Supplementary Appendix to Envelope theorem in dynamic economic models with recursive utility[☆]

Juan Pablo Rincón-Zapatero[a], Yanyun Zhao[b,*]

[a] *Department of Economics, Universidad Carlos III de Madrid, 28903 Getafe (Madrid), Spain*

[b] *Wenlan School of Business, Zhongnan University of Economics and Law, 430073 Wuhan, China*

1. Nonsmooth chain rule and Inada condition

In this section, we firstly introduce some preliminary notations and auxiliary lemmas. Then the proof of the main result will be presented later.

1.1. Examples of aggregators

There are many examples of aggregators studied in the literature. We consider the following three examples:

Time additively separable aggregator: consider the aggregator $\mathrm{W}(x, y, z) = u(x, y) + \beta z$, where $u : \mathrm{X} \times \mathrm{X} \to \mathbb{R}$ is a utility function and $\beta \in (0, 1)$ is the discount factor. Under suitable assumptions, $\mathrm{U}(\boldsymbol{x}) = \sum_{t=0}^{\infty} \beta^t u(x_t, x_{t+1})$ is induced from this aggregator. More details could be found in Stokey, Lucas and Prescott (1989, Ch.5).

Epstein–Heynes aggregator: consider the aggregator $\mathrm{W}(x, y, z) = (-1 + z)e^{-v(x,y)}$ where $v : \mathrm{X} \times \mathrm{X} \to \mathbb{R}$ is a continuous and differentiable function satisfying $v > 0, \partial v/\partial x > 0, \partial v/\partial y > 0$. Under suitable assumptions, this aggregator yields the utility function as follows:

$$U(\boldsymbol{x}) = -\sum_{t=1}^{\infty} \exp\left[-\sum_{\tau=1}^{t} v(x_{\tau-1}, x_\tau)\right]$$

Koopmans-Diamond-Williamson aggregator: the Koopmans-Diamond-Williamson aggregator (hereafter the KDW aggregator) is given by $\mathrm{W}(x, y, z) = (\beta/d) \ln(1 + a(f(x) - y)^b + dz)$ where $a, b, d, \beta > 0$ with $b, \beta < 1$ and with $f : \mathbb{R}_+ \to \mathbb{R}_+$, $0 \le y \le f(x)$. In fact, there is no closed–form expression from the KDW aggregator. However, Becker and Boyd

[☆] Support from the Ministerio de Economía y Competitividad (Spain), grants ECO 2014-56384-P, MDM 2014-0431, and Comunidad de Madrid, MadEco-CM S2015/HUM-3444 is gratefully acknowledged. The second author is grateful for university starting fund. We thank participants at the XXVI European Workshop on General Equilibrium Theory in Salamanca for useful comments.

[*] Corresponding Author. Email: yanyunzhao@zuel.edu.cn.

(1997, Ch.3) show that there is a unique recursive utility function for the KDW aggregator. This result is extended in Marinacci and Montrucchio (2010) to cover cases where $\beta \geq 1$. In fact, a whole class of upcounting aggregators are identified in Marinacci and Montrucchio (2010) that admit a unique recursive utility function, and they are the so-called Thompson aggregators.

1.2. Nonsmooth chain rule

For a concave function $F: \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$, the effective domain is the set dom $F = \{x \in \mathbb{R}^n : F(x) > -\infty\}$. We say F is proper if dom F is nonempty. As usual, if F is given in a convex subset $C \subseteq \mathbb{R}^n$, we extend F to $\mathbb{R}^n$ by defining $F = -\infty$ on the complement of C and then the extended function is concave. The set $\partial F(x_0)$ defined by

$$\partial F(x_0) := \{q \in \mathbb{R}^n : F(x) - F(x_0) \leq q^T(x - x_0), \forall x \in \mathbb{R}^n\}$$

is the superdifferential of F at x_0. In our finite dimensional setting, the superdifferential is nonempty at interior points of the domain of F. The elements q of $\partial F(x_0)$ are supergradients. When F is differentiable, the superdifferential reduces to a singleton, which single element is the gradient of F at x_0: $\partial F(x_0) = \{\nabla F(x_0)\}$ when $\nabla F(x_0)$ exists. Reciprocally, if F is continuous in the interior of its domain and $\partial F(x_0)$ contains only one point (which is the gradient of F), then F is differentiable at x_0.

Let $f = (\ell, f_1, \ldots, f_d)$, where $\ell: \mathbb{R}^m \to \mathbb{R}^k$ is linear and each $f_i: \mathbb{R}^m \to \mathbb{R} \cup \{-\infty\}$ is concave and let $F: \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ be concave. We will consider the composition $F \circ f$. It is well-known that $F \circ f$ is not concave unless further conditions are imposed. Let the usual ordering on $\mathbb{R}^d$, that is, for $s = (s_1, \ldots, s_d)$ and $s' = (s'_1, \ldots, s'_d)$ elements of $\mathbb{R}^d$, we say $s \leq s'$ if $s_i \leq s'_i$ for each i. The function F is said to be isotone on $S \subseteq \mathbb{R}^d$ if for any $r \in \mathbb{R}^k$, $F((r, s)) \leq F((r, s'))$ whenever $s, s' \in S$ and $s \leq s'$. Let the correspondence

$$M(f_1, \ldots, f_d)(a) := \{s \in \mathbb{R}^d : (f_1, \ldots, f_d)(a) \leq s\}$$

and its image, $R(M(f_1, \ldots, f_d)) := \bigcup_{a \in \mathbb{R}^m} M(f_1, \ldots, f_d)(a)$.

Theorem 1.1. *Let $f = (\ell, f_1, \ldots, f_d)$, where $\ell: \mathbb{R}^m \to \mathbb{R}^k$ is linear and each $f_i: \mathbb{R}^m \to \mathbb{R} \cup \{-\infty\}$ is concave and continuous on the interior of its domain; let $F: \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ be concave, and isotone on $R(M(f_1, \ldots, f_d))$. Then $F \circ f$ is concave and if*

$$\text{dom } \ell \cap \bigcap_{i=1}^{d} \text{int}(\text{dom } f_i) \neq \emptyset,$$

then for any a with $f(a)$ finite, the superdifferential of $F \circ f$ is given by

$$\partial(F \circ f)(a) = \left\{\ell^*(\alpha_1, \ldots, \alpha_k)^T\right\} + \alpha_{k+1}\partial f_1(a) + \cdots + \alpha_{k+d}\partial f_d(a), \tag{1}$$

where $(\alpha_1, \ldots, \alpha_{k+d}) \in \partial F(a)$ and ℓ^* is the adjoint matrix of ℓ.

Proof. Let $f^1 := (f_1, \ldots, f_d)$. Let $a_1, a_2 \in \mathbb{R}^d$ and let $\lambda \in [0,1]$; let, to simplify notation, $a^\lambda = \lambda a_1 + (1-\lambda)a_2$. We have

$$\begin{aligned}
(F \circ f)(a^\lambda) &= F((\ell(a^\lambda), f^1(a^\lambda))) \\
&= F((\lambda\ell(a_1) + (1-\lambda)\ell(a_2), f^1(a^\lambda))) \\
&\geq F((\lambda\ell(a_1) + (1-\lambda)\ell(a_2), \lambda f^1(a_1) + (1-\lambda)f^1(a_2))) \\
&= F((\lambda(\ell(a_1), f^1(a_1)) + (1-\lambda)(\ell(a_2), f^1(a_2))) \\
&\geq \lambda F((\ell(a_1), f^1(a_1)) + (1-\lambda)F((\ell(a_2), f^1(a_2))) \\
&= \lambda(F \circ f)(a_1) + (1-\lambda)(F \circ f)(a_2).
\end{aligned}$$

Thus, $F \circ f$ is concave. Now, if $f(a)$ is finite and F, f and $F \circ f$ are concave, then the chain rule for concave functions, states (see e.g. Ward and Borwein (1987); isotonicity of F with respect to all variables plays no role in establishing this formula)

$$\partial(F \circ f)(a) = \alpha_1\partial\ell_1(a) + \ldots + \alpha_k\partial\ell_k(a) + \alpha_{k+1}\partial f_1(a) + \cdots + \alpha_{k+d}\partial f_d(a),$$

where $(\alpha_1, \ldots, \alpha_{k+d}) \in \partial F(f(a))$ and $\ell = (\ell_1, \ldots, \ell_k)$. Since each ℓ_i is a linear mapping, we have $\partial\ell_i(a) = \{\ell_i\}$, where we use the same notation to designate the mapping ℓ_i and its representation as matrix. Hence, we get (1). □

Proposition 1. *Consider a dynamic optimization problem* $(\mathrm{X}, \Gamma, \mathrm{W})$. *Then, for any concave function $v : X \longrightarrow \mathbb{R} \cup \{-\infty\}$, the function $V(x,y) := W(x,y,v(y))$ is concave and for any (x_0, y_0) with $v(y_0)$ finite, the superdifferential of V is given by*

$$\partial V(x_0, y_0) = \{(\alpha_1, \alpha_2 + \alpha_3 q) \,:\, (\alpha_1, \alpha_2, \alpha_3) \in \partial W(x_0, y_0, v(y_0)),\ q \in \partial v(y_0)\}\,. \tag{2}$$

Proof. Letting $f = (\ell, v)$, where ℓ is the identity mapping of $\mathbb{R}^n \times \mathbb{R}^n$ and where v is concave by assumption, we have that $w = W \circ f$ is concave by Theorem 1.1, since W is increasing with respect to the third component. The expression (2) comes from (1), since the superdifferential of the function $(x,y) \to v(y)$ is $\{0\} \times \partial v(y)$. □

Of course, when W is differentiable, we can plug $\alpha_i = D_i W(x_0, y_0, v(y_0))$, $i = 1,2,3$ into (2).

Consider the convex set $\mathrm{A} = \{(x, y) : \; y \in \Gamma(x), \; x \in \mathbb{R}^n\}$. If $(x_0, y_0) \in A$, the normal cone to A at (x_0, y_0) is defined as

$$N_{\mathrm{A}}(x_0, y_0) := \{(\xi_1, \xi_2) \in \mathbb{R}^{2n} : \; (\xi_1, \xi_2) \cdot (x - x_0, y - y_0) \leq 0, \text{ for all } (x, y) \in \mathrm{A}\}. \quad (3)$$

The indicator function of A is the convex function defined by

$$\delta_{\mathrm{A}}(x, y) := \begin{cases} 0, & (x, y) \in \mathrm{A} \\ \infty, & (x, y) \notin \mathrm{A}. \end{cases}$$

It is well-known that $\partial \delta_{\mathrm{A}}(x, y) = N_{\mathrm{A}}(x, y)$.

Theorem 1.2. *Consider a dynamic optimization problem* $(\mathrm{X}, \Gamma, \mathrm{W})$. *Assume that the value function* $\mathcal{J}$ *is concave and finite on* X. *Then, for any* $(x_0, y_0) \in \mathrm{A}$, *the superdifferential of* $\mathcal{J}$ *at* x_0 *is characterized as follows.*

$$\partial \mathcal{J}(x_0) = \Big\{ q_0 \in \mathbb{R}^n \; : \; \exists (\alpha_1, \alpha_2, \alpha_3) \in \partial W(x_0, y_0, \mathcal{J}(x_0)), \; \exists (p_1, p_2) \in N_{\mathrm{A}}(x_0, y_0), \\ \exists q \in \partial v(y_0) \textit{ such that } q_0 = \alpha_1 - p_1, \; \alpha_2 + \alpha_3 q = p_2 \Big\}. \quad (4)$$

Proof. Let $V(x, y) = W(x, y, \mathcal{J}(y))$. By Proposition 1, w is concave and its superdifferential is given in (2). The Bellman equation can be written

$$\mathcal{J}(x) = \max_{y \in \mathbb{R}^n} \Big\{ V(x, y) - \delta_{\mathrm{A}}(x, y) \Big\}.$$

Let $x_0 \in X$ and let y_0 be a maximizing argument in the Bellman equation and that we assume that exists. Then, by Proposition 4.3 in Aubin (1993), $q_0 \in \partial \mathcal{J}(x_0)$ if and only if $(q_0, 0) \in \partial \left(V - \delta_{\mathrm{A}} \right)(x_0, y_0)$. The theorem follows by observing that $\partial \left(V - \delta_{\mathrm{A}} \right)(x_0, y_0) = \partial V(x_0, y_0) - N_{\mathrm{A}}(x_0, y_0)$ and substituting the expression for $\partial V(x_0, y_0)$ from Proposition 1. □

1.3. Inada condition

Let $\{x_{t+1}\}_{t=0}^{\infty}$ be an optimal path starting at $x_0 \in X$. We conclude from Theorem 1.2 that for any $t \geq 0$

$$\partial \mathcal{J}(x_t) \neq \emptyset \;\text{ if and only if }\; \Big(\partial W(x_t, x_{t+1}, \mathcal{J}(x_{t+1})) \neq \emptyset \quad \text{and} \quad \partial \mathcal{J}(x_{t+1}) \neq \emptyset \Big). \quad (5)$$

This means that if the superdifferential of the value function is not empty at x_t, then it is not empty along an optimal path at time $t+1, t+2, \ldots$; also, in this case an optimal path never visits regions of A where the superdifferential of W is empty. This allows us to establish a generalization of one of the Inada conditions for dynamic problems with recursive utility.

Corollary 1.1. *Consider a dynamic optimization problem* $(\mathrm{X}, \Gamma, \mathrm{W})$*. Assume that the value function* $\mathcal{J}$ *is concave and finite on* X *and let* $x_0 \in X$*,* $y_0 \in \mathcal{H}(x_0)$ *such that* $\partial W(x_0, y_0, \mathcal{J}(y_0)) = \emptyset$*. Then* x_0 *is at the boundary of* X*.*

Proof. By (5), $\partial W(x_0, y_0, \mathcal{J}(y_0)) = \emptyset$ implies $\partial \mathcal{J}(x_0) = \emptyset$; since for any concave function its superdifferential is not empty in the interior of its domain and dom $\mathcal{J} = X$, we conclude that x_0 is a boundary point of X. $\square$

Let the dynamic problem $(\mathrm{X}, \Gamma, \mathrm{W})$ be given as follows. The state space is $X = \mathbb{R}^n_+$, the technological correspondence $\Gamma(x) = \{y \in \mathbb{R}^n : 0 \leq y \leq f(x)\}$ where $f = (f_1, \ldots, f_n)$ and each f_i is concave and nondecreasing, with dom$f_i = X$, and an aggregator $W(x, y, z) = w(f(x) - y, z)$, where the function w is concave in (c, z) and increasing in the variable z. Then W is concave by Theorem 1.1.

Let $\frac{\partial w}{\partial h}$ denote the directional derivative of w in direction h in the sense of convex analysis. Also, let $c = f(x) - y$.

We have the following result.

Corollary 1.2. *Consider a dynamic optimization problem* $(\mathrm{X}, \Gamma, \mathrm{W})$ *as described in Corollary 1.1 above. Assume that the value function* $\mathcal{J}$ *is concave and finite on* X *and suppose*

$$\frac{\partial w}{\partial h}(c_0, z) = -\infty \tag{6}$$

for $c_0 = (c^0_1, \ldots, c^0_n)$ *and all* $z \in \mathbb{R}$*, where* $c^0_i = 0$ *for some* i *and some direction* h*. Then for any* $x_0 > 0$*, an optimal* $y_0 \in \mathcal{H}(x_0)$ *has* i*-th component* $y^0_i \neq f_i(x_0)$*.*

Proof. Suppose, by way of contradiction, that $y^0_i = f_i(x_0)$, i.e. $c^0_i = 0$. Condition (6) means $\partial w(c_0, z) = \emptyset$, thus $\partial W(x_0, f(x_0), z) = \emptyset$ for all z. By Corollary 1.1, x_0 is at the boundary of $\mathbb{R}^n_+$, contradicting that $x_0 > 0$. $\square$

To illustrate Corollary 1.2, consider the KDW aggregator described in Example 1.1, where $\mathrm{W}(x, y, z) = (\beta/d) \ln(1 + a(f(x) - y)^b + dz)$, $a, b, d, \beta > 0$, $b, \beta < 1$ and with $f : \mathbb{R}_+ \to \mathbb{R}_+$, concave and nondecreasing, and $0 \leq y \leq f(x)$. Note that for this problem $w(c, z) = (\beta/d) \ln(1 + ac^b + dz)$. It is easy to see that the Inada condition on the production function $f'(0^+) = \infty$ implies (6). Hence an optimal y_0 from $x_0 > 0$ satisfies $0 \leq y_0 < f(x_0)$, that is, the optimal consumption is positive, $c_0 > 0$.

Now we are in a position to prove theorem 3.1 in the main context.

Proof. Let $x_0 \in \text{int}(X)$. Then the superdifferential $\mathcal{J}(x_0)$ is nonempty as described in Theorem 1.2. Under assumption (D3), the normal cone to A at (x_0, y_0) with $x_0 \in \text{int}(X)$ is given by

$$
\begin{aligned}
-N_{\mathrm{A}}(x_0, y_0) = \Big\{(p_1, p_2) \in \mathbb{R}^{2n} \,:\, (p_1, p_2) = \sum_{i \in s(x_0, y_0)} \lambda^i (D_x g^i(x_0, y_0), D_y g^i(x_0, y_0)), \\
\lambda^i \geq 0,\ \forall i \in s(x_0, y_0)\Big\}.
\end{aligned} \tag{7}
$$

Let $(x_{t+1})_{t=0}^{\infty}$ be an optimal path from x_0. Then, using (4) and (7), we have that for any $q_0 \in \partial\mathcal{J}(x_0)$, there exist $\lambda^i \geq 0$ for some $i \in s(x_0, x_1)$, and $q_1 \in \partial\mathcal{J}(x_1)$, such that

$$
\begin{aligned}
q_0 &= D_x \mathrm{W}(x_0, x_1, \mathcal{J}(x_1)) + \sum_{i \in s(x_0, x_1)} \lambda^i D_x g^i(x_0, x_1) \\
-D_y \mathrm{W}(x_0, x_1, \mathcal{J}(x_1)) &= D_z \mathrm{W}(x_0, x_1, \mathcal{J}(x_1)) q_1 + \sum_{i \in s(x_0, x_1)} \lambda^i D_y g^i(x_0, x_1).
\end{aligned}
$$

Following identical steps as in Lemma 5.1 in Rincón-Zapatero and Santos (2009), but with the obvious adaptations to our case, we get

$$
\begin{aligned}
q_0 &= D_x \mathrm{W}(x_0, x_1, \mathcal{J}(x_1)) \\
&\quad - D_x g_s(x_0, x_1)^{\top} D_y g_s^{+}(x_0, x_1) \left\{D_y \mathrm{W}(x_0, x_1, \mathcal{J}(x_1)) + D_z \mathrm{W}(x_0, x_1, \mathcal{J}(x_1)) q_1\right\}.
\end{aligned} \tag{8}
$$

For $t = 1, 2, \ldots$, we define the following condensed notations:

$$
\begin{aligned}
\beta_t &= \prod_{i=1}^{t} D_z \mathrm{W}(x_{i-1}, x_i, \mathcal{J}(x_i)) \\
\mathcal{G}_t &= \prod_{i=1}^{t} G(x_{i-1}, x_i)
\end{aligned}
$$

By simple iterations of (8) from $t = 1$ to $t = T > 1$, it follows that $q_0 \in \partial\mathcal{J}(x_0)$ if and only if there exists $q_T \in \partial\mathcal{J}(x_T)$ such that

$$
q_0 = \sum_{t=0}^{T-1} \beta_t \mathcal{G}_t \{D_x \mathrm{W}(x_t, x_{t+1}, \mathcal{J}(x_{t+1})) + G(x_t, x_{t+1}) D_y \mathrm{W}(x_t, x_{t+1}, \mathcal{J}(x_{t+1}))\} + \beta_T \mathcal{G}_T q_T.
$$

□

References

Aubin, J., 1993. Optima and Equilibria: An Introduction to Nonlinear Analysis. Springer-Verlag.

Becker, R., Boyd, J., 1997. Capital Theory, Equilibrium Analysis and Recursive Utility. Blackwell.

Marinacci, M., Montrucchio, L., 2010. Unique solutions for stochastic recursive utilities. Journal of Economic Theory 145, 1776–1804.

Rincón-Zapatero, J., Santos, M., 2009. Differentiability of the value function without interiority assumptions. Journal of Economic Theory 144, 1948–1964.

Stokey, N., Lucas, R., Prescott, E., 1989. Recursive Methods in Economic Dynamics. Harvard University Press.

Ward, D.E., Borwein, J.M., 1987. Nonsmooth calculus in finite dimensions. SIAM Journal on Control and Optimization 25, 1312–1340.