ONLINE APPENDIX: MARKOV EQUILIBRIA WITH QUASI-HYPERBOLIC DISCOUNTING

ZHIGANG FENG
Department of Economics, University of Nebraska-Omaha

MANUEL S. SANTOS
Department of Economics, University of Miami

Appendix A Some Definitions

Lipschitz Functions A function $f : D \to \mathbb{R}$ with $D \subseteq \mathbb{R}^n$ is called **Lipschitz continuous** (or simply Lipschitz) on D if there exists a real constant $L \geq 0$, known as the **Lipschitz constant**, such that for all $x_1, x_2 \in D$,

$$|f(x_1) - f(x_2)| \leq L\|x_1 - x_2\|.$$

Geometrically, this means that the slope of the line segment connecting any two points on the graph of the function is bounded in absolute value by L. A fundamental result, **Rademacher's theorem**, states that any Lipschitz function defined on an open subset of $\mathbb{R}^n$ is differentiable almost everywhere.

The Generalized Gradient For locally Lipschitz functions, which may not be differentiable everywhere, the concept of the derivative can be extended to the **generalized gradient**, also known as the **Clarke subdifferential**. The formal definition is based on the **generalized directional derivative**. For a locally Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ at a point x_0, the generalized directional derivative in the direction h is:

$$f^\circ(x_0; h) = \limsup_{y \to x_0, t \to 0^+} \frac{f(y + th) - f(y)}{t}. \tag{49}$$

The generalized gradient, denoted $\partial f(x_0)$, is then the set of vectors $\zeta \in \mathbb{R}^n$ such that $\langle \zeta, h \rangle \leq f^\circ(x_0; h)$ for all $h \in \mathbb{R}^n$. A more intuitive characterization of $\partial f(x_0)$ is in terms of the traditional gradient $\nabla f(x)$. Then, $\partial f(x_0)$ is the convex hull of the set of all limit points $\nabla f(x_i)$ for all sequences $x_i \to x_0$ where $\nabla f(x_i)$ exists:

$$\partial f(x_0) = \mathrm{co}\left\{ \lim_{i \to \infty} \nabla f(x_i) \mid x_i \to x_0, \nabla f(x_i) \text{ exists} \right\}. \tag{50}$$

For example, for $f(x) = |x|$ at $x_0 = 0$, the derivatives near the origin are either 1 or -1. The set of limit points is $\{-1, 1\}$, and its convex hull is the interval $[-1, 1]$. Thus, $\partial f(0) = [-1, 1]$. If f is continuously differentiable at x_0, the generalized gradient reduces to a single point, $\partial f(x_0) = \{\nabla f(x_0)\}$.

Proposition A.1. *(Clarke (1975), Proposition 1.13).*
The following are equivalent:

(a) $\partial f(x) = \{\zeta\}$, *a singleton.*

(b) $\nabla f(x)$ *exists,* $\nabla f(x) = \zeta$, *and* ∇f *is continuous at* x *relative to the set upon which it exists.*

The Dini Derivative Dini derivatives are a generalization of the standard derivative used to study functions with limited differentiability, such as continuous but non-differentiable functions; e.g. (Saks, 1937). Their primary advantage is that they always exist for any function (though they may take on the values ±∞). Dini derivatives are fundamental in real analysis and optimization for characterizing the local behavior and optimality conditions of non-smooth functions.

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. The four Dini derivatives at a point x_0 in its domain are defined using one-sided limits:

- The **upper right Dini derivative**:

$$D^+ f(x_0) = \limsup_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h}.$$

- The **lower right Dini derivative**:

$$D_+ f(x_0) = \liminf_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h}.$$

- The **upper left Dini derivative**:

$$D^- f(x_0) = \limsup_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h}.$$

- The **lower left Dini derivative**:

$$D_- f(x_0) = \liminf_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h}.$$

A function is differentiable at x_0 if and only if all four Dini derivatives exist, are finite, and are equal to each other. A powerful result that describes the relationship between these derivatives for any function is the *Denjoy-Young-Saks* theorem.

Theorem A.2. *(The Denjoy-Young-Saks Theorem). Let $f : \mathbb{R} \to \mathbb{R}$ be an arbitrary function. Then, at almost every point x in its domain, exactly one of the following four possibilities holds:*

1. *f is differentiable at x (i.e., all four Dini derivatives are finite and equal).*
2. *$D^+ f(x) = D^- f(x) = \infty$ and $D_+ f(x) = D_- f(x) = -\infty$.*
3. *$D^+ f(x) = \infty$, $D_- f(x) = -\infty$, and $D_+ f(x)$ and $D^- f(x)$ are finite and equal.*
4. *$D^- f(x) = \infty$, $D_+ f(x) = -\infty$, and $D^+ f(x)$ and $D_- f(x)$ are finite and equal.*

Two other results on Dini derivatives are of interest for this paper.

Theorem A.3. *(Giorgi and Komlósi (1992), Theorem 1.16). Let $f(x)$ be defined and continuous on (a, b). If some of its Dini derivatives is bounded (above and below) on (a, b) then $f(x)$ is Lipschitz continuous on (a, b).*

Theorem A.4. *(Giorgi and Komlósi (1992), Theorem 1.17). Let $f(x)$ be defined on (a, b) and let $x_0 \in (a, b)$. If $f(x)$ attains a local maximum at x_0 then $D^+ f(x_0) \leq 0$ and $D_- f(x_0) \geq 0$.*

Appendix B Danskin's Envelope Theorem

Danskin (1967) proves a general envelope theorem, which applies to our framework under interiority of equilibrium solutions [e.g., Rincon-Zapatero and Santos (2009)]. We shall follow

Bernhard and Rapaport (1995) and Clarke (1975), and present a simplified variant of this classical result.

Let $\mathbb{U}$ be a compact topological space, and g a map from $\mathbb{R}^n \times \mathbb{U}$ into $\mathbb{R}$, assumed to be jointly continuous and C^1 with respect to the first variable. Let

$$J(x) = \max_{u \in \mathbb{U}} g(x, u)$$

and

$$M(x) = \{u \in \mathbb{U} | g(x, u) = J(x)\} .$$

Theorem B.1. *(The Envelope Theorem). Function J has a directional derivative at x in the direction h,*

$$DJ(x; h) = \max_{u \in M(x)} \sum_{i=1}^{n} h_i \cdot D_i g(x, u) \tag{51}$$

for every x and h in $\mathbb{R}^n$, where $D_i g$ stands for the partial derivative with respect to component x_i of x, and u is fixed.

A vast literature has extended Danskin's theorem in several dimensions: weaker continuity and differentiability assumptions, constrained optimization, and infinite-dimensional spaces. Here, we present an extension of these classical results under weaker continuity conditions.

Let $\partial J(x)$ stand for the generalized gradient in the sense of Clarke (1975), and $D_x J(x, u)$ stand for the derivative of mapping $J(\cdot, u)$ with respect to component x for fixed u.

Theorem B.2. *(Clarke (1975), Theorem 2.1) Let $\mathbb{U}$ be a sequentially compact space, and let $g : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}$ have the following properties:*

1. *$g(x, u)$ is upper semicontinuous in (x, u).*
2. *g is locally Lipschitz in x, uniformly for u in U.*
3. *$g_x^\circ(x, u; \cdot) = D_x g(x, u; \cdot)$, the derivatives being with respect to x.*
4. *$\partial_x g(x, u)$ is upper semicontinuous in (x, u).*

Then, if we let $J(x) = \max\{g(x, u) : u \in \mathbb{U}\}$,

1. *J is locally Lipschitz.*
2. *$D_x J(x; h)$ exists.*

3. *$D_x J(x;h) = J^\circ(x;h) = \max\{\zeta \cdot h : \zeta \in \partial_x g(x,u), u \in M(x)\}$, where $M(x) = \{u \in \mathbb{U} : g(x,u) = J(x)\}$.*

4. *$\partial J(x)$ is the convex hull of $\{\partial_x g(x,u) : u \in M(x)\}$.*

Lemma B.3 (Addendum to Proof of Theorem 2.12)**.** *Let $k^* = g(k^*)$ be an interior steady state. Then, under the above Assumptions for the Economic Growth Model (Section 2) the right-hand derivative $V'_+(k^*)$ of $V(\cdot)$ at k^* is well defined, and the right-hand derivative $g'_+(k^*)$ of $g(\cdot)$ at k^* is also well defined.*

Proof. The proof is divided into three parts, and builds on Theorems B.1 and B.2 of the Online Appendix (and consequently Theorem 3.1 of the paper). Furthermore, Bellman's equation, $\beta V(k) = W(k) - (1-\beta)u(f(k) - g(k))$, entails that the derivative of V is bounded if and only if the derivative of g is bounded, for all k where such derivatives exist. This is quite relevant at a steady state solution k^*, where $k^* = g(k^*)$ and g maps points of an interval into itself.

Part (1). The current value function W has a well-defined right-hand derivative $W'_+(k^*)$. This is just a consequence of Theorem B.1. Observe that by Theorem B.2, function W is Lipschitz continuous. Therefore, by the Lipschitz property we can write $W(k^* + a) - W(k^*) = \int_{k^*}^{k^*+a} W'(k)dk$ for all $a > 0$. Since $\lim_{a\to 0^+} W'(k^* + a) = W'_+(k^*)$, the right-hand derivative $W'_+(k^*)$ is obviously well defined; further, for every $\xi > 0$ there is $\alpha > 0$ such that $|W'_+(k^*) - \frac{\int_{k^*}^{k^*+a} W'(k)dk}{a}| < \xi$ for all $0 < a < \alpha$. Hence, possible discontinuities in the policy function $g(k^* + a)$ do not have a first-order effect on the continuity properties of the term $\frac{\int_{k^*}^{k^*+a} W'(k)dk}{a}$ as a function of a for sufficiently small a, and such term defines the right-hand derivative $W'_+(k^*)$. Finally, let us note that by the envelope theorem, $\int_{k^*}^{k^*+a} W'(k)dk = \int_{k^*}^{k^*+a} u'(f(k) - g(k))f'(k)dk$.

Part (2). The continuation value function V has a well-defined right-hand derivative $V'_+(k^*)$. This part is an extension of Theorem 2.6, using first-order condition $u'(f(k)-g(k)) = \beta\delta V'(g(k))$ at points of existence of $V'(g(k))$. As already discussed, both $W(k)$ and $g(k)$ are differentiable at almost all k, and so $V(k)$ is differentiable at almost all k. In addition, at a maximum point Clarke's generalized directional derivative [see (49)] must be less than or equal to zero. This means that for fixed $\xi > 0$ we have $V'(k^* + a) \leq \frac{u'(c^*)}{\beta\delta} + \xi$ for all $a > 0$ small enough; recall that $V'(k^* + a)$ exists at almost all points. Further, by Bellman's equation, $\beta V(k) = W(k) - (1 - \beta)u(f(k) - g(k))$, the derivative of g must be bounded at all points $k^* + a$ for all $a > 0$ sufficiently small, whenever $g'(k^* + a)$ does exist.

Observe that $k^* = g(k^*)$ is an interior steady state, and g is strictly monotone increasing and continuous from the right. Consequently, $lim\ sup_{a \to 0^+} V'(k^* + a) \geq \frac{u'(c^*)}{\beta\delta}$, where again the *lim sup* is taken over those points of existence of $V'(k^* + a)$. By way of contradiction, assume that there is a subsequence such that $lim_{a_j \to 0^+} V'(k^* + a_j) \leq \frac{u'(c^*)}{\beta\delta} - \xi$, for some $\xi > 0$. Since the policy function g is monotonically increasing, along this subsequence function g must have unbounded slope (or be discontinuous) at some points k arbitrarily close to k^* with $k > k^*$. Accordingly, the continuation value function V should also have unbounded slope for points k arbitrarily close to k^*. This can be ruled out by similar arguments as in the previous paragraph. Taken all these inequalities together about the limiting behavior of the derivative $V'(k)$, as we approach k^* from the right the generalized gradient of V at k^* as defined in (50) must be unique (from the right side). Then, by Proposition A.1 of the Online Appendix the right-hand derivative $V'_+(k^*)$ must be well defined, and must be equal to $\frac{u'(c^*)}{\beta\delta}$.

Therefore, we can write: $V(k^* + a) - V(k^*) = \int_{k^*+a}^{k^*} V'(k)dk + \textit{high-order terms}$, where $\int_{k^*}^{k^*+a} V'(k)dk = \frac{1}{\beta\delta} \int_{k^*}^{k^*+a} u'(f(k) - g(k))dk$. Thus, we can conclude that if the right-hand derivative $W'_+(k^*)$ is well defined, then the right-hand derivative $V'_+(k^*)$ should also be well defined. Indeed, one can readily see that $\frac{\int_{k^*}^{k^*+a} u'(f(k)-g(k))dk}{a}$ shares the same continuity properties as a function of a approaching $a = 0$ as the integral $\frac{\int_{k^*}^{k^*+a} u'(f(k)-g(k))f'(k)dk}{a}$ in Part (1).

Part (3). The policy function g has a well-defined right-hand derivative $g'_+(k^*)$. As in Theorem 2.7, by the inverse function theorem and the chain rule applied to utility function u in Bellman's equation $\beta V(k) = W(k) - (1 - \beta)u(f(k) - g(k))$ and Parts (1)-(2) of this proof, we get that the right-hand derivative $g'_+(k^*)$ is well defined. □

Appendix C Results for Theorem 3.3

Lemma C.1 (Addendum to Proof, Part (ii) of Theorem 3.8)**.** *Under the conditions of Theorem 3.8, we must have*

$$W^*(\theta) \leq \max_{\theta_+ \in \Gamma(\theta)} U(\theta, \theta_+) + \beta\delta V^*(\theta_+)$$

for all $\theta \in \Theta$.

Proof. We break the proof in a few simple steps.

(i) Let us define a new function $\hat{V}_n^*(\theta) \geq V_{\varepsilon_{n'}}^*(\theta)$ for all θ in Θ and all $n' \geq n$. More specifically, let $\hat{V}_n^*(\theta) = \sup_{n' \geq n}\{V_{\varepsilon_{n'}}^*(\theta)\}$ for each θ in Θ. Then, $\hat{V}_n^*(\theta) \geq V_{\varepsilon_{n'}}^*(\theta)$ for all θ in Θ and $n' \geq n$, where $V_{\varepsilon_{n'}}^*(\theta)$ refers to a continuation value function for an ε-equilibrium (for some $\varepsilon_{n'}$) as defined in part (ii) of the proof of Theorem 3.8. Incidentally, note that $\hat{V}_n^*(\theta)$ is lower semicontinuous, since it is the *sup* of an infinite sequence of continuous functions $V_{\varepsilon_n}^*(\theta)$. As already discussed, there is no restriction of generality to take the upper closure of $\hat{V}_n^*(\theta)$. Hence, we suppose that $\hat{V}_n^*(\theta)$ is a continuous mapping. Moreover, observe that $\hat{V}_n^*(\theta)$ is non-increasing in n for every θ.

(ii) Accordingly, let us define a new current value function $\hat{W}_n^*(\theta) \geq W_{\varepsilon_{n'}}^*(\theta)$ for all θ in Θ and all $n' \geq n$. Certainly, $\hat{W}_n^*(\theta) = \max_{\theta_+ \in \Gamma(\theta)} U(\theta, \theta_+) + \beta\delta\hat{V}_n^*(\theta_+)$ for all θ in Θ, satisfies these properties. Moreover, $\hat{W}_n^*(\theta)$ is Lipschitz continuous.

(iii) For each θ_0 we can choose a subsequence of maximizers $\{\hat{\theta}_{1n}^*(\theta_0)\}_{n\geq 0}$ converging to some $\theta_1^*(\theta_0)$ because Θ is a compact set. Here,

$$\hat{\theta}_{1n}^*(\theta_0) = \arg\max_{\theta_+ \in \Gamma(\theta_0)} \{U(\theta_0, \theta_+) + \beta\delta\hat{V}_n^*(\theta_+)\}.$$

That is, $\hat{\theta}_{1n}^*(\theta_0)$ is a maximizer under value function $\hat{V}_n^*(\theta_+)$ and initial condition θ_0.

(iv) Finally, we claim that $\max_{\theta_+ \in \Gamma(\theta)} U(\theta, \theta_+) + \beta\delta V^*(\theta_+) \geq W^*(\theta)$ *for all θ in Θ,* which proves the lemma. Let $\alpha > 0$ be an arbitrarily small number. Then, we first claim that $V^*(\theta_1^*(\theta_0)) + \alpha > \hat{V}_{n'}^*(\theta_+)$ for θ_+ in a neighborhood of $\theta_1^*(\theta_0)$ and all $n' \geq n$. Observe that $\{\hat{V}_n^*(\theta_1^*(\theta_0))\}_{n\geq 0}$ converges to $V^*(\theta_1^*(\theta_0))$ for fixed θ_0, as $n \to \infty$; moreover, $\hat{V}_n^*(\theta_+)$ is a continuous function. Since $\hat{V}_n^*(\theta_+)$ is also nonincreasing in n, we must have $V^*(\theta_1^*(\theta_0)) + \alpha > \hat{V}_{n'}^*(\theta_1)$ for all θ_1 in a neighborhood of $\theta_1^*(\theta_0)$ for all $n' \geq n$ for n sufficiently large. Therefore,

$$\alpha + U(\theta_0, \theta_1^*(\theta_0)) + \beta\delta V^*(\theta_1^*(\theta_0)) \geq U(\theta_0, \hat{\theta}_{1n'}^*(\theta_0)) + \beta\delta\hat{V}_{n'}^*(\hat{\theta}_{1n'}^*(\theta_0)), \tag{52}$$

for all $n' \geq n$ and n large enough. Then,

$$\begin{aligned}\max_{\theta_+ \in \Gamma(\theta_0)} \{U(\theta_0, \theta_+) + \beta\delta V^*(\theta_+)\} + \alpha &\geq U(\theta_0, \hat{\theta}_{1n'}^*(\theta_0)) + \beta\delta\hat{V}_{n'}^*(\hat{\theta}_{1n'}^*(\theta_0)) \\ &\geq \hat{W}_{n'}^*(\theta_0) \\ &\geq W^*(\theta_0) - \alpha\end{aligned}$$

Since our choices θ_0 and $\alpha > 0$ were arbitrary, it must hold true that $\max_{\theta_+ \in \Gamma(\theta)}\{U(\theta, \theta_+) + \beta\delta V^*(\theta_+)\} \geq W^*(\theta)$ for all $\theta \in \Theta$. □

Lemma C.2. *Under the conditions of Theorem 3.8,*
(i) Let V^* be the lim sup of the sequence $\left\{V^*_{\varepsilon_n}\right\}_{n\geq 0}$.
(ii) Let $W^*_{\varepsilon_n}$ be the value function for $V^*_{\varepsilon_n}$ for each $n \geq 0$, and W^* be the value function for V^*; see (37).
(iii) Let the sequence $\left\{W^*_{\varepsilon_n}\right\}_{n\geq 0}$ converge to W^* in the sup norm.
(iv) For given θ_0, let $\theta^*_{1\varepsilon_n}(\theta_0)$ be an optimal point for function $W^*_{\varepsilon_n}$ under continuation value function $V^*_{\varepsilon_n}$.
Suppose that $\theta^*_1(\theta_0)$ is a limit point of a sequence of optimal points $\{\theta^*_{1\varepsilon_n}(\theta_0)\}_{n\geq 0}$. Then, $\theta^*_1(\theta_0)$ is an optimal point for value function W^* under continuation value function V^*.

Proof. For every $\alpha > 0$ there is some n large enough such that

$$\begin{aligned} W^*(\theta_0) &\geq U(\theta_0, \theta^*_1(\theta_0)) + \beta\delta V^*(\theta^*_1(\theta_0)) \\ &\geq U(\theta_0, \theta^*_{1\varepsilon_n}(\theta_0)) + \beta\delta V^*_{\varepsilon_n}(\theta^*_{1\varepsilon_n}(\theta_0)) - \alpha \\ &\geq W^*_{\varepsilon_n}(\theta_0) - \alpha \\ &\geq W^*(\theta_0) - 2\alpha, \end{aligned}$$

where the first inequality comes from the definition of W^*, see (ii) in the statement of this lemma; the second inequality comes from the definition of V^* as the limit sup; see (i) in the statement of this lemma, and equation (52) in the proof of Lemma C.1; the third inequality comes from (ii) in the statement of this lemma; and the fourth inequality comes from (iii) in the statement of this lemma for n large enough. Since $\alpha > 0$ has been chosen arbitrarily small, the proof is thus established.

□

Let $\tilde{\Lambda}^*_{\varepsilon_n}$ be the correspondence of optimal solutions for value functions $W^*_{\varepsilon_n}$ and $V^*_{\varepsilon_n}$; see (37). Note that correspondence $\tilde{\Lambda}^*_{\varepsilon_n}$ should be distinguished from $\Lambda^*_{\varepsilon_n}$, which is computed after the ε_n-derivative in (41).

Corollary C.3. *Under the conditions of the previous Lemma C.2, we must have:*
$\limsup \left\{\tilde{\Lambda}^*_{\varepsilon_n}\right\}_{n\geq 0} \subseteq \Lambda^*$.

The proof of Corollary C.3 follows as a simple extension of Lemma C.2 since Θ is a compact set. Hence, every sequence of optimal points $\{\theta^*_{1\varepsilon_n}(\theta_n)\}_{n\geq 0}$ must have a convergent subsequence over both $\theta^*_{1\varepsilon_n}$ and θ_n.

Remark C.4. Following Hildenbrand (1974), p. 16], note that *lim sup* $\left\{\tilde{\Lambda}^*_{\varepsilon_n}\right\}_{n\geq 0} \subseteq \Lambda^*$ entails that for every $\alpha > 0$, there is n such that

$$\tilde{\Lambda}^*_{\varepsilon_{n'}} \subseteq B_\alpha(\Lambda^*) \tag{53}$$

for all $n' \geq n$, where $B_\alpha(\Lambda^*)$ denotes the α-neighborhood of set Λ^*.

Lemma C.5 (Addendum to Proof, Part (iv) of Theorem 3.8). *For every $\theta_0 \in \Theta^{W^*}$, the sequence $\left\{\Lambda^*_{\varepsilon_n}(\theta_0)\right\}_{n\geq 0}$ converges to $\theta_1^*(\theta_0)$.*

Proof. Note that $W^*(\theta)$ is differentiable at $\theta_0 \in \Theta^{W^*}$, and the derivative $DW^*(\theta)$ varies continuously with θ. Further, by the previous Remark C.4 for every $\alpha > 0$, there is ε_n small enough such that the ε-derivative as defined in (40) is computed within the set $B_\alpha(\Lambda^*)$, and is roughly equal to $DW(\theta_0)$. In other words, by the envelope theorem all derivatives in the integrand of (40) take the form $DW^*_{\varepsilon_n}(\theta) = D_1U(\theta, \theta^*_{1\varepsilon_n}(\theta))$, and all $\theta^*_{1\varepsilon_n}(\theta)$ become arbitrarily close to $\theta_1^*(\theta_0)$ for θ closed enough to θ_0 for n sufficiently large, since $\Lambda^\star$ is an upper hemi-continuous correspondence that is single-valued at θ_0. Therefore, by Assumption 3(iv) we must have that the sequence $\left\{\Lambda^*_{\varepsilon_n}(\theta_0)\right\}_{n\geq 0}$ converges to $\theta_1^*(\theta_0)$. $\square$

Appendix D Fixed-Point Theorems

In 1912, Brouwer proved the first fixed-point theorem in Euclidean spaces: Every continuous mapping from the *n-simplex* to itself has a fixed point. In 1930, Schauder extended Brouwer's fixed-point theorem from Euclidean spaces to Banach spaces.

Theorem D.1. *(Schauder's Fixed Point Theorem). Let Σ be a nonempty closed convex subset of a Banach space $\mathbb{X}$. If $f : \Sigma \to \Sigma$ is continuous with a compact image, then f has a fixed point.*

Schauder also came up with the well-known conjecture: Every continuous function, from a nonempty compact and convex set in a (Hausdorff) topological vector space into itself, has a fixed point.

R. Cauty provided an answer to Schauder's Conjecture. In the international conference of Fixed Point Theory and its Applications in 2005, T. Dobrowolski remarked that there is a gap in the proof. Therefore, Schauder's Conjecture is still unsolved.

In 1934, Tychonoff extended Schauder's fixed-point theorem from a Banach space to a locally convex topological vector space.

Theorem D.2. *(Tychonoff's Fixed Point Theorem). Let $\mathbb{X}$ be a Hausdorff locally convex topological vector space. Assume that Σ is a compact convex subset of $\mathbb{X}$. Then, every continuous function $f : \Sigma \to \Sigma$ has a fixed point.*

Our discussion here follows Li (2021), where one can find further applications and extensions of these theorems. Theorem D.1 applies directly to our framework.

Appendix E Numerical Algorithm for Computing the Markov Equilibrium

The following algorithm details a numerical procedure for computing the Markov equilibrium in the quasi-hyperbolic discounting model. The iteration is performed on the pair (V, g)—the continuation value function and the policy function, respectively. This iteration includes a smoothing step to handle potential discontinuities in the policy function and ensure numerical stability.

Algorithm 1.

MPE Solver with ε-Approximate Derivatives

```
Tolerances and Initialization
Tolerances: tol > 0, jump threshold $\tau > 0$, derivative radius $\varepsilon > 0$ (initial), target $\varepsilon_{\min} > 0$.
Base state grid $\Theta = \{\theta_1, \dots, \theta_N\}$ and initial envelope guess $W_{\text{guess}}$ (bounded, Lipschitz).
Initial correspondence/policy guess $\Lambda_{\text{guess}}$ (defined where single-valued).
Weight kernel $w(\theta, \theta')$ for local averaging; let $\omega(\theta, \theta') := \dfrac{w(\theta, \theta')}{\sum_{\theta'' \in \mathcal{N}_\varepsilon(\theta)} w(\theta, \theta'')}$.

Outer continuation loop on $\varepsilon$
repeat
  $(W_{\text{new}}, \Lambda_{\text{new}}) \leftarrow (W_{\text{guess}}, \Lambda_{\text{guess}})$
  Inner fixed-point loop at fixed $\varepsilon$
  repeat
    $W_{\text{old}} \leftarrow W_{\text{new}}$
    Step 1: $\varepsilon$-approximate derivative and implied policy
    for each $\theta \in \Theta$ do
      Neighborhood: $\mathcal{N}_\varepsilon(\theta) := \{\theta' \in \Theta : \|\theta' - \theta\| \le \varepsilon\}$
```

```
            Compute finite-difference derivative estimates $\widehat{DW}(\theta' \mid W_{\text{old}})$ at each $\theta' \in \mathcal{N}_\varepsilon(\theta)$.
            Weighted finite average:
                $D_\varepsilon W_{\text{old}}(\theta) := \sum_{\theta' \in \mathcal{N}_\varepsilon(\theta)} \omega(\theta, \theta')\, \widehat{DW}(\theta' \mid W_{\text{old}})$
            Recover next state from the gradient relation:
                $\tilde{g}_\varepsilon(\theta) \leftarrow h\big(D_\varepsilon W_{\text{old}}(\theta),\, \theta\big),$
        where $h(x, \theta)$ satisfies $x = D_1 U(\theta, \theta_+)$ and returns $\theta_+ = h(x, \theta)$.
            Define guess for single-valued correspondence $\tilde{\Lambda}_\varepsilon(\theta) = \{\tilde{g}_\varepsilon(\theta)\}$.
          **Step 2: Continuation value (upper closure)**
          **for** each $\theta \in \Theta$ **do**
            $\tilde{V}_{\text{raw}}(\theta) \leftarrow \frac{1}{\beta}\Big(W_{\text{old}}(\theta) - (1-\beta)\, U(\theta, \tilde{g}_\varepsilon(\theta))\Big)$
          *Upper envelope on $\Theta$ (local-sup majorant)*
          **for** each $\theta \in \Theta$ **do**
            $\mathcal{N}_\rho(\theta) \leftarrow \{\theta' \in \Theta : \ \|\theta' - \theta\| \le \rho(\theta)\}$
            $\tilde{V}^{suc}(\theta) \leftarrow \max_{\theta' \in \mathcal{N}_\rho(\theta)} \tilde{V}_{\text{raw}}(\theta')$
          $\Theta_+ \leftarrow \Theta$
          **Step 3: Bellman update and maximizers**
          **for** each $\theta \in \Theta$ **do**
            $\hat{W}(\theta) \leftarrow \max_{\theta_+ \in \Gamma(\theta)} \Big\{U(\theta, \theta_+) + \beta\delta\, \tilde{V}^{suc}(\theta_+)\Big\}$
            $\hat{\Lambda}(\theta) \leftarrow \arg\max_{\theta_+ \in \Gamma(\theta)} \Big\{U(\theta, \theta_+) + \beta\delta\, \tilde{V}^{suc}(\theta_+)\Big\}$
          **Value update**
          $W_{\text{new}} \leftarrow \hat{W}$
          **Convergence check**
          **if** $\sup_{\theta \in \Theta} |W_{\text{new}}(\theta) - W_{\text{old}}(\theta)| < \texttt{tol}$ **then**
            **break**
        **until** convergence is reached
        **Continuation step:** $\varepsilon \leftarrow \varepsilon/2$;   $(W_{\text{guess}}, \Lambda_{\text{guess}}) \leftarrow (W_{\text{new}}, \hat{\Lambda})$
      **until** $\varepsilon \le \varepsilon_{\min}$
      **Outputs:** $W(\cdot)$, $V(\cdot) = \tilde{V}^{suc}(\cdot)$, $\Lambda(\cdot)$, and policy $g(\cdot)$.
```
