APPENDIX TO: "USING INVERSE EULER EQUATIONS TO SOLVE MULTIDIMENSIONAL DISCRETE-CONTINUOUS DYNAMIC MODELS: A GENERAL METHOD"

LORETTI I. DOBRESCU* AND AKSHAY SHANKER†

Appendix Appendix D Supplementary appendix to Section 2

D.1 Additional notation. In addition to the notational conventions introduced in the main paper, we list, for clarity, the following standard mathematical conventions used in this online appendix. We use $\|\cdot\|$ to denote to the Euclidean norm. For a set $S \subset \mathbb{R}^n$, S^o refers to its interior, and for $x \in \mathbb{R}^n$, $\mathbb{B}_\epsilon(x)$ refers to the ϵ-ball about x. For an $m \times n$ matrix $\mathbf{A}$ and $n \times p$ matrix $\mathbf{B}$, $\mathbf{AB}$ denotes the matrix product. When $\mathbf{B}$ is also an $m \times n$ matrix, $\mathbf{A} \circ \mathbf{B}$ will denote the Hadamard product. For two euclidean column vectors a and b, we use $a \cdot b$ to denote the inner-product. For a function $\mathrm{f}\colon \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^m$, partial derivatives and the Jacobian with respect to the argument x are denoted using $\partial_x \mathrm{f}(x, y)$. Moreover, letting $\mathrm{f} = (\mathrm{f}_1, \dots, \mathrm{f}_m)$, the Jacobian can be written as $\partial_x \mathrm{f}(x, y) = [\nabla_x \mathrm{f}_1(x, y), \dots, \nabla_x \mathrm{f}_m(x, y)]^\top$, where $\nabla_x \mathrm{f}_1(x, y)$ denotes gradient column vectors and $\top$ denotes the matrix transpose.

We use $\mathbf{0}$ to refer to the null vector in a vector space, use $\mathbb{R}_{+\infty}(\mathbb{R}_{-\infty})$ to denote the space $\mathbb{R} \cup \{+\infty\}(\mathbb{R} \cup \{-\infty\})$, and $\bar{\mathbb{R}}$ to denote the space $\mathbb{R} \cup \{+\infty, -\infty\}$. The effective domain of a functional f, $\operatorname{dom} \mathrm{f}$, will be all $\boldsymbol{x} \in \boldsymbol{X}$ such that $\mathrm{f}(\boldsymbol{x}) \in \mathbb{R}$.

Recall that for a real topological vector space $\boldsymbol{X}$, $\boldsymbol{X}^\star$ denotes the topological dual space - i.e., the space of all continuous linear functions on $\boldsymbol{X}$ (Luenberger, 1997; Aliprantis and Border, 2006). For $\boldsymbol{x}$ with $\boldsymbol{x} \in \boldsymbol{X}$, and $\boldsymbol{x}^\star$ with $\boldsymbol{x}^\star \in X^\star$, we use $\boldsymbol{x}^\star \boldsymbol{x}$ to refer to the evaluation of $\boldsymbol{x}^\star$ at $\boldsymbol{x}$ - i.e. an evaluation of the mapping $\boldsymbol{x} \mapsto \langle \boldsymbol{x}^\star, \boldsymbol{x} \rangle$. Moreover, let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be vector spaces. For a function f with $\mathrm{f}\colon \boldsymbol{X} \to \boldsymbol{Y}$, and for $\boldsymbol{y}^\star \in \boldsymbol{Y}^\star$, we will use $\boldsymbol{y}^\star \mathrm{f}$ to denote the mapping $\boldsymbol{x} \mapsto \langle \mathrm{f}(\boldsymbol{x}), \boldsymbol{y}^\star \rangle$, and $\boldsymbol{y}^\star f(\boldsymbol{x})$ to denote the evaluation of the mapping.

D.2 Additional results.

Definition 1. *(**Supdifferential**) Let $\boldsymbol{X}$ be a normed space, and let $\mathrm{f}\colon \boldsymbol{X} \to \mathbb{R}_{-\infty}$. The supdifferential of* f *at $\boldsymbol{x}$, denoted by $\partial^S \mathrm{f}(\boldsymbol{x})$, is the set such that (i) if $\boldsymbol{x} \in \operatorname{dom} \mathrm{f}$,*

* Corresponding author. School of Economics, University of New South Wales, Sydney 2052 Australia. Email: dobrescu@unsw.edu.au.

† School of Economics, University of New South Wales, Australia and ARC Center for Excellence in Population Ageing Research (CEPAR), Sydney 2052 Australia. Email: a.shanker@unsw.edu.au.

‡ The latest version of this appendix can be found here here.

then $\partial^S \mathrm{f}(\boldsymbol{x})$ *is the set of* $x^\star \in \boldsymbol{X}^\star$ *such that*

$$\mathrm{f}(\boldsymbol{s}) - \mathrm{f}(\boldsymbol{x}) \leq \langle \boldsymbol{x}^\star, \boldsymbol{s} - \boldsymbol{x} \rangle, \qquad \forall \boldsymbol{s} \in \boldsymbol{X},$$

and (ii) if $\boldsymbol{x} \notin \operatorname{dom} \mathrm{f}$*, then* $\partial^S \mathrm{f}(\boldsymbol{x}) = \emptyset$.

Note that the elements of $\partial^S \mathrm{f}(\boldsymbol{x})$ are supderivatives of f at $\boldsymbol{x}$.

Definition 2. ***(Concave function)*** *Let* $\boldsymbol{X}$ *be a vector space, and let* $\boldsymbol{Y}$ *be a vector space specified with a positive cone* $\boldsymbol{P}$*. The mapping* $\mathrm{f}\colon \boldsymbol{X} \to \boldsymbol{Y}$ *is concave if* $\operatorname{dom} \mathrm{f}$ *is convex, and for all* $\eta \in (0,1)$*, we have*

$$\mathrm{f}(\eta \boldsymbol{x} + (1-\eta)\boldsymbol{y}) \geq \eta \mathrm{f}(\boldsymbol{x}) + (1-\eta)\mathrm{f}(\boldsymbol{y}), \qquad \boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} \mathrm{f}.$$

For a proof of the following, see Proposition 16.7 by Bauschke and Combettes (2010).

Fact 1. ***(Partial and total supdifferentials)*** *Let* $\mathrm{f}\colon \prod_{i\in\mathcal{F}} \boldsymbol{X}_i \to \mathbb{R}_{-\infty}$*, where* $\mathcal{F}$ *with* $N = |\mathcal{F}|$ *indexes a finite family of normed vector spaces* $(\boldsymbol{X}_i)_{i\in\mathcal{F}}$*. Since* $\mathcal{F}$ *is finite, the norm dual of* $\prod_{i\in\mathcal{F}} X_i$ *will be* $\prod_{i\in\mathcal{F}} X_i^\star$*, with* $x^\star(x) = \langle \boldsymbol{x}, \boldsymbol{x}^\star \rangle = \sum_{i\in\mathcal{F}} \langle x_i, x_i^\star \rangle$*. Let* $\partial^S \mathrm{f}(\boldsymbol{x})$ *denote the subdifferential of* f *at* $\boldsymbol{x}$*. The evaluation mapping defines the i'th partial supdifferential as* $\operatorname{ev}(\partial^S \mathrm{f}, \boldsymbol{i}) =\colon \partial^S \mathrm{f}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)_i$*. Fix* $\boldsymbol{x} \in \boldsymbol{X}$*, then define* $\mathrm{R}_i\colon \boldsymbol{X}_i \to \boldsymbol{X}$ *as follows: the* $\boldsymbol{j}'\boldsymbol{th}$ *component of* $\mathrm{R}_i y$ *equals* $\boldsymbol{y}$ *if* $\boldsymbol{i} = j$*, and* $\boldsymbol{x}_j$ *otherwise. We define the following notation:*

$$\partial^S \mathrm{f}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)_i \colon = \partial^S (f \circ \mathrm{R}_i)(\boldsymbol{x}_i) \tag{1}$$

and moreover, we have that $\partial^S \mathrm{f}(\boldsymbol{x}) \subset \prod_{i\in\mathcal{F}} \partial^S (\mathrm{f} \circ \mathrm{R}_i)(\boldsymbol{x}_i)$.

Claim D.1. *Let* $\boldsymbol{X} = L^2(\Omega, \mathbb{R}^N)$ *and* $\boldsymbol{\phi} \in \boldsymbol{X}$*, consider* $\mathrm{f} : \mathbb{R}^N \to \mathbb{R}_{-\infty}$*, and fix* $\boldsymbol{x} \in \boldsymbol{X}$*. If*

$$\langle \boldsymbol{\phi}, \boldsymbol{s} - \boldsymbol{x} \rangle \geq \int \mathrm{f}(\boldsymbol{s}) - \mathrm{f}(\boldsymbol{x}) \, d\,\mathbb{P}, \qquad \forall \boldsymbol{s} \in \boldsymbol{X},$$

then for any $\boldsymbol{s} \in \boldsymbol{X}$*,* $\boldsymbol{\phi}(\boldsymbol{s} - \boldsymbol{x}) \geq \mathrm{f}(\boldsymbol{s}) - \mathrm{f}(\boldsymbol{x})$ *a.e.*

Proof. Assume by contradiction that there exists $\tilde{\boldsymbol{s}} \in \boldsymbol{X}$ and a non-null set $A \subseteq \Omega$ with strictly positive measure such that for all $\omega \in A$,

$$\boldsymbol{\phi}(\tilde{\boldsymbol{s}}(\omega) - \boldsymbol{x}(\omega)) < \mathrm{f}(\tilde{\boldsymbol{s}}(\omega)) - \mathrm{f}(\boldsymbol{x}(\omega)).$$

Define a function $\boldsymbol{s} : \Omega \to \mathbb{R}^N$ by

$$\boldsymbol{s}(\omega) = \begin{cases} \tilde{\boldsymbol{s}}(\omega) & \text{if } \omega \in A, \\ \boldsymbol{x}(\omega) & \text{otherwise,} \end{cases}$$

where $\tilde{\boldsymbol{s}}(\omega)$ is chosen such that the strict inequality above is satisfied for $\omega \in A$. By construction of $\boldsymbol{s}$, $\boldsymbol{s}(\omega) - \boldsymbol{x}(\omega) = 0$ for $\omega \notin A$, and $\boldsymbol{s}(\omega) - \boldsymbol{x}(\omega) = \tilde{\boldsymbol{s}}(\omega) - \boldsymbol{x}(\omega)$ for $\omega \in A$. Thus,

$$\langle \boldsymbol{\phi}, \boldsymbol{s} - \boldsymbol{x} \rangle = \int_A \boldsymbol{\phi}(\tilde{\boldsymbol{s}}(\omega) - \boldsymbol{x}(\omega)) \, \mathbb{P}(d\omega),$$

Given the strict inequality on A, integrating over A yields

$$\langle \boldsymbol{\phi}, \boldsymbol{s} - \boldsymbol{x} \rangle < \int_A (\mathrm{f}(\tilde{\boldsymbol{s}}(\omega)) - \mathrm{f}(\boldsymbol{x}(\omega)))\, \mathbb{P}(d\omega), = \int \mathrm{f}(\boldsymbol{s}) - \mathrm{f}(\boldsymbol{x})\, d\mathbb{P},$$

which contradicts the assumption that

$$\langle \boldsymbol{\phi}, \boldsymbol{s} - \boldsymbol{x} \rangle \geq \int (\mathrm{f}(\boldsymbol{s}) - f(\boldsymbol{x}))\, dP \quad \text{for all } \boldsymbol{s} \in \boldsymbol{X}.$$

Thus, our contradictory assumption must be false, implying that there cannot exist a non-null set A where the inequality is violated. Consequently,

$$\boldsymbol{\phi}(\boldsymbol{s}(\omega) - \boldsymbol{x}(\omega)) \geq \mathrm{f}(\boldsymbol{s}(\omega)) - \mathrm{f}(\boldsymbol{x}(\omega))$$

a.e., as was to be shown. □

Proposition D.1. ***(Global maxima of integral operator)*** *If (i)* $\mathrm{f}(\cdot, \boldsymbol{y})$ *is concave almost everywhere (a.e.) for fixed* $\boldsymbol{y}$*, (ii)* $\partial \mathrm{f}_x(\boldsymbol{x}, \boldsymbol{y}) = 0$ *a.e. for* $\boldsymbol{x} \in \operatorname{dom} \mathrm{f}$*, and (iii)* $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{X}^\star$*, then* $\boldsymbol{x}$ *is a global maximum of* I_f*.*

Proof. Let $\boldsymbol{x} \in \operatorname{dom} \mathrm{f}$. By Fermat's theorem, $\boldsymbol{x}$ is a global maximum of I_f if and only if $\mathbf{0} \in \partial^S \mathrm{I}_\mathrm{f}(\boldsymbol{x})$. By the definition of the supdifferential, $\mathbf{0} \in \partial^S \mathrm{I}_\mathrm{f}(\boldsymbol{x})$ if and only if for all $\boldsymbol{s} \in \boldsymbol{X}$,

$$\mathrm{I}_\mathrm{f}(\boldsymbol{s}) - \mathrm{I}_\mathrm{f}(\boldsymbol{x}) \leq \langle \mathbf{0}, \boldsymbol{s} - \boldsymbol{x} \rangle.$$

Since $\partial_x \mathrm{f}(x, y) = 0$ a.e., we will show that (i) $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$, and (ii) $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) \in \partial^S \mathrm{I}_\mathrm{f}(\boldsymbol{x})$. Showing $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$ is straightforward since for any $\boldsymbol{s}^\star \in \boldsymbol{X}^\star$, we have

$$\begin{aligned}(2)\quad \mathbb{P}\left\{\boldsymbol{s}^\star + \partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) \neq \boldsymbol{s}^\star\right\} &= \mathbb{P}\left\{\omega \mid \boldsymbol{s}^\star(\omega) + \partial_x \mathrm{f}(\boldsymbol{x})(\omega) \neq \boldsymbol{s}^\star(\omega)\right\} = \\ &\mathbb{P}\left\{\omega \mid \boldsymbol{s}^\star(\omega) + \partial_x \mathrm{f}(\boldsymbol{x}(\omega)) \neq \boldsymbol{s}^\star(\omega)\right\} = \mathbb{P}\left\{\omega \mid \boldsymbol{s}^\star(\omega) \neq \boldsymbol{s}^\star(\omega)\right\} = 0,\end{aligned}$$

and thus $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$.

Next, to show $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) \in \partial^S \mathrm{I}_\mathrm{f}(\boldsymbol{x})$, by the definition of the integral operator,

$$\mathrm{I}_\mathrm{f}(\boldsymbol{s}) - \mathrm{I}_\mathrm{f}(\boldsymbol{x}) = \int \mathrm{f}(\boldsymbol{s}(\omega), \boldsymbol{y}(\omega)) - \mathrm{f}(\boldsymbol{x}(\omega), \boldsymbol{y}(\omega))\mathbb{P}(\mathrm{d}\omega).$$

If $\boldsymbol{s} \notin \operatorname{dom} \mathrm{f}$, then $\int \mathrm{f}(\boldsymbol{s}(\omega), \boldsymbol{y}(\omega)) - \mathrm{f}(\boldsymbol{x}(\omega), \boldsymbol{y}(\omega)))\mathbb{P}(d\omega) = -\infty$, thus $\int \mathrm{f}(\boldsymbol{s}(\omega), \boldsymbol{y}(\omega)) - \mathrm{f}(\boldsymbol{x}(\omega), \boldsymbol{y}(\omega))\mathbb{P}(d\omega) \leq \langle \partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{s} - \boldsymbol{x} \rangle$. Thus, $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) \in \partial^S \mathrm{I}_\mathrm{f}(\boldsymbol{x})$.

On the other hand, if $\boldsymbol{s} \in \operatorname{dom} \mathrm{f}$, then

$$\begin{aligned}&\int \mathrm{f}(\boldsymbol{s}(\omega), \boldsymbol{y}(\omega)) - \mathrm{f}(\boldsymbol{x}(\omega), \boldsymbol{y}(\omega))\mathbb{P}(\mathrm{d}\omega) \\ &\qquad \leq \int \langle \partial_x \mathrm{f}(\boldsymbol{x}(\omega), \boldsymbol{y}(\omega))), \boldsymbol{s}(\omega) - \boldsymbol{x}(\omega) \rangle \mathbb{P}(d\omega) = \langle \partial_x \mathrm{f}(\boldsymbol{x}), \boldsymbol{s} - \boldsymbol{x} \rangle,\end{aligned}$$

where inequality follows since

$$\mathbb{P}\left\{\omega \mid \partial_x \mathrm{f}(\boldsymbol{x}(\omega), \boldsymbol{y}(\omega))(\boldsymbol{s}(\omega) - \boldsymbol{x}(\omega)) < \mathrm{f}(\boldsymbol{s}(\omega), \boldsymbol{y}(\omega)) - \mathrm{f}(\boldsymbol{x}(\omega), \boldsymbol{y}(\omega))\right\} = 0,$$

by the concavity of f. Thus, $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) \in \partial^S \mathrm{I}_\mathrm{f}(\boldsymbol{x})$, and since $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$, we have $\mathbf{0} \in \partial^S \mathrm{I}_\mathrm{f}(\boldsymbol{x})$. □

Remark D.1. *Consider the setting of Proposition D.1 and consider the mapping* $\boldsymbol{x} \mapsto \partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y})$ *on* $\boldsymbol{X}$*. The evaluation* $\partial_x \mathrm{f}(x, y)$ *will be a Euclidean vector of length* N *for all* $x \in \operatorname{dom} \mathrm{f}$*. As such, if* $\partial_x \mathrm{f}(\cdot, y)$ *is measurable, then* $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y})$ *is an* $\mathbb{R}^N$*-valued random variable on* $(\Omega, \Sigma, \mathbb{P})$*, and if* $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y})$ *has finite variance, then* $\partial_x \mathrm{f}(\boldsymbol{x}, \boldsymbol{y}) \in L^2(\Omega, \mathbb{R}^N, \Sigma)$*.*

Theorem D.1. ***(Fermat's theorem)****A function* f *on a normed space* $\boldsymbol{X}$ *attains its maximum at* $\boldsymbol{x} \in \boldsymbol{X}$ *if and only if* $\boldsymbol{0} \in \partial^S \mathrm{f}(\boldsymbol{x})$*.*

Proof. See Theorem 16.2 in Bauschke and Combettes (2010). ◻

Appendix Appendix E Supplementary appendix to Section 3

Given a sequence of multipliers $\{\boldsymbol{\mu}_t\}_{t=0}^T$ and $\{\boldsymbol{\lambda}_t\}_{t=0}^T$, where $\boldsymbol{\lambda}_t \in L^2\left(\Omega, \mathbb{R}^{N_M}, \mathscr{F}_t\right)$ and $\boldsymbol{\mu}_t \in L^2\left(\Omega, \mathbb{R}^{N_g}, \mathscr{F}_t\right)$, and a sequence of discrete choices $\vec{\boldsymbol{d}}_0$, define the perturbation function $\mathbf{H}_t \colon \boldsymbol{M}_t \times \boldsymbol{Y}_t \to \bar{\mathbb{R}}$ as follows:

$$\mathbf{H}_t(\boldsymbol{m}, \boldsymbol{y}) \colon = \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}, \boldsymbol{d}_t, \boldsymbol{y}) + \boldsymbol{\lambda}_{t+1} \mathbf{F}_t^m(\boldsymbol{w}_t, \boldsymbol{d}_t, \boldsymbol{y}) + \boldsymbol{\mu}_t \mathrm{g}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}, \boldsymbol{y}) - \boldsymbol{\lambda}_t \boldsymbol{m}. \tag{3}$$

The function $\mathbf{H}_t$ is a discrete time equivalent of a deterministic Hamiltonian, generalized to a vector space.[1]

Claim E.1. *If* $\{\boldsymbol{x}_t, , \boldsymbol{d}_t, \boldsymbol{y}_t\}_{t=0}^T$ *solves* (**DV**) *and* $\exists$ *measurable* $\mathrm{y}_t \colon X_t \to Y_t$ *and* $\mathrm{d}_t \colon X_t \to D_t$ *s.t.* $\boldsymbol{y}_t = \mathrm{y}_t(\boldsymbol{x}_t)$ *and* $\boldsymbol{d}_t = \mathrm{d}_t(\boldsymbol{x}_t)$ *a.e., then* $\{\mathrm{y}_t, \mathrm{d}_t\}_{t=0}^T$ *solves* (**DP**)*.*

Proof. Let $\{\tilde{\mathrm{y}}_t, \tilde{\mathrm{d}}_t\}_{t=0}^T$ be any other feasible sequence of functions in the sense of (**DP**). Then, for each t, let $\tilde{\boldsymbol{y}}_t = \tilde{\mathrm{y}}_t(\boldsymbol{x}_t)$, $\tilde{\boldsymbol{d}}_t = \tilde{\mathrm{d}}_t(\boldsymbol{x}_t)$, and $\tilde{\boldsymbol{x}}_{t+1} = \mathrm{F}_t(\tilde{\boldsymbol{d}}_t, \tilde{\boldsymbol{y}}_t, \boldsymbol{w}_t)$ hold a.e. Since $\{\boldsymbol{x}_t, \boldsymbol{d}_t, \boldsymbol{y}_t\}_{t=0}^T$ solves (**DV**), we must have

$$\sum_{t=0}^T \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) \geq \sum_{t=0}^T \mathbf{u}_t(\boldsymbol{w}_t, \tilde{\boldsymbol{z}}_t, \tilde{\boldsymbol{m}}_t, \tilde{\boldsymbol{d}}_t, \tilde{\boldsymbol{y}}_t). \tag{4}$$

However, we also have that

$$\sum_{t=0}^T \mathbf{u}_t(\boldsymbol{w}_t, \tilde{\boldsymbol{z}}_t, \tilde{\boldsymbol{m}}_t, \tilde{\boldsymbol{d}}_t, \tilde{\boldsymbol{y}}_t) = \sum_{t=0}^T \mathbb{E}_0 \mathrm{u}_t(\boldsymbol{w}_t, \tilde{\boldsymbol{z}}_t, \tilde{\boldsymbol{m}}_t, \tilde{\mathrm{d}}_t(\tilde{\boldsymbol{x}}_t), \tilde{\mathrm{y}}_t(\tilde{\boldsymbol{x}}_t)), \tag{5}$$

and

$$\sum_{t=0}^T \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) = \sum_{t=0}^T \mathbb{E}_0 \mathrm{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \mathrm{d}_t(\boldsymbol{x}_t), \mathrm{y}_t(\boldsymbol{x}_t)). \tag{6}$$

To complete the proof, note that since $\{\tilde{\mathrm{y}}_t, \tilde{\mathrm{d}}_t\}_{t=0}^T$ was any other feasible sequence,

$$\mathbf{v}_0(\boldsymbol{x}_0) = \sum_{t=0}^T \mathbb{E}_0 \mathrm{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \mathrm{d}_t(\boldsymbol{x}_t), \mathrm{y}_t(\boldsymbol{x}_t)) \geq \sum_{t=0}^T \mathbb{E}_0 \mathrm{u}_t(\boldsymbol{w}_t, \tilde{\boldsymbol{z}}_t, \tilde{\boldsymbol{m}}_t, \tilde{\mathrm{d}}_t(\boldsymbol{x}_t), \tilde{\mathrm{y}}_t(\boldsymbol{x}_t)). \tag{7}$$

[1] A similar discrete time perturbation function can be found in Sorger (2015), where it is referred to as an 'M-function'. However, the M-function by Sorger (2015) is defined on Euclidean spaces, not on a general vector space.

□

Remark E.1. *A similar argument verifies that* $\mathbf{v}_0^{\vec{\boldsymbol{d}}_0}(\boldsymbol{x}_0) = \mathrm{v}_0^{\vec{\boldsymbol{d}}_0}(\boldsymbol{x}_0)$.

E.1 Sufficiency of the generalized Euler equation. We start with general sufficient conditions for a feasible sequence to solve (**CS-DV**).

Proposition E.1. *(**Sequential sufficiency**) If there exists a feasible sequence* $\{\boldsymbol{m}_t, \boldsymbol{y}_t\}_{t=0}^T$ *and multipliers* $\{\boldsymbol{\lambda}_t\}_{t=0}^T$ *and* $\{\boldsymbol{\mu}_t\}_{t=0}^T$ *such that*

(1) $\boldsymbol{m}_t \in \boldsymbol{M}_t$, $\boldsymbol{y}_t \in \boldsymbol{Y}_t$, $\boldsymbol{\lambda}_t \in L^2(\Omega, \mathbb{R}^{N_M}, \mathscr{F}_t)^\star$, *and* $\boldsymbol{\mu}_t \in L^2(\Omega, \mathbb{R}^{N_g}, \mathscr{F}_t)^\star$ *for each* t,
(2) $\boldsymbol{\lambda}_t \geq \boldsymbol{0}$ *and* $\boldsymbol{\mu}_t \geq \boldsymbol{0}$ *for each* t,
(3) $\boldsymbol{\mu}_t \mathrm{g}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) = 0$ *for each* t,
(4) $(\boldsymbol{m}_t, \boldsymbol{y}_t) \in \arg\max_{\boldsymbol{m},\boldsymbol{y}} \boldsymbol{H}_t(\boldsymbol{m}, \boldsymbol{y})$ *for each* t, *and*
(5) if $T = \infty$, *then* $\lim_{t\to\infty} \boldsymbol{\lambda}_{t+1}\boldsymbol{m}_{t+1} = 0$,

then $\{\boldsymbol{m}_t, \boldsymbol{y}_t\}_{t=0}^\infty$ *solves* (**CS-DV**).

Proof. Let conditions (1).- (5). of Proposition E.1 hold for a feasible sequence $\{\boldsymbol{m}_t, \boldsymbol{y}_t\}_{t=0}^T$, and let $\{\tilde{\boldsymbol{m}}_t, \tilde{\boldsymbol{y}}_t\}_{t=0}^T$ be a feasible sequence. Let $\Delta_{\bar{T}} := \sum_{t=0}^{\bar{T}} \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \sum_{t=0}^{\bar{T}} \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t)$ for any $\bar{T} \in \mathbb{N}$, with $\bar{T} \leq T$. If $T < \infty$, fix $\boldsymbol{\lambda}_{T+1} = \mathbf{0}$, $\boldsymbol{\lambda}_{T+1} \in L^2(\Omega, \mathbb{R}^{N_M}, \mathscr{F}_T)^\star$, and let $\boldsymbol{m}_{T+1}, \tilde{\boldsymbol{m}}_{T+1} \in \boldsymbol{M}_T$. We have the following:

$$
\begin{aligned}
(8)\qquad \Delta_{\bar{T}} \geq \sum_{t=0}^{\bar{T}} \Big[& \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t) \\
& + \boldsymbol{\lambda}_{t+1}\mathbf{F}_t^m(\boldsymbol{d}_t, \boldsymbol{y}_t, \boldsymbol{w}_t) - \boldsymbol{\lambda}_{t+1}\boldsymbol{m}_{t+1} - \boldsymbol{\lambda}_{t+1}\mathbf{F}_t^m(\boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t, \boldsymbol{w}_t) + \boldsymbol{\lambda}_{t+1}\tilde{\boldsymbol{m}}_{t+1} \\
& + \boldsymbol{\mu}_t \mathrm{g}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \boldsymbol{\mu}_t \mathrm{g}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t) \Big] \\
= \sum_{t=0}^{T} \Big[& \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t) \\
& + \boldsymbol{\lambda}_{t+1}\mathbf{F}_t^m(\boldsymbol{d}_t, \boldsymbol{y}_t, \boldsymbol{w}_t) - \boldsymbol{\lambda}_t\boldsymbol{m}_t - \boldsymbol{\lambda}_{t+1}\mathbf{F}_t^m(\boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t, \boldsymbol{w}_t) + \boldsymbol{\lambda}_t\tilde{\boldsymbol{m}}_t \\
& + \boldsymbol{\mu}_t \mathrm{g}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \boldsymbol{\mu}_t \mathrm{g}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t) \Big] \\
(9)\qquad + \quad & \boldsymbol{\lambda}_{T+1}(\tilde{\boldsymbol{m}}_{T+1} - \boldsymbol{m}_{T+1}) \\
(10)\qquad = \sum_{t=0}^{T} & \mathbf{H}_t(\boldsymbol{m}_t, \boldsymbol{y}_t) - \mathbf{H}_t(\tilde{\boldsymbol{m}}_t, \tilde{\boldsymbol{y}}_t) + \boldsymbol{\lambda}_{T+1}(\tilde{\boldsymbol{m}}_{T+1} - \boldsymbol{m}_{T+1}) \\
(11)\qquad \geq\ & \boldsymbol{\lambda}_{T+1}(\tilde{\boldsymbol{m}}_{T+1} - \boldsymbol{m}_{T+1}).
\end{aligned}
$$

Inequality (8) follows from $\boldsymbol{m}_{t+1} = \mathbf{F}_t^m(\boldsymbol{d}_t, \boldsymbol{y}_t, \boldsymbol{w}_{t+1})$, $\tilde{\boldsymbol{m}}_{t+1} = \mathbf{F}_t^m(\boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t, \boldsymbol{w}_{t+1})$, conditions 2. and 3. of Proposition E.1, and the fact that $\{\tilde{\boldsymbol{m}}_t, \tilde{\boldsymbol{y}}_t\}_{t=0}^T$ is feasible. Equality (9) follows from rearranging the RHS terms in the summation sign on the first line and the assumption $\boldsymbol{m}_0 = \tilde{\boldsymbol{m}}_0$. The second equality (10) follows from the definition of the function $\mathbf{H}_t$ by equation (3). The final inequality (11) follows by condition 4. of Proposition E.1.

Finally, if $T < \infty$, we have $\boldsymbol{\lambda}_{T+1}(\tilde{\boldsymbol{m}}_{T+1} - \boldsymbol{m}_{T+1}) = 0$, and thus $\Delta_T \geq 0$. If $T = \infty$, by condition (5). of Proposition E.1, we have

$$\sum_{t=0}^{\infty} \mathbf{u}_t(\boldsymbol{w}_t, z_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \sum_{t=0}^{\infty} \mathbf{u}_t(\boldsymbol{w}_t, z_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t) \geq \lim_{\bar{T}\to\infty} \boldsymbol{\lambda}_{\bar{T}}(\tilde{\boldsymbol{m}}_{\bar{T}+1} - \boldsymbol{m}_{\bar{T}+1}) \geq 0, \tag{12}$$

and so $\Delta_\infty \geq 0$, completing the proof. □

The connection of the S-function to the Hamiltonian is given by the following claim:

Claim E.2. *Let $\{\tilde{\mathrm{y}}_t, \tilde{\mathrm{d}}_t\}_{t=0}^T$ be a measurable sequence of feasible policy functions, and let $\{\boldsymbol{x}_t, \boldsymbol{d}_t, \boldsymbol{y}_t\}_{t=0}^T$ be generated by $\{\tilde{\mathrm{y}}_t, \tilde{\mathrm{d}}_t\}_{t=0}^T$. If $\boldsymbol{\lambda}_{t+1} = \lambda_t(\boldsymbol{x}_t)$ and $\boldsymbol{\mu}_t = \mu_t(\boldsymbol{x}_t)$ for each t, then*

$$\boldsymbol{H}_t(\boldsymbol{m}, \boldsymbol{y}) = \mathrm{I}_{\mathbf{S}(\cdot, \boldsymbol{\mu}_t, \boldsymbol{\lambda}_t \,|\, \mathbb{E}_t \boldsymbol{\lambda}_{t+1})}. \tag{13}$$

Proof. Recall the definition of the S-function

$$\begin{aligned}\mathbf{S}_t\,(m, y, \mu, \lambda \,|\, \mathbb{E}_t \boldsymbol{\lambda}_{t+1}) = \mathrm{u}_t(w_t, z_t, d_t, m) &+ \mathbb{E}_t \lambda_{t+1}(x_{t+1}) \mathrm{F}_t^m(d_t, y, w_{t+1}) \\ &- \lambda m \\ &+ \mu \mathrm{g}(w_t, z_t, m, y),\end{aligned} \tag{S}$$

where $\mathrm{y}_t(x_t)$, $w_{t+1})$, $w_{t+1} = \mathrm{F}_t^w(w_t, \eta_{t+1})$, $x_{t+1} = (w_{t+1}, m_{t+1}, z_{t+1})$, and $d_t = \mathrm{d}_t(x_t)$. Moreover, we have $x_t \in X_t \subset \mathbb{R}^{N_W + N_Z + N_M}$, and similarly $\lambda \in \mathbb{R}^{N_M}$ and $\mu \in \mathbb{R}^{N_g}$. Thus, the integral operator with kernel $\mathbf{S}_t(\cdot, \boldsymbol{\mu}_t, \boldsymbol{\lambda}_t \,|\, \mathbb{E}_t \boldsymbol{\lambda}_{t+1})$ can be defined by

$$\begin{aligned}\mathrm{I}_{\mathbf{S}(\cdot, \boldsymbol{\mu}_t, \boldsymbol{\lambda}_t \,|\, \mathbb{E}_t \boldsymbol{\lambda}_{t+1})}(\boldsymbol{m}, \boldsymbol{y}) &= \int \mathbf{S}_t(\boldsymbol{m}(\omega), \boldsymbol{y}(\omega), \boldsymbol{\mu}_t(\omega), \boldsymbol{\lambda}_t(\omega))\, \mathbb{P}(\mathrm{d}\omega) \\ &= \mathbf{u}_t(\boldsymbol{w}_t, z_t, \boldsymbol{d}_t, \boldsymbol{m}) + \boldsymbol{\lambda}_{t+1} \mathbf{F}_t^m(\boldsymbol{d}_t, \boldsymbol{y}, \boldsymbol{w}_{t+1}) - \boldsymbol{\lambda}_t \boldsymbol{m} + \boldsymbol{\mu}_t \mathbf{g}_t(\boldsymbol{w}_t, z_t, \boldsymbol{m}, \boldsymbol{y}) = \mathbf{H}_t(\boldsymbol{m}, \boldsymbol{y}),\end{aligned}$$

where the last equality follows from the definition of the S-function and by the definition of the Hilbert space problem in appendix A.2. □

We can now prove the first part of Proposition 3.1.

Proof of Proposition 3.1 Item (1). Let $\{\mathrm{y}_t, \mathrm{d}_t\}_{t=0}^T$ be a sequence of measurable feasible policy functions, and let $\{\boldsymbol{x}_t, \boldsymbol{d}_t, \boldsymbol{y}_t\}_{t=0}^T$ be generated by $\{\mathrm{y}_t, \mathrm{d}_t\}_{t=0}^T$. Moreover, let $\boldsymbol{\lambda}_{t+1} = \lambda_t(\boldsymbol{x}_t)$ and $\boldsymbol{\mu}_t = \mu_t(\boldsymbol{x}_t)$ for each t. Let conditions (i)-(iv) of Proposition 3.1, Part (1) (i.e., the proposition to be proved) hold. We will show that $\{\boldsymbol{m}_t, \boldsymbol{y}_t\}_{t=0}^T$ solves Problem (**CS-DV**) given $\vec{\boldsymbol{d}}_0 = \{\boldsymbol{d}_t\}_{t=0}^T$, and thus $\{\mathrm{y}_t\}_{t=0}^T$ solves (**FS-DP**) given $\vec{\boldsymbol{d}}_0$ by Claim E.1.

To do so, we check that $\{\boldsymbol{m}_t, \boldsymbol{y}_t\}_{t=0}^T$ and $\{\boldsymbol{\mu}_t, \boldsymbol{\lambda}_t\}_{t=0}^T$ satisfy conditions (1)-(5) of Proposition E.1. Conditions (1)-(3) follow immediately from the conditions of the proposition to be proved. For condition (4), we will show that $\{\boldsymbol{m}_t, \boldsymbol{y}_t\} \in \arg\max_{\boldsymbol{m},\boldsymbol{y}} \mathbf{H}_t(\boldsymbol{m},\boldsymbol{y})$ for each t. By Claim E.2, $\mathbf{H}_t(\boldsymbol{m},\boldsymbol{y}) = \mathrm{I}_{\mathbf{S}(\cdot,\boldsymbol{\mu}_t,\boldsymbol{\lambda}_t \,|\, \mathbb{E}_t\boldsymbol{\lambda}_{t+1})}$ and we can write

$$\begin{aligned}\mathbf{H}_t(\boldsymbol{m},\boldsymbol{y}) &= \mathrm{I}_{\mathbf{S}(\cdot,\boldsymbol{\mu}_t,\boldsymbol{\lambda}_t \,|\, \mathbb{E}_t\boldsymbol{\lambda}_{t+1})}(\boldsymbol{m},\boldsymbol{y})\\ &= \int \mathbf{S}_t(\boldsymbol{m}(\omega), \boldsymbol{y}(\omega), \boldsymbol{\mu}_t(\omega), \boldsymbol{\lambda}_t(\omega))\,\mathbb{P}(\mathrm{d}\omega).\end{aligned}$$

Thus, we will need to show that $\partial_{(\boldsymbol{m},\boldsymbol{y})}\mathbf{S}_t(\boldsymbol{m}_t, \boldsymbol{y}_t, \boldsymbol{\mu}_t, \boldsymbol{\lambda}_t \,|\, \mathbb{E}_t\boldsymbol{\lambda}_{t+1}) \in L^2\left(\Omega, \mathbb{R}^{N_M+N_Y}, \mathscr{F}_t\right)^\star$, and that $\partial_{(\boldsymbol{m},\boldsymbol{y})}\mathbf{S}_t(\boldsymbol{m}_t, \boldsymbol{y}_t, \boldsymbol{\mu}_t, \boldsymbol{\lambda}_t \,|\, \mathbb{E}_t\boldsymbol{\lambda}_{t+1}) = \mathbf{0}$. Square integrability is immediate from the definition of the S-function and condition (ii) of Proposition 3.1. Moreover,

$$\partial_{(\boldsymbol{m},\boldsymbol{y})}\mathbf{S}_t(\boldsymbol{m}_t, \boldsymbol{y}_t, \boldsymbol{\mu}_t, \boldsymbol{\lambda}_t \,|\, \mathbb{E}_t\boldsymbol{\lambda}_{t+1}) = \partial_{(m,y)}\mathbf{S}_t(\boldsymbol{m}_t, \mathrm{y}_t(\boldsymbol{x}_t), \mu_t(\boldsymbol{x}_t), \lambda_t(\boldsymbol{x}_t) \,|\, \mathbb{E}_t\boldsymbol{\lambda}_{t+1}).$$

Applying condition (i) of Proposition 3.1, $\mathbf{H}_t(\boldsymbol{m},\boldsymbol{y}) = \mathrm{I}_{\mathbf{S}(\cdot,\boldsymbol{\mu}_t,\boldsymbol{\lambda}_t \,|\, \mathbb{E}_t\boldsymbol{\lambda}_{t+1})}$ then follows from Claim 2. Finally, the transversality condition (5) of Proposition E.1, follows from condition (iii) of Proposition 3.1. □

Proof of Proposition 3.1 Item (2). Let $\{\mathrm{y}_t, \mathrm{d}_t\}_{t=0}^T$ be a sequence of measurable feasible policy functions. Assume the conditions for Proposition 3.1 Item (1). hold. In the proof, we will show that

$$\mathrm{v}_0(x_0) = \sum_{t=0}^{T} \mathbb{E}_0 \mathrm{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t), \qquad x_0 \in X_0, \tag{14}$$

holds for a stochastic recursive sequence $\{\boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t\}_{t=0}^T$ generated by $\{\mathrm{y}_t, \mathrm{d}_t\}_{t=0}^T$. We proceed by induction: first make the inductive assumption that for $t+1 < T$, $\{\mathrm{y}_t, \mathrm{d}_t\}_{j=t+1}^T$ satisfies

$$\mathrm{v}_{t+1}(x_{t+1}) = \max_{\{\mathrm{y}_j,\mathrm{d}_j\}_{j=t+1}^T} \sum_{j=t+1}^{T} \mathbb{E}_{t+1} \mathrm{u}_j(\boldsymbol{w}_j, \boldsymbol{z}_j, \boldsymbol{m}_j, \boldsymbol{d}_j, \boldsymbol{y}_j), \qquad x_{t+1} \in X_{t+1}, \tag{15}$$

for a stochastic recursive sequence $\{\boldsymbol{m}_j, \boldsymbol{d}_j, \boldsymbol{y}_j\}_{j=t+1}^T$ adapted to the filtration $\left\{\mathscr{F}_j\right\}_{j=t+1}^T$ and generated by $\{\mathrm{y}_j, \mathrm{d}_j\}_{j=t+1}^T$. Letting $t = T$, $\mathrm{v}_T(x_T) = \mathrm{u}_T(\boldsymbol{w}_T, \boldsymbol{z}_T, \boldsymbol{m}_T, \boldsymbol{d}_T, \boldsymbol{y}_T)$ will hold by the assumption of Proposition 3.1 Item (2). Next, fix $\boldsymbol{x}_t$ and note that given $\boldsymbol{d}_t$, we can let $x' = \mathrm{F}_t(d_t, y_t, w_{t+1})$. From the Bellman Principle of Optimality (Hernandez-Lerma and Lasserre (2012), Section 3), we have:

$$\max_{\vec{\boldsymbol{d}}'} \mathbb{E}_t \mathrm{v}_{t+1}^{\vec{\boldsymbol{d}}'}(x') \;=\; \mathbb{E}_t \max_{\{\boldsymbol{m}_j,\boldsymbol{d}_j,\boldsymbol{y}_j\}_{j=t+1}^T} \sum_{j=t+1}^{T} \mathbb{E}_{t+1} \mathrm{u}_j(\boldsymbol{w}_j, \boldsymbol{z}_j, \boldsymbol{m}_j, \boldsymbol{d}_j, \boldsymbol{y}_j) \;=\; \mathbb{E}_t \mathrm{v}_{t+1}(x'), \tag{16}$$

where maximization in the second term is over feasible stochastic recursive sequences with $\boldsymbol{x}_{t+1} = x'$. Moreover, $\mathrm{Q}_t^{\vec{\boldsymbol{d}}'}(x_t, d_t, y_t) = \mathrm{u}_t(w_t, z_t, m_t, d_t, y_t) + \mathbb{E}_t \mathrm{v}_{t+1}^{\vec{\boldsymbol{d}}'}(x')$ for any sequence $\vec{\boldsymbol{d}}'$. However, by the assumption of Proposition 3.1 Item (2)., we have

$$\mathrm{v}_t(x_t) = \max_{d_t, \vec{\boldsymbol{d}}', y_t} \mathrm{u}_t(w_t, z_t, m_t, d_t, y_t) + \mathbb{E}_t \mathrm{v}_{t+1}(x') = \mathrm{Q}_t^{\vec{\boldsymbol{d}}_{t+1}}(x_t, \mathrm{d}_t(x_t), \mathrm{y}_t(x_t)), \qquad x_t \in X_t,$$

where $\vec{\boldsymbol{d}}_{t+1} = \{\boldsymbol{d}_{t+1}, \ldots, \boldsymbol{d}_T\}$ is generated by $\{\mathrm{y}_j, \mathrm{d}_j\}_{j=t+1}^{T}$ and (x_t).

Now, the sequence of policy function $\{\mathrm{y}_{d_t,t}, \mathrm{y}_{t+1}, \ldots, \mathrm{y}_T\}$ solves (**FS-DP**) given $\vec{\boldsymbol{d}}_t = \{d_t, \boldsymbol{d}_{t+1}, \ldots\}$ starting at time t. As such, we have that $\mathrm{v}_t^{\vec{\boldsymbol{d}}_t}(x_t) = \mathrm{Q}_t^{\vec{\boldsymbol{d}}_t}(x_t, d_t, \mathrm{y}_t^{d_t}(x_t))$. However,

$$\mathrm{v}_t^{\vec{\boldsymbol{d}}_t}(x_t) = \max_y \mathrm{u}_t(w_t, z_t, m_t, d_t, y) + \mathbb{E}_t \mathrm{v}_{t+1}^{\vec{\boldsymbol{d}}}(x'), \tag{17}$$

thus, $\mathrm{v}_t(x_t) = \mathrm{u}_t(w_t, z_t, m_t, \mathrm{d}_t(x_t), \mathrm{y}_t(x_t)) + \mathbb{E}_t \mathrm{v}_{t+1}(x')$, where $x' = \mathrm{F}_t^m(\mathrm{d}_t(x_t), \mathrm{y}_t(x_t), w_{t+1})$. Noting (15), we then have

$$\mathrm{v}_t(x_t) = \sum_{t=0}^{T} \mathbb{E}_0 \mathrm{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t). \tag{18}$$

Thus, we have that $\mathrm{v}_t(x_t) = \sum_{t=0}^{T} \mathbb{E}_0 \mathrm{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t)$ for a stochastic recursive sequence $\{\boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t\}_{t=0}^{T}$ generated by $\{\mathrm{y}_t, \mathrm{d}_t\}_{t=0}^{T}$. By the principle of induction, (15) will hold for any T. $\square$

E.2 Value and shadow value function properties. The next proof is without loss of generality with $t = 0$. The proof for $t > 0$ is analogous by reformulating the initial period problem statements.

Proof of Corollary 3.1. We will show that $\partial_m \mathrm{u}_0(w_t, z_t, m_t, d_t, y_t) + \mu_0^\top \partial_m \mathrm{g}_0(w_t, z_t, m_t, d_t, y_t) \in \partial_m^S \mathrm{v}_0^{\vec{\boldsymbol{d}}_0}(w_0, z_0, m_0)$. Let $\left\{\boldsymbol{m}_t, \boldsymbol{y}_t\right\}_{t=0}^{T}$ be optimal given m_0, and let $\left\{\tilde{\boldsymbol{m}}_t, \tilde{\boldsymbol{y}}_t\right\}_{t=0}^{T}$ be optimal given any $\tilde{m}_0 \in M_0$. Consider $\Delta_{\bar{T}} := \sum_{t=0}^{\bar{T}} \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \sum_{t=0}^{\bar{T}} \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t)$ for any $\bar{T} \in \mathbb{N}$, with $\bar{T} \leq T$. If $T < \infty$, fix $\boldsymbol{\lambda}_{T+1} = \mathbf{0}$, $\boldsymbol{\lambda}_{T+1} \in L^2\left(\Omega, \mathbb{R}^{N_M}, \mathscr{F}_T\right)$, and let $\boldsymbol{m}_{T+1}, \tilde{\boldsymbol{m}}_{T+1} \in \boldsymbol{M}_T$. Following analogous steps as equations (8) - (11), and noting Claim E.1 and Remark E.1, we have

$$\begin{aligned}
\mathrm{v}_0^{\vec{\boldsymbol{d}}_0}(m_t) - \mathrm{v}_0^{\vec{\boldsymbol{d}}_0}(\tilde{m}_t) = \Delta_{\bar{T}} \geq & \sum_{t=0}^{\bar{T}} \Big[\mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t) \\
& \quad + \boldsymbol{\lambda}_{t+1}\mathbf{F}_t^m(\boldsymbol{d}_t, \boldsymbol{y}_t, \boldsymbol{w}_t) - \boldsymbol{\lambda}_{t+1}\boldsymbol{m}_{t+1} - \boldsymbol{\lambda}_{t+1}\mathbf{F}_t^m(\boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t, \boldsymbol{w}_t) + \boldsymbol{\lambda}_{t+1}\tilde{\boldsymbol{m}}_{t+1} \\
& \quad + \boldsymbol{\mu}_t \mathbf{g}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \boldsymbol{\mu}_t \mathbf{g}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t)\Big] \\
= & \sum_{t=0}^{T} \Big[\mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \mathbf{u}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t) \\
& \quad + \boldsymbol{\lambda}_{t+1}\mathbf{F}_t^m(\boldsymbol{d}_t, \boldsymbol{y}_t, \boldsymbol{w}_t) - \boldsymbol{\lambda}_t\boldsymbol{m}_t - \boldsymbol{\lambda}_{t+1}\mathbf{F}_t^m(\boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t, \boldsymbol{w}_t) + \boldsymbol{\lambda}_t\tilde{\boldsymbol{m}}_t \\
& \quad + \boldsymbol{\mu}_t \mathbf{g}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \boldsymbol{m}_t, \boldsymbol{d}_t, \boldsymbol{y}_t) - \boldsymbol{\mu}_t \mathbf{g}_t(\boldsymbol{w}_t, \boldsymbol{z}_t, \tilde{\boldsymbol{m}}_t, \boldsymbol{d}_t, \tilde{\boldsymbol{y}}_t)\Big] \\
& + \boldsymbol{\lambda}_{T+1}(\tilde{\boldsymbol{m}}_{T+1} - \boldsymbol{m}_{T+1}) \\
= & \sum_{t=0}^{T} \mathbf{H}_t(\boldsymbol{m}_t, \boldsymbol{y}_t) - \mathbf{H}_t(\tilde{\boldsymbol{m}}_t, \tilde{\boldsymbol{y}}_t) + \boldsymbol{\lambda}_{T+1}(\tilde{\boldsymbol{m}}_{T+1} - \boldsymbol{m}_{T+1}) \\
\geq & \ \mathbf{H}_0(\boldsymbol{m}_0, \boldsymbol{y}_0) - \mathbf{H}_0(\tilde{\boldsymbol{m}}_t, \tilde{\boldsymbol{y}}_0) + \boldsymbol{\lambda}_{T+1}(\tilde{\boldsymbol{m}}_{T+1} - \boldsymbol{m}_{T+1}).
\end{aligned}$$

Multiplying both sides by -1, we get (note that the w_0, z_0 arguments in the future choice-specific value functions are dropped for ease of notation)

$$\mathrm{v}_0^{\vec{d}_0}(\tilde{m}_t) - \mathrm{v}_0^{\vec{d}_0}(m_t) \leq \mathbf{H}_0(\tilde{\boldsymbol{m}}_0, \tilde{\boldsymbol{y}}_0) - \mathbf{H}_0(\boldsymbol{m}_0, \boldsymbol{y}_0) + \boldsymbol{\lambda}_{T+1}(\boldsymbol{m}_{T+1} - \tilde{\boldsymbol{m}}_{T+1}).$$

Now, since $\mathbf{H}_0$ is concave in $(m_0, \boldsymbol{y}_0)$, and $\partial_y \mathbf{S}_0(m_0, \boldsymbol{y}_0, \boldsymbol{\mu}_0, \boldsymbol{\lambda}_0 \,|\, \Xi_t) = 0$ a.e., recalling Claim E.2, we have that

$$\mathrm{v}_0^{\vec{d}_0}(\tilde{m}_t) - \mathrm{v}_0^{\vec{d}_0}(m_t) \leq \partial_m \mathbf{S}_0(m_0, \boldsymbol{y}_0)(\tilde{m}_0 - m_0) + \boldsymbol{\lambda}_{T+1}(\boldsymbol{m}_{T+1} - \tilde{\boldsymbol{m}}_{T+1}), \tag{19}$$

where we have set $\boldsymbol{\lambda}_0 = 0$. If $\boldsymbol{\lambda}_{T+1} = \mathbf{0}$, then $\mathrm{v}_0^{\vec{d}_0}(\tilde{m}_t) - \mathrm{v}_0^{\vec{d}_0}(m_t) \leq \partial_m \mathbf{S}_0(m_0, \boldsymbol{y}_0)(\tilde{m}_0 - m_0)$. Since $\tilde{m}_0$ is arbitrary, we have that $\partial_m \mathbf{S}_0(m_0, \boldsymbol{y}_0) \in \partial^S \mathrm{v}_0^{\vec{d}_0}(m_0)$. On the other hand, if $T = \infty$, applying the transversality condition (Item (1.iii). of Proposition 3.1) delivers the required conclusion. ◻

We can also show that the shadow value function is differentiable.

Lemma E.1. *If Assumption 2.1, 3.1-3.3 and 3.3- 4.1 hold, then $\lambda_t^{\vec{d}_t}$ and $\mu_t^{\vec{d}_t}$ are continuous in $\boldsymbol{m}_t$.*

Proof. For any arbitrary $j \geq t$, by Theorem 3.1,

$$\begin{aligned}\mu_j^{\vec{d}_t}(w_j, z_j, m_j, d_j, y_j) \\ = \partial_y \mathrm{g}_j^l(w_j, z_j, m_j, d_j, y_j)^{\top,-1} \left[\partial_y \mathrm{u}_t(w_j, z_j, m_j, d_j, y_j)\right] + \mathbb{E}_{x_j} \lambda_{j+1}^{\vec{d}_t}(x'),\end{aligned} \tag{20}$$

where $x' = \mathrm{F}_j^m\left(w_j, y_j, d_j\right)$, and $\mu_j^{\vec{d}_j}$ and $\lambda_{j+1}^{\vec{d}_t}$ are the multiplier functions furnished as the necessary conditions to problem (**CS-DV**). Fix ξ_j, then let $y_j = \mathrm{y}_j^{\vec{d}_{t+1}}(x_t)$. By Assumption 4.1, $\mathrm{y}_j^{\vec{d}_t}$ is continuous in $\boldsymbol{m}_j$. As such, by Assumption 3.1, $\mu_j^{\vec{d}_j}$ is continuous in $\boldsymbol{m}_j$ if $\lambda_{j+1}^{\vec{d}_j}$ is continuous in $\boldsymbol{m}_{j+1}$. If T is finite, then $\lambda_{T+1} = 0$, which allows us to arrive at the result by induction. If T is infinite, the transversality condition (Proposition 3.1, (1).(iii)) and iterating (20) using **L** gives the required result.

◻

Appendix Appendix F Supplementary appendix to Section 4

F.1 Additional invertibilility results.

Corollary F.1. *If $a \in \pi_A K_{l,t}$ and there exists $I = \bigtimes_{i=1}^{N_M}[\hat{m}_{i,t}^{\min}, \hat{m}_{i,t}^{\max}] \subset \bar{M}_t$ and $(\iota_1, \ldots, \iota_{N_M}) \in \{-1, 1\}^{N_M}$ such that for each $j \in \{1, \ldots, N_M\}$ and every $\bar{m}_{k,t} \in [\hat{m}_{k,t}^{\min}, \hat{m}_{k,t}^{\max}]$ with $k \neq j$*

$$\begin{aligned}\iota_j \bar{\mathbf{E}}_{a,t}\left(\bar{m}_1, \ldots, \bar{m}_{j-1}, \ldots, \hat{m}_j^{\min}, \bar{m}_{j+1}, \ldots \bar{m}_{N_M}, \mathrm{em}_{\tilde{A}}(\tilde{a})\right) \leq \\ \iota_j \bar{\mathbf{E}}_{a,t}\left(\bar{m}_1, \ldots, \bar{m}_{j-1}, \ldots, \hat{m}_j^{\max}, \bar{m}_{j+1}, \ldots \bar{m}_{N_M}, \mathrm{em}_{\tilde{A}}(\tilde{a})\right),\end{aligned} \tag{21}$$

then $\Theta_t^F(a)$ *is well-defined.*

Proof. See Theorem 3.1 in Mawhin (2013). □

The corollary below follows from Proposition 5.16 in Lee (2013).

Corollary F.2. *If Claim 4.2 holds, then* $\sigma_t(\bar{C})$ *is an* $N_{\bar{A}}$*-dimensional submanifold of* A_t. *Moreover, if* $N_{\bar{A}} < N_g + N_Y$, *then for each* $y, \mu \in \sigma_t(\bar{C})$, *there exists an open set* $\boldsymbol{O} \subset \mathbb{R}^{N_{\bar{A}}}$, *an open set* $\boldsymbol{U} \subset \mathbb{R}^{N_g+N_Y}$ *with* $y, \mu \in \boldsymbol{U}$, *and a mapping* $\mathrm{sm}\colon \mathbb{R}^{N_{\bar{A}}} \to \mathbb{R}^{N_g+N_Y-N_{\bar{A}}}$ *such that*

$$Y_t \times \Upsilon_t \cap U = \Big\{ y, \mu \,\Big|\, \Big(y_{\iota_y(n_1+1)}, \dots, y_{\iota_y(N_Y)}, \mu_{\iota_\mu(n_2+1)}, \dots, \mu_{\iota_\mu(n_{N_g})} \Big) = \mathrm{sm}\Big(y_{\iota_y(1)}, \dots, y_{\iota_y(n_1)}, \mu_{\iota_\mu(1)}, \dots, \mu_{\iota_\mu(n_2)} \Big) \Big\}, \qquad n_1 + n_2 = N_{\bar{A}}$$

where ι_μ *and* ι_y *are permutations of the indices of the multipliers and post-states, respectively.*

F.2 Additional exogenous grid results.

Remark F.1. *By Item (1). of Proposition 3.1, any* $v_{t,i}^{\#} \in \mathcal{V}_t^F$ *will correspond to a sequence of future discrete choices* $\mathbf{d}_{t+1}(x'_{t+1,i})$ *that are optimal at time* $t+1$. *In particular,* $v_{t,i}^{\#} = \mathrm{u}_t(x_{t,i}^{\#}, d_t, y_{t,i}^{\#}) + \mathbb{E}_{x_t^{\#}} \mathrm{v}_{t+1}^{\vec{\boldsymbol{d}}_{t+1}}(x'_{t+1,i}) = \mathrm{v}_t^{\vec{\boldsymbol{d}}_t}(x_{t,i}^{\#})$, *where* $\vec{\boldsymbol{d}}_t = [d_t, \mathbf{d}_{t+1}(y_{t+1,i}^{\#})]$, $x_{t,i}^{\#} = (w_t, z_t, \bar{m}_{t,i}^{\#})$, *and* $x'_{t+1,i} = \mathrm{F}_t(d_t, y_{t,i}^{\#}, w_{t+1})$.

Appendix Appendix G Supplementary appendix to Section 5

G.1 Intersection points using RFC. We can use the first order information in the RFC algorithm to easily add an approximation of the intersection points of two different future choice-specific value functions. In Figure 1, the point $\times$ is an approximation of the intersection of the future choice-specific value functions. When a neighboring point l lies above a tangent plane to the future choice-specific value function at j, the intersection point can be found by finding the intersection of the two lines $p \mapsto (m_{t,j}^{\#}, v_{t,j}^{\#}) + p \circ \nabla v_{t,j}^{\#}(m_{t,l}^{\#} - m_{t,j}^{\#})$ and $p \mapsto (m_{t,l}^{\#}, v_{t,l}^{\#}) + p \circ \nabla v_{t,l}^{\#}(m_{t,j}^{\#} - m_{t,l}^{\#})$.

We can also add the optimal policy at the intersection points. In particular, suppose $m_{t,\times}^{\#}$ is an approximated intersection point between $m_{t,j}^{\#}$ and $m_{t,l}^{\#}$. We can perform a nearest neighbor search around the point $m_{t,j}^{\#}$ in the original grid $\mathcal{A}_t$, and find neighboring points $\left\{ m_{t,i}^{\#}, y_{t,i}^{\#} \right\}_i$ such that the observable values do not 'jump' from the point j. An interpolant can be constructed using these neighboring points to locally approximate the optimal policy given the optimal future discrete choice implied by $m_{t,j}^{\#}$. This interpolant can be evaluated at $m_{t,\times}^{\#}$ to produce an approximate policy $y_{t,\times}^{\#}$. To ensure jumps in the policy are approximated, an approximate policy interpolated using neighbors to the policy at $m_{t,l}^{\#}$ can also be attached at an arbitrarily small distance from $m_{t,\times}^{\#}$. We note that in applications with many states, it may be more practicable to approximate policies at the intersection by nearest neighbor interpolation, a procedure that performs well in our Section 6.1 application.

G.2 Local topology of future discrete choices. In what follows, Theorem 4.1 or Claim 4.2 conditions will be in force. Let $\bar{M}_t^{\vec{\boldsymbol{d}}_{t+1}}$ denote the set of points in $\bar{M}_t$ that 'lead to' the future sequence of discrete choices $\vec{\boldsymbol{d}}_{t+1}$, defined by

$$\bar{M}^{\vec{\boldsymbol{d}}_{t+1}} :\; = \left\{ \bar{m}_t \in \bar{M}_t \,\middle|\, \vec{\boldsymbol{d}}_{t+1} = \arg\max_{\vec{\boldsymbol{d}}'} \bar{\mathrm{v}}_{d_t,t}^{\vec{\boldsymbol{d}}'}(\bar{m}_t) \right\}, \tag{22}$$

where $\vec{\boldsymbol{d}}'$ is feasible, and $\bar{\mathrm{v}}_{d_t,t}^{\vec{\boldsymbol{d}}'}$ is the value function conditional on current choice d_t and future sequence $\vec{\boldsymbol{d}}'$. For the following, let leb be the Lebesgue measure on the Borel sets of $\bar{M}_t$.

Lemma G.1. *There exists $U = \{U_i\}_{i\in\mathcal{F}}$ such that (i) U_i is open and $\mathrm{leb}(U) = \mathrm{leb}(\bar{M}_t)$, and (ii) $\bar{\mathbf{d}}_t$ is constant on U_i.*

Proof. Let $\bar{\mathbf{d}}_t$ be the function mapping the active state to an optimal sequence of future discrete choices. For any open set $O \in \bar{M}_t$, there exists a compact set C with a non-empty interior such that $O \subset C$. Since $\lambda_t(\bar{m}_t) \in \partial^S \bar{\mathrm{v}}_t^{\vec{\boldsymbol{d}}}(\bar{m}_t)$ and λ_t is continuous (see Lemma E.1), $\partial_{\bar{M}} \bar{\mathrm{v}}_t^{\vec{\boldsymbol{d}}}$ is bounded on O for each $\vec{\boldsymbol{d}}$. Next, the number of intersections between any two $\bar{\mathrm{v}}_t^{\vec{\boldsymbol{d}}_j}$ and $\bar{\mathrm{v}}_t^{\vec{\boldsymbol{d}}_k}$ can at most be finite, implying there are at most countably many intersection points in O that are singletons. Thus, there must exist a collection of open sets $V = \{V_i\}_{i\in\mathcal{F}} \subset O$ such that $\bar{\mathbf{d}}_t$ is constant on V_i and $\mathrm{leb}(V) = \mathrm{leb}(O)$. □

Appendix Appendix H Supplementary appendix to Section 6

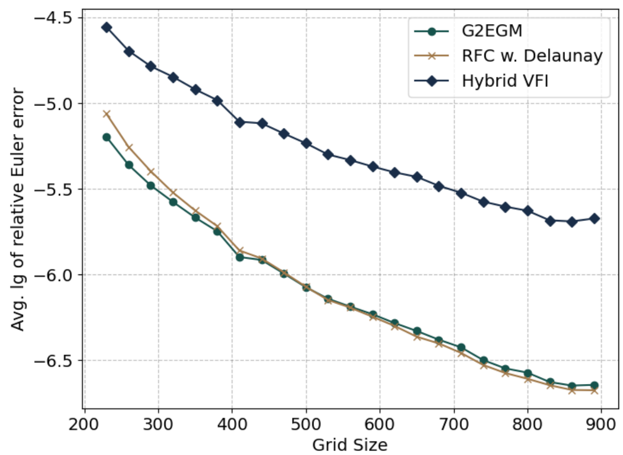

Figure H.1. Euler error benchmarking for Application 1

H.1 Computation for Section 6.1.

H.2 Computation for Section 6.2. Once we find the time t solution (by following the steps in Box 2 below), let $\mathrm{y}_{f,d,t}$ and $\mathrm{y}_{h,d,t}$ denote the optimal choice-specific policy functions conditional on the time t discrete choice, with $d \in \{0,1\}$. Given the optimal policy functions, the iteration of the shadow values is given by $\lambda_{h,t}(x_t) = \mathrm{d}_t(x_t)\partial_c\varphi^u(c_t, y_{h,t}) + \beta\mathbb{E}_{x_t}(1-\mathrm{d}_t(x_t))\left[\mathbb{E}_{x_t}\lambda_{h,t+1}(x_{t+1}) + \partial_{y_h}\varphi^u(c_t, y_{h,t})\right]$ and $\lambda_{a,t}(x_t) = \partial_{y_h}\varphi^u(c_t, y_{h,t})$, where $x_t = (w_t, m_{f,t}, m_{h,t})$, $c_t = m_{f,t} + \mathrm{d}_t(x_t)m_{h,t} - \mathrm{d}_t(x_t)\mathrm{y}_{a,\mathrm{d}_t(x_t),t}(x_t) - \mathrm{y}_{h,\mathrm{d}_t(x_t),t}(x_t)(1+\tau)$, and $y_{h,t} = \mathrm{y}_{h,\mathrm{d}_t(x_t),t}(x_t)$. For the adjuster's pseudo-code, recall that

$$c_t = \partial_c\varphi^{u,-1}((1+r)\beta\lambda_{f,t+1}(m_{f,t+1}, m_{h,t+1}), y_{h,t}). \tag{23}$$

We apply these steps below for each w_t in the exogenous shock grid. With $\mathrm{y}_{h,1,t}$, $\mathrm{y}_{f,1,t}$ approximated on a uniform grid, we can then construct $\mathrm{y}_{h,t}$ and $\mathrm{y}_{f,t}$ in a standard way by solving the current period discrete choice.

Box 2: Exogenous grid and RFC for the adjuster

(1) Fix ξ_t and a uniform grid over values of the exogenous variable $y_{h,t}$, $\mathcal{Y}_{h,t}$. Initialise $\mathcal{A}_t$,,$\mathcal{V}_t$), and $\mathcal{M}_t$) as empty arrays.

(2) For each $y^{\#}_{h,t,i} \in \mathcal{Y}_{h,t}$:

(i) Evaluate P roots of (25) in terms of $y_{f,t}$, with c_t given by (23), and $y_{h,t}$ fixed as $y^{\#}_{h,t,i}$. Collect the roots in a tuple $(y^{\#}_{f,t,i_0}, \dots, y^{\#}_{f,t,i_j}, y^{\#}_{f,t,i_P})$. Append $((y^{\#}_{f,t,i_0}, y^{\#}_{h,t,i}), \dots, (y^{\#}_{f,t,i_P}, y^{\#}_{h,t,i}))$ to $\mathcal{A}_t$.

(ii) For each root in $(y^{\#}_{f,t,0}, \dots, y^{\#}_{f,t,i_j}, \dots, y^{\#}_{f,t,i_P})$, evaluate the endogenous grid points $(\bar{m}^{\#}_{t,0}, \dots, \bar{m}^{\#}_{t,j}, \dots, \bar{m}^{\#}_{t,P})$ using the budget constraint

$$\bar{m}^{\#}_{t,i_j} = y^{\#}_{f,t,i_j} + (1+\tau) y^{\#}_{h,t,i} + c_t,$$

and evaluate $(v^{\#}_{t,0}, \dots, v^{\#}_{t,i_j}, \dots, v^{\#}_{t,i_P})$ as

$$v^{\#}_{t,i_j} = u(c_t, y^{\#}_{h,t,i}) + \beta \mathbb{E}_{w_t} \mathrm{v}_{t+1}(w_{t+1}, (1+r) y^{\#}_{f,t,i_j}, y^{\#}_{h,t,i}).$$

(iii) Append $(\bar{m}^{\#}_{t,0}, \dots, \bar{m}^{\#}_{t,j}, \dots, \bar{m}^{\#}_{t,P})$ to $\mathcal{M}_t$, and $(v^{\#}_{t,0}, \dots, v^{\#}_{t,i_j}, \dots, v^{\#}_{t,i_P})$ to $\mathcal{V}_t$.

(3) Apply RFC to $\mathcal{V}_t$, $\mathcal{A}_t$, and $\mathcal{M}_t$ and recover $\mathcal{M}^{\mathrm{RFC}}_t$, $\mathcal{V}^{\mathrm{RFC}}_t$, and $\mathcal{Y}^{\mathrm{RFC}}_t$.

References

Aliprantis, C. D. and Border, K. C. (2006). *Infinite dimensional analysis: A hitchhiker's guide*. Springer-Verlag, Berlin, 3 edition. 1

Bauschke, H. H. and Combettes, P. L. (2010). *Convex analysis and monotone operator theory in Hilbert spaces*. Canadian Mathematical Society, Springer. 2, 4

Hernandez-Lerma, O. and Lasserre, J. (2012). *Discrete-time Markov control processes: Basic optimality criteria*. Stochastic modelling and applied probability. Springer New York. 7

Lee, J. M. (2013). *Introduction to smooth manifolds*. Graduate Texts in Mathematics. Springer. 10

Luenberger, D. (1997). *Optimization by vector space methods*. Professional Series. Wiley. 1

Mawhin, J. (2013). Variations on Poincaré-Miranda's theorem. *Advanced Nonlinear Studies*, 13(1):209–217. 10

Sorger, G. (2015). *Dynamic economic analysis: Deterministic models in discrete time*. Thesis. Cambridge University Press. 4