation. Suppose that there is the government, and her policy is represented by some **government expenditure rule** $g(k)$. Because some of y_t is used for $g(k_t)$, the difference equation of k_t has changed: the new equation is

$$k_{t+1} = f(k_t) + (1 - d)k_t - g(k_t) - c_t.$$

Hence, we must define $F(k) = f(k) + (1 - d)k - g(k)$, and if g is not convex, then F may not be concave, even if g is smooth. Even in this case, the arguments in the previous paragraph can be applicable and we can easily check that i) and iv) of Theorem 3.3 hold. If u is concave, then it is locally Lipschitz. Hence, if F is locally Lipschitz, then $w(k, k') = u(F(k) - k')$ is also locally Lipschitz and ii) holds. Thus, again if w is regular at $(\bar{k}, k)$ for all $k \in G(\bar{k})$, then Theorem 3.3 is applicable. On the other hand, Theorem 3.1 is inapplicable even when u and F are C^1 because T may not be convex.

¹¹In economics, it is usual that u is assumed to be C^2 and f is assumed to be C^1 . Hence, the assumptions stated in this paragraph are weaker than the usual economic assumptions.

In conclusion, Theorem 3.3 is applicable for many models that represent natural economic situations, and in some of these models, Theorem 3.1 is inapplicable.

4. CONCLUSION

In this paper, we examined the envelope theorem for dynamic programming models, and derived a relaxed envelope theorem for these models. Although this result is related to Morukhovich and Sagara [4], the assumption is weakened. We think that this result will be useful in many areas of economic research.

APPENDIX A. PROOFS OF SEVERAL RESULTS

A.1. Proof of Proposition 2.1. First, we show 1). Suppose that $f : U \rightarrow Y$ is locally Lipschitz, $C \subset U$ is compact, and f is not Lipschitz on C . Then, for each positive integer n , there exists $x_n, y_n \in C$ such that $\|f(x_n) - f(y_n)\| > n\|x_n - y_n\|$. Because C is compact and $x_n \in C$ for all n , we can assume without loss of generality that there exists $x^* = \lim_{n \rightarrow \infty} x_n$. By the triangle inequality,

$$\begin{aligned} n\|x_n - y_n\| &< \|f(x_n) - f(y_n)\| \leq \|f(x_n) - f(x^*)\| + \|f(y_n) - f(x^*)\| \\ &\leq 2 \max\{\|f(z) - f(x^*)\| \mid z \in C\} \equiv A < +\infty, \end{aligned}$$

which implies that

$$\|y_n - x^*\| \leq \|y_n - x_n\| + \|x_n - x^*\| \leq \frac{A}{n} + \|x_n - x^*\| \rightarrow 0.$$

Therefore, we have that $\lim_{n \rightarrow \infty} y_n = x^*$. Because f is locally Lipschitz, there exist $r > 0$ and $L > 0$ such that f is L -Lipschitz on $B_r(x^*) \cap U$. Because $x_n, y_n \in U$ and x_n, y_n converge to x^* , there exists n such that $n > L$ and $x_n, y_n \in B_r(x^*)$. However, this implies that

$$n\|x_n - y_n\| < \|f(x_n) - f(y_n)\| \leq L\|x_n - y_n\|,$$

which is a contradiction.

Next, we show 2). By 1), the ‘only if’ part has already been proved. Hence, it suffices to show the ‘if’ part. Suppose that f is Lipschitz on any compact subset $C \subset U$. Choose any $x \in U$. Because U is an open set, there exists $r > 0$ such that $B_{2r}(x) \subset U$. Because $X = \mathbb{R}^n$, $\bar{B}_r \subset B_{2r}(x)$ is a compact set, and thus, f is Lipschitz on this set. Hence, there exists $L > 0$ such that if $y, z \in B_r(x) \subset \bar{B}_r(x)$, then $\|f(y) - f(z)\| \leq L\|y - z\|$, as desired. This completes the proof. $\square$

A.2. Proof of Proposition 2.2. Suppose that (x_n) is a sequence on U that converges to x , and (t_n) is a positive monotone sequence that converges to 0. Choose any $r > 0$ and $L > 0$ such that f is L -Lipschitz on $B_r(x)$. Then, there exists N such that if $n \geq N$, then $x_n, x_n + t_n v \in B_r(x)$, and thus,

$$\frac{\|f(x_n + t_n v) - f(x_n)\|}{t_n} \leq \frac{L\|x_n + t_n v - x_n\|}{t_n} = L\|v\|,$$

which implies that $|f^\circ(x; v)| \leq L\|v\|$.

Next, choose any $v \in X$. By the definition of $\limsup$, there exist a sequence (x_n) that converges to x and a positive monotone sequence (t_n) that converges to 0 such that

$$\lim_{n \rightarrow \infty} \frac{f(x_n + t_n v) - f(x_n)}{t_n} = f^\circ(x; v).$$

If $a > 0$, then for $s_n = a^{-1}t_n$,

$$f^\circ(x; av) \geq \limsup_{n \rightarrow \infty} \frac{f(x_n + s_n av) - f(x_n)}{s_n} = a \lim_{n \rightarrow \infty} \frac{f(x_n + t_n v) - f(x_n)}{t_n} = a f^\circ(x; v).$$

By symmetrical arguments, we can show that $a f^\circ(x; v) \geq f^\circ(x; av)$, and thus, we find that $f^\circ(x; av) = a f^\circ(x; v)$. Because it is clear that $f^\circ(x; 0) = 0$, we have shown that

$$f^\circ(x; av) = a f^\circ(x; v)$$

for any $a \geq 0$.

Choose any $v, w \in X$. Then, there exist a sequence (x_n) that converges to x and a positive monotone sequence (t_n) that converges to 0 such that

$$f^\circ(x; v + w) = \lim_{n \rightarrow \infty} \frac{f(x_n + t_n(v + w)) - f(x_n)}{t_n}.$$

Define $y_n = x_n + t_n w$. Then, (y_n) converges to x , and thus,

$$\begin{aligned} f^\circ(x; v + w) &= \lim_{n \rightarrow \infty} \frac{f(x_n + t_n(v + w)) - f(x_n)}{t_n} \\ &= \lim_{n \rightarrow \infty} \left[\frac{f(y_n + t_n v) - f(y_n)}{t_n} + \frac{f(x_n + t_n w) - f(x_n)}{t_n} \right] \\ &\leq \limsup_{n \rightarrow \infty} \frac{f(y_n + t_n v) - f(y_n)}{t_n} + \limsup_{n \rightarrow \infty} \frac{f(x_n + t_n w) - f(x_n)}{t_n} \\ &\leq f^\circ(x; v) + f^\circ(x; w). \end{aligned}$$

This completes the proof. $\square$

A.3. Proof of Proposition 2.3. Choose any $v \in X$. Suppose that (x_n) is a sequence on U that converges to x , (t_n) is a positive monotone sequence that converges to 0, and

$$\lim_{n \rightarrow \infty} \frac{f(x_n + t_n v) - f(x_n)}{t_n} = f^\circ(x; v).$$

Then, for $y_n = x_n + t_n v$, (y_n) converges to x , and

$$f^\circ(x; v) = \lim_{n \rightarrow \infty} \frac{(-f)(y_n - t_n v) - (-f)(y_n)}{t_n} \leq (-f)^\circ(x; -v).$$

Because f and $-f$ can be interchanged, we can easily show that $(-f)^\circ(x; -v) \leq f^\circ(x; v)$. Therefore, we obtain the following equation.

$$(-f)^\circ(x; -v) = f^\circ(x; v).$$

Hence, if $p \in \partial^\circ f(x)$, then

$$(-f)^\circ(x; v) = f^\circ(x; -v) \geq \langle p, -v \rangle = \langle -p, v \rangle,$$

and thus, $-p \in \partial^\circ(-f)(x)$. Conversely, if $-p \in \partial^\circ(-f)(x)$, then

$$f^\circ(x; v) = (-f)^\circ(x; -v) \geq \langle -p, -v \rangle = \langle p, v \rangle,$$

and thus, $p \in \partial^\circ f(x)$. Therefore, our claim is correct. This completes the proof. $\square$

A.4. Proof of Proposition 2.5. Suppose that $p \in \partial f(x)$. By definition,

$$\begin{aligned} f^\circ(x; v) &= \limsup_{y \rightarrow x, t \downarrow 0} \frac{f(y + tv) - f(y)}{t} \\ &\geq \limsup_{t \downarrow 0} \frac{f(x + tv) - f(x)}{t} \geq \lim_{t \downarrow 0} \frac{\langle p, tv \rangle}{t} = \langle p, v \rangle, \end{aligned}$$

and thus $p \in \partial^\circ f(x)$.

Next, suppose that $p \in X'$ and $p \notin \partial f(x)$. Then, there exists $v \in X$ such that $f(x + v) - f(x) < \langle p, v \rangle$, and thus, there exist $a \in \mathbb{R}$ and an open neighborhood V of x such that $y \in V$ implies that $f(y + v) - f(y) < a < \langle p, v \rangle$. If $0 < t \leq 1$, then $y + tv = (1 - t)y + t(y + v)$, and thus

$$\frac{f(y + tv) - f(y)}{t} \leq f(y + v) - f(y) < a,$$

which implies that $f^\circ(x; v) \leq a < \langle p, v \rangle$. Hence, $p \notin \partial^\circ f(x)$, which completes the proof. $\square$

A.5. Proof of Proposition 2.6. First, suppose that f is Gâteaux differentiable at x . Then, for any $v \in X$,

$$\begin{aligned} \langle D_G f(x), v \rangle &= \lim_{t \downarrow 0} \frac{f(x + tv) - f(x)}{t} \\ &\leq \limsup_{y \rightarrow x, t \downarrow 0} \frac{f(y + tv) - f(y)}{t} = f^\circ(x; v), \end{aligned}$$

which implies that $D_G f(x) \in \partial^\circ f(x)$.

Next, suppose that f is strongly differentiable at x . Then, by definition,

$$f^\circ(x; v) = \langle D_s f(x), v \rangle$$

for any $v \in X$. Therefore, $D_s f(x) \in \partial^\circ f(x)$. Suppose that $p \in X'$ and $p \neq D_s f(x)$. Then, there exists $v \in X$ such that $\langle p, v \rangle > \langle D_s f(x), v \rangle$, which implies that $\langle p, v \rangle > f^\circ(x; v)$ and thus $p \notin \partial^\circ f(x)$. Hence, $\partial^\circ f(x) = \{D_s f(x)\}$.

Conversely, suppose that $\partial^\circ f(x) = \{p\}$. First, choose any $v \in X$, and define $V = \{av | a \in \mathbb{R}\}$. For $av \in V$, define $\langle q, av \rangle = af^\circ(x; v)$. If $a \geq 0$, then by Proposition 2.2, $\langle q, av \rangle = f^\circ(x; av)$. If $a < 0$, then again by Proposition 2.2,

$$0 = f^\circ(x; 0) \leq f^\circ(x; av) + f^\circ(x; -av),$$

and thus,

$$\langle q, av \rangle = -\langle q, -av \rangle = -f^\circ(x; -av) \leq f^\circ(x; av).$$

Therefore, by the Hahn-Banach theorem, there exists $r \in \partial^\circ f(x)$ such that the restriction of r to V coincides with q . However, because $\partial^\circ f(x) = \{p\}$, $r = p$, and in particular,

$$f^\circ(x; v) = \langle p, v \rangle.$$

Hence,

$$\begin{aligned}
\liminf_{y \rightarrow x, t \downarrow 0} \frac{f(y+tv) - f(y)}{t} &= -\limsup_{y \rightarrow x, t \downarrow 0} \frac{f(y+tv-tv) - f(y+tv)}{t} \\
&= -f^\circ(x; -v) = -\langle p, -v \rangle = \langle p, v \rangle \\
&= f^\circ(x, v) = \limsup_{y \rightarrow x, t \downarrow 0} \frac{f(y+tv) - f(y)}{t},
\end{aligned}$$

which implies that

$$\lim_{y \rightarrow x, t \downarrow 0} \frac{f(y+tv) - f(y)}{t} = \langle p, v \rangle.$$

Hence, f is strongly differentiable at x and $D_s f(x) = p$.

Finally, suppose that f is strongly differentiable at x , and let $p = D_s f(x)$. Because f is locally Lipschitz, there exist $r > 0$ and $L > 0$ such that f is L -Lipschitz on $B_r(x)$. Choose any compact set $C \subset X$. Fix $\varepsilon > 0$, and choose any $v \in C$. Then, there exists $\delta(v) \in]0, r[$ such that if $\|y - x\| < \delta(v)$ and $0 < t < \delta(v)$, then

$$|f(y+tv) - f(y) - t\langle p, v \rangle| < t \frac{\varepsilon}{L + \|p\| + 1}.$$

Let

$$r' = \min \left\{ \frac{\varepsilon}{L + \|p\| + 1}, 1 \right\}.$$

If $w \in C$ and $\|w - v\| < r'$, then

$$\begin{aligned}
|f(x+tw) - f(x) - t\langle p, w \rangle| &\leq |f((x+t(w-v))+tv) - f(x+t(w-v)) - t\langle p, v \rangle| \\
&\quad + |f(x+t(w-v)) - f(x)| + t|\langle p, w - v \rangle| \\
&< t \frac{\varepsilon}{L + \|p\| + 1} + t(L + \|p\|)\|w - v\| < t\varepsilon.
\end{aligned}$$

Because $\{B_{r'}(v)\}_{v \in C}$ is an open covering of C , there is a finite subcovering $\{B_{r'}(v_i)\}_{i=1}^N$. Let $\delta = \min\{\delta(v_1), \dots, \delta(v_N)\}$. If $0 < t < \delta$, then for any $v \in C$,

$$|f(x+tv) - f(x) - t\langle p, v \rangle| < t\varepsilon.$$

This implies that $p = D_H f(x)$. Clearly, $D_H f(x) = D_G f(x)$, and

$$f^\circ(x; v) = \langle D_s f(x), v \rangle = \langle D_G f(x), v \rangle = f'(x; v),$$

which implies that f is regular at x . This completes the proof. $\square$

A.6. Proof of Proposition 2.7. First, choose any bounded set $B \subset X$, and choose $M > 0$ such that $\|v\| \leq M$ for all $v \in B$. Fix $\varepsilon > 0$, and choose $r > 0$ such that if $y \in B_r(x)$, then

$$\|D_G f(y) - D_G f(x)\| < \frac{\varepsilon}{M}.$$

Let $t > 0$ be so small that $t < \frac{r}{M}$. Then, $x + tv \in B_r(x)$. Because of the mean value theorem, there exists $\theta \in [0, 1]$ such that for all $t > 0$,

$$\frac{f(x+tv) - f(x)}{t} = \langle D_G f(x + \theta tv), v \rangle.$$

Therefore,

$$\begin{aligned} & \left| \frac{f(x+tv) - f(x)}{t} - \langle D_G f(x), v \rangle \right| \\ &= |\langle D_G f(x + \theta tv) - D_G f(x), v \rangle| \\ &\leq \|D_G f(x + \theta tv) - D_G f(x)\| \|v\| < \varepsilon, \end{aligned}$$

as desired. Hence, $D_G f(x) = Df(x)$. Next, choose any $v \in X$. Let (x_n) be a sequence on U that converges to x , and (t_n) be a decreasing positive sequence that converges to 0. Without loss of generality, we can assume that $x_n, x_n + t_n v \in B_r(x)$ for all n . Then, there exists $\theta_n \in [0, 1]$ such that

$$\frac{f(x_n + t_n v) - f(x_n)}{t_n} = \langle D_G f(x_n + \theta_n t_n v), v \rangle \rightarrow \langle D_G f(x), v \rangle,$$

which implies that f is strongly differentiable at x and $D_s f(x) = D_G f(x)$. This completes the proof. $\square$

A.7. Proof of Proposition 2.8. We treat only the case in which f is convex and L -Lipschitz: the remaining cases can easily be proved using the result of this case. Choose any $x \in U$ and $v \in X$. Then, for $t, s > 0$ with $s > t$,

$$-L\|v\| \leq \frac{f(x+tv) - f(x)}{t} \leq \frac{f(x+sv) - f(x)}{s} \leq L\|v\|.$$

Therefore,

$$f'(x; v) = \inf_{t>0} \frac{f(x+tv) - f(x)}{t} \in [-L\|v\|, L\|v\|].$$

Now, choose a sequence (x_n) on U that converges to x and a sequence (t_n) of positive real numbers that converges to 0 such that

$$f^\circ(x; v) = \lim_{n \rightarrow \infty} \frac{f(x_n + t_n v) - f(x_n)}{t_n}.$$

Fix $\delta > 0$, and let $s_n = \max\{t_n, \|x_n - x\|/\delta\}$. Then, $s_n \rightarrow 0$. Because $t \mapsto \frac{f(x_n+tv) - f(x_n)}{t}$ is nondecreasing,

$$\begin{aligned} f^\circ(x; v) &\geq \limsup_{n \rightarrow \infty} \frac{f(x_n + s_n v) - f(x_n)}{s_n} \\ &\geq \liminf_{n \rightarrow \infty} \frac{f(x_n + s_n v) - f(x_n)}{s_n} \\ &\geq \lim_{n \rightarrow \infty} \frac{f(x_n + t_n v) - f(x_n)}{t_n} = f^\circ(x; v), \end{aligned}$$

and thus, we can assume without loss of generality that $t_n = s_n$ and $\|x_n - x\| \leq \delta t_n$. Then,

$$\left| \frac{f(x_n + t_n v) - f(x_n)}{t_n} - \frac{f(x + t_n v) - f(x)}{t_n} \right| \leq 2L\delta,$$

and thus,

$$f^\circ(x; v) \geq f'(x; v) \geq f^\circ(x; v) - 2L\delta.$$

Because $\delta > 0$ is arbitrary, we obtain that $f^\circ(x; v) = f'(x; v)$. This completes the proof. $\square$

A.8. Proof of Proposition 3.6. Because $\limsup_{k \rightarrow +\infty} \frac{F(k)}{k} < 1$, there exists K^* such that if $k \geq K^*$, then $F(k) \leq k$. Without loss of generality, we can assume that $\bar{k} < K^*$, and thus the feasible sequence (k_t) on the model (3.4) must satisfy $k_t \in [0, K^*]$ for all $t \geq 1$.

Let

$$\Pi(k) = \{(k_t) \in [0, K^*]^{\mathbb{N}} | k_0 = k, 0 \leq k_{t+1} \leq F(k_t)\},$$

$$\Pi^* = \cup_{k \in [0, K^*]} \Pi(k).$$

Choose any $k \in [0, K^*]$ and $(k_t) \in \Pi(k)$. By the assumption of K^* , $F(k_t) - k_{t+1} \leq K^*$, and thus, $u(F(k_t) - k_{t+1}) \leq u(F(K^*))$. This implies that the following function

$$U((k_t)) = \sum_{t=0}^{\infty} \delta^t u(F(k_t) - k_{t+1})$$

is well-defined and bounded from above. By Tychonoff's theorem, $[0, K^*]^{\mathbb{N}}$ is compact with respect to the product topology, and because $\Pi(k)$ is a closed subset of this set, it is also compact. We first show that U is actually continuous on Π^* .

Choose a convergent sequence (k_t^n) on Π^* and $\lim_{n \rightarrow \infty} k_t^n = k_t^*$ for each t . Because $0 \leq k_t^n \leq K^*$ and $0 \leq k_{t+1}^n \leq F(k_t^n)$ for all n , we have that $0 \leq k_t^* \leq K^*$ and $0 \leq k_{t+1}^* \leq F(k_t^*)$, and thus $(k_t^*) \in \Pi(k_0^*) \subset \Pi^*$. First, suppose that $U((k_t^*)) = -\infty$. Choose any $a \in \mathbb{R}$. Then, there exists $T > 0$ such that

$$\sum_{t=0}^T \delta^t (u(F(k_t^*) - k_{t+1}^*) - u(F(K^*))) + \frac{u(F(K^*))}{1 - \delta} < a.$$

This implies that, there exists N such that if $n \geq N$, then

$$\sum_{t=0}^T \delta^t (u(F(k_t^n) - k_{t+1}^n) - u(F(K^*))) + \frac{u(F(K^*))}{1 - \delta} < a.$$

Therefore,

$$\sum_{t=0}^{\infty} \delta^t u(F(k_t^n) - k_{t+1}^n) \leq \sum_{t=0}^T \delta^t (u(F(k_t^n) - k_{t+1}^n) - u(F(K^*))) + \sum_{t=0}^{\infty} \delta^t u(F(K^*)) < a,$$

which implies that $\limsup_{n \rightarrow \infty} U((k_t^n)) = -\infty$, and in this case, $\lim_{n \rightarrow \infty} U((k_t^n)) = U((k_t^*))$. Hence, we can assume that $U((k_t^*)) > -\infty$. In particular, $u(F(k_t^*) - k_{t+1}^*) > -\infty$ for all $t \geq 0$.

Next, choose any $\varepsilon > 0$. There exists N_t such that if $n \geq N_t$, then

$$|u(F(k_t^n) - k_{t+1}^n) - u(F(k_t^*) - k_{t+1}^*)| < (1 - \delta)\varepsilon.$$

Define $M_1 = N_1$, and if M_T is defined, then define $M_{T+1} = \max\{N_{T+1}, M_T + 1\}$. Then, (M_T) is an increasing sequence. If $n \geq M_T$,

$$\begin{aligned} \sum_{t=0}^{\infty} \delta^t u(F(k_t^n) - k_{t+1}^n) &= \sum_{t=0}^T \delta^t u(F(k_t^n) - k_{t+1}^n) + \sum_{t=T+1}^{\infty} \delta^t u(F(k_t^n) - k_{t+1}^n) \\ &< \sum_{t=0}^T \delta^t u(F(k_t^*) - k_{t+1}^*) + (1 - \delta^{T+1})\varepsilon + \frac{\delta^{T+1}}{1 - \delta} u(F(K^*)) \\ &\rightarrow U((k_t^*)) + \varepsilon, \end{aligned}$$

and

$$\begin{aligned}
\sum_{t=0}^{\infty} \delta^t u(F(k_t^*) - k_{t+1}^*) &= \sum_{t=0}^T \delta^t u(F(k_t^*) - k_{t+1}^*) + \sum_{t=T+1}^{\infty} \delta^t u(F(k_t^*) - k_{t+1}^*) \\
&< \sum_{t=0}^T \delta^t u(F(k_t^n) - k_{t+1}^n) + (1 - \delta^{T+1})\varepsilon + \frac{\delta^{T+1}}{1 - \delta} u(F(K^*)) \\
&\rightarrow U((k_t^n)) + \varepsilon.
\end{aligned}$$

Therefore, we have $|U((k_t^n)) - U((k_t^*))| \leq \varepsilon$ for sufficiently large n , as desired.

Hence, U is continuous on Π^* . Next, we will show that the set-valued mapping Π is continuous. Choose $k \in [0, K^*]$ and let (k^n) be a sequence on $[0, K^*]$ that converges to k . First, suppose that there exist an open neighborhood U of $\Pi(k)$ and a sequence $((k_t^n))$ on Π^* such that $(k_t^n) \in \Pi(k^n)$ for each n , and for all N , there exists $n \geq N$ such that $(k_t^n) \notin U$. Taking a subsequence, we can assume that $(k_t^n) \notin U$ for all n . Because (k_t^n) is a sequence of $[0, K^*]^{\mathbb{N}}$, we can assume without loss of generality that $\lim_{n \rightarrow \infty} k_t^n = k_t^*$ for each $t \geq 0$. Using the same arguments as in the previous paragraphs, we can show that $(k_t^*) \in \Pi(k) \subset U$, which contradicts our initial assumption. Therefore, we have that Π is upper hemi-continuous. Second, suppose that U is an open set in Π^* such that there exists $(k_t^*) \in U \cap \Pi(k)$. Define $k_0^n = k^n$, and if k_t^n is defined, then define $k_{t+1}^n = \min\{k_{t+1}^*, F(k_t^n)\}$. Because F is continuous, $\lim_{n \rightarrow \infty} k_1^n = k_1^*$, and if $\lim_{n \rightarrow \infty} k_t^n = k_t^*$, then $\lim_{n \rightarrow \infty} k_{t+1}^n = k_{t+1}^*$. Hence, $(k_t^n) \in \Pi(k^n)$ and $\lim_{n \rightarrow \infty} (k_t^n) = (k_t^*)$, which implies that $(k_t^n) \in \Pi(k^n) \cap U$ for sufficiently large n . Therefore, Π is lower hemi-continuous, as desired.

Therefore, Π is a compact-valued continuous mapping on $[0, K^*]$, and by Berge's maximum theorem, the solution mapping to (3.4) is nonempty-valued, and the value function $\bar{V}$ is continuous on $[0, K^*]$. Let (k_t^*) be the solution to (3.4) with $k_0^* = \bar{k}$. Because $k_t^* \leq K^*$, we have that

$$\bar{V}(\bar{k}) = \sum_{t=0}^{\infty} \delta^t u(F(k_t^*) - k_{t+1}^*) \leq \frac{1}{1 - \delta} u(F(K^*)) < +\infty.$$

Suppose that there exists ε^* in the claim of this proposition. Let $k^+ = \min\{\varepsilon^*, \bar{k}\}$. Then,

$$\bar{V}(\bar{k}) \geq u(F(\bar{k}) - k^+) + \sum_{t=1}^{\infty} \delta^t u(F(k^+) - k^+) > -\infty.$$

Therefore, $\bar{V}(\bar{k}) \in \mathbb{R}$. This completes the proof. $\square$

Acknowledgements

The author is grateful to anonymous referees for their helpful suggestions.

REFERENCES

- [1] L.M. Benveniste, J.A. Scheinkman, On the differentiability of the value function in dynamic models of economics, *Econometrica* 47 (1979), 727-732.
- [2] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
- [3] J. Heinonen, *Lectures on Lipschitz Analysis*, Lectures at the 14th Jyväskylä Summer School, 2004.
- [4] B.S. Mordukhovich, N. Sagara, Subdifferentials of nonconvex integral functionals in Banach spaces with applications to stochastic dynamic programming, *J. Convex Anal.* 25 (2018), 643-673.
- [5] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1996.

- [6] N. Sagara, Cores and Weber sets for fuzzy extensions of cooperative games, *Fuzzy Sets and Systems* 272 (2015), 102-114.
- [7] N. Stokey, R. Lucas, t Recursive Methods in Economic Dynamics, Harvard University Press, Massachusetts, 1989.