Strong Contraction Property of Euler Mapping in Dynamic Discrete Choice Models with Finite Dependence

Victor Aguirregabiria *
University of Toronto, CEPR

Arvind Magesan †
University of Calgary

September 15, 2025

This note is an excerpt from the paper *"Solving Discrete Choice Dynamic Programming Models Using Euler Equations"* by the same authors. In discrete choice Markov Decision Processes that satisfy a finite dependence property, we define an Euler fixed-point operator that solves the model and prove that it is a contraction mapping with modulus δ strictly smaller than the model's discount factor β. This result implies that the Euler mapping not only ensures convergence but does so more efficiently than the standard value function operator.

1 Model

1.1 Framework

We consider a stationary Markov Decision Process (MDP) with a discrete action space, following Rust (1994). In each discrete time period t, an agent selects an action a_t to maximize expected intertemporal utility:[1]

$$\mathbb{E}_t \left[\sum_{j=0}^{\infty} \beta^j \, \Pi_{a_{t+j}}(\mathbf{s}_{t+j}) \right],$$

where $\beta \in (0,1)$ is the discount factor, $\Pi_{a_t}(\mathbf{s}_t)$ denotes the per-period payoff function, and $\mathbf{s}_t \in \mathcal{S}$ is the state vector, which evolves according to a controlled Markov process with transition

*Department of Economics, University of Toronto. 150 St. George Street, Toronto, ON, M5S 3G7, Canada, victor.aguirregabiria@utoronto.ca.

†Department of Economics, University of Calgary. 2500 University Drive, N.W Calgary, AB, Canada, arvindmagesan@gmail.com.

[1] In this paper, we adopt the convention of denoting actions as subscripts and states as arguments enclosed in parentheses for all functions that depend on actions or states, e.g., $\Pi_a(\mathbf{s})$.

probability function $\Pr(\mathbf{s}_{t+1} \mid a_t = a, \mathbf{s}_t) = f_a(\mathbf{s}_{t+1} \mid \mathbf{s}_t)$. The action a_t is drawn from a finite set $\mathcal{A} = \{0, 1, \ldots, J\}$.

The value function $V(\mathbf{s}_t)$ satisfies the Bellman equation:

$$V(\mathbf{s}_t) = \max_{a_t \in \mathcal{A}} \left\{ \Pi_{a_t}(\mathbf{s}_t) + \beta \int V(\mathbf{s}_{t+1})\, f_{a_t}(\mathbf{s}_{t+1} \mid \mathbf{s}_t)\, d\mathbf{s}_{t+1} \right\}. \tag{1}$$

The optimal decision rule $\alpha(\mathbf{s}_t) : \mathcal{S} \to \mathcal{A}$ selects the action that maximizes the right-hand side of equation (1).

Following the standard dynamic discrete choice literature (e.g., Rust, 1994), we partition the state vector as $\mathbf{s}_t = (\mathbf{x}_t, \boldsymbol{\varepsilon}_t)$, where $\mathbf{x}_t \in \mathcal{X}$ is observed by the researcher, and $\boldsymbol{\varepsilon}_t = \{\varepsilon_{a,t} : a \in \mathcal{A}\}$ is a vector of action-specific unobserved shocks. We impose the standard assumptions of *additive separability* and *conditional independence*. Additive separability implies that the payoff function decomposes as

$$\Pi_{a_t}(\mathbf{s}_t) \;=\; \pi_{a_t}(\mathbf{x}_t) + \varepsilon_{a_t,t}, \tag{2}$$

where $\pi_{a_t}(\mathbf{x}_t)$ is the deterministic component of utility. Conditional independence means that the transition density factors as $f_{a_t}(\mathbf{s}_{t+1} \mid \mathbf{s}_t) = f_{a_t}(\mathbf{x}_{t+1} \mid \mathbf{x}_t) \cdot g(\boldsymbol{\varepsilon}_{t+1})$, where $g(\cdot)$ is the density of the unobserved shocks $\boldsymbol{\varepsilon}_t$, assumed to be absolutely continuous, continuously differentiable, and with finite moments.

Let $V^{\sigma}(\mathbf{x}_t)$ denote the *integrated value function*, obtained by integrating the value function over the distribution of unobservables: $V^{\sigma}(\mathbf{x}_t) \equiv \int V(\mathbf{x}_t, \boldsymbol{\varepsilon}_t)\, g(\boldsymbol{\varepsilon}_t)\, d\boldsymbol{\varepsilon}_t.$. The *integrated Bellman equation* is then given by:

$$V^{\sigma}(\mathbf{x}_t) = \int \max_{a \in \mathcal{A}} \{ v_a(\mathbf{x}_t) + \varepsilon_{a,t} \}\; g(\boldsymbol{\varepsilon}_t)\, d\boldsymbol{\varepsilon}_t, \tag{3}$$

where $v_a(\mathbf{x}_t)$ denotes the *choice-specific value function*, defined as:

$$v_a(\mathbf{x}_t) \equiv \pi_a(\mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1} \in \mathcal{X}} V^{\sigma}(\mathbf{x}_{t+1})\, f_a(\mathbf{x}_{t+1} \mid \mathbf{x}_t). \tag{4}$$

The integrated Bellman equation defines a fixed-point mapping in the space of value functions. This mapping is the *value function (VF) operator*:

$$\mathbf{V}^{\sigma} \;=\; \Gamma_{VF}\left(\mathbf{V}^{\sigma}\right) \tag{5}$$

where $\mathbf{V}^{\sigma} = \{V^{\sigma}(\mathbf{x}_t) : \mathbf{x}_t \in \mathcal{X}\}$. This mapping is a contraction with Lipschitz constant equal to the discount factor β (see Puterman, 1994, and Rust, 1996).

For some of the results presented below, it is useful to express the right-hand side of the integrated Bellman equation using the more compact notation of *McFadden's social surplus function*. Given a vector of choice-specific values $\mathbf{v} = \{v_a : a \in \mathcal{A}\} \in \mathbb{R}^{J+1}$, the social surplus is defined as:

$$S(\mathbf{v}) \equiv \int \max_{a\in\mathcal{A}} \{v_a + \varepsilon_{a,t}\} \; g(\boldsymbol{\varepsilon}_t)\, d\boldsymbol{\varepsilon}_t. \tag{6}$$

Using this notation, the integrated Bellman equation can be written as:

$$V^\sigma(\mathbf{x}_t) \;=\; S(\mathbf{v}(\mathbf{x}_t)) \tag{7}$$

where $\mathbf{v}(\mathbf{x}_t) = \{v_a(\mathbf{x}_t) : a \in \mathcal{A}\}$ denotes the vector of choice-specific values functions.

1.2 Finite dependence and Euler mapping

Define the value difference $\widetilde{v}_a(\mathbf{x}_t) \equiv v_a(\mathbf{x}_t) - v_0(\mathbf{x}_t)$, where alternative 0 is chosen as the baseline without loss of generality. Let $\widetilde{\mathbf{v}}(\mathbf{x}_t) = \{\widetilde{v}_a(\mathbf{x}_t) : a \in \mathcal{A} \setminus \{0\}\}$ denote the vector of value differences for the remaining J alternatives. In the rest of this subsection, we derive a fixed-point mapping for the vector $\widetilde{\mathbf{v}}(\mathbf{x}_t)$. We refer to this mapping as the *Euler mapping*, as it is closely related to the Euler equations that characterize optimality in dynamic programming problems (see Aguirregabiria and Magesan, 2013).

Using value differences, we can rewrite equation (7), which relates the value function and the social surplus function, as follows:

$$V^\sigma(\mathbf{x}_t) \;=\; v_0(\mathbf{x}_t) + S([0, \widetilde{\mathbf{v}}(\mathbf{x}_t)]) \tag{8}$$

where $[0, \widetilde{\mathbf{v}}]$ denotes the vector of choice-specific values with alternative 0 normalized to zero.

Combining the definition of value difference $\widetilde{v}_a(\mathbf{x}_t)$ with equations (4) and (8), we have:

$$\widetilde{v}_a(\mathbf{x}_t) = \widetilde{\pi}_a(\mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1}} \Big[v_0(\mathbf{x}_{t+1}) + S([0, \widetilde{\mathbf{v}}(\mathbf{x}_{t+1})]) \Big] \Big[f_a\left(\mathbf{x}_{t+1} \mid \mathbf{x}_t\right) - f_0\left(\mathbf{x}_{t+1} \mid \mathbf{x}_t\right) \Big] \tag{9}$$

with $\widetilde{\pi}_a(\mathbf{x}_t) \equiv \pi_a(\mathbf{x}_t) - \pi_0(\mathbf{x}_t)$. Replacing $v_0(\mathbf{x}_{t+1})$ with its expression in equation (4):

$$\begin{aligned} \widetilde{v}_a(\mathbf{x}_t) \;&=\; \widetilde{\pi}_a(\mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1}} \Big[\pi_0(\mathbf{x}_{t+1}) + S([0, \widetilde{\mathbf{v}}(\mathbf{x}_{t+1})]) \Big] \Big[f_a\left(\mathbf{x}_{t+1} \mid \mathbf{x}_t\right) - f_0\left(\mathbf{x}_{t+1} \mid \mathbf{x}_t\right) \Big] \\ &+\; \beta^2 \sum_{\mathbf{x}_{t+1}} \sum_{\mathbf{x}_{t+2}} V^\sigma(\mathbf{x}_{t+2})\, f_0\left(\mathbf{x}_{t+2} \mid \mathbf{x}_{t+1}\right) \Big[f_a\left(\mathbf{x}_{t+1} \mid \mathbf{x}_t\right) - f_0\left(\mathbf{x}_{t+1} \mid \mathbf{x}_t\right) \Big] \end{aligned} \tag{10}$$

We consider models with the following *Two-period Finite Dependence* property.

DEFINITION. **Two-Period Finite Dependence.** [Arcidiacono and Miller (2011, 2019)] *For arbitrary choices a_t and a_{t+1}, define the two-periods-forward transition probability of the state variables:*

$$f^{(2)}_{a_t,a_{t+1}}(\mathbf{x}_{t+2} \mid \mathbf{x}_t) \equiv \sum_{\mathbf{x}_{t+1}} f_{a_{t+1}}(\mathbf{x}_{t+2} \mid \mathbf{x}_{t+1})\ f_{a_t}(\mathbf{x}_{t+1} \mid \mathbf{x}_t) \tag{11}$$

A controlled transition probability $f_{a_t}(\mathbf{x}_{t+1} \mid \mathbf{x}_t)$ has the (global) two-period finite dependence property if for every state $\mathbf{x}_t \in \mathcal{X}$ and choice $a \in \mathcal{A}$ we have:

$$f^{(2)}_{a_t=a,a_{t+1}=0}(\mathbf{x}_{t+2} \mid \mathbf{x}_t) = f^{(2)}_{a_t=0,a_{t+1}=0}(\mathbf{x}_{t+2} \mid \mathbf{x}_t), \text{ for any } \mathbf{x}_{t+2} \in \mathcal{X} \quad \blacksquare \tag{12}$$

EXAMPLE. **Multi-Armed Bandit Models**. A multi-armed bandit model is a discrete choice MDP in which the only endogenous state variable is the decision made in the previous period. Formally, the state vector is given by $\mathbf{x}_t = (y_t, \mathbf{z}_t)$, where $y_t = a_{t-1}$, and $\mathbf{z}_t$ is a vector of exogenous state variables with transition density $f_z(\mathbf{z}_{t+1} \mid \mathbf{z}_t)$. This class of models encompasses a wide range of important economic applications, including market entry and exit, demand with switching costs, occupational choice, migration, and store location decisions.

In multi-armed bandit models, the distribution of $\mathbf{x}_{t+2}$ conditional on $(\mathbf{x}_t, a_t, a_{t+1})$ does not depend on a_t. That is,

$$f^{(2)}_{a_t=a,a_{t+1}=a'}(\mathbf{x}_{t+2} \mid \mathbf{x}_t) = 1\{y_{t+2} = a'\} \sum_{\mathbf{z}_{t+1}} f_z(\mathbf{z}_{t+2} \mid \mathbf{z}_{t+1})\ f_z(\mathbf{z}_{t+1} \mid \mathbf{z}_t) \tag{13}$$

It is straightforward to verify that this equation implies that this class of models exhibits a global two-period finite dependence property. $\blacksquare$

Under the two-period finite dependence property, the last term on the right-hand side of equation (10) equals zero. As a result, we can rewrite that equation as the following mapping from value differences to value differences:

$$\widetilde{v}_a(\mathbf{x}_t) = c_a(\mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1}} S([0, \widetilde{\mathbf{v}}(\mathbf{x}_{t+1})]) \Big[f_a(\mathbf{x}_{t+1} \mid \mathbf{x}_t) - f_0(\mathbf{x}_{t+1} \mid \mathbf{x}_t) \Big] \tag{14}$$

where function $c_a(\mathbf{x}_t)$ is defined as $\widetilde{\pi}_a(\mathbf{x}_t) + \beta \sum_{\mathbf{x}_{t+1}} \pi_0(\mathbf{x}_{t+1}) \Big[f_a(\mathbf{x}_{t+1} \mid \mathbf{x}_t) - f_0(\mathbf{x}_{t+1} \mid \mathbf{x}_t) \Big]$. This function depends only on the model's primitives—specifically, the per-period payoff function and the transition probabilities—and is independent of any endogenous or equilibrium objects.

Equation (14) defines a fixed-point mapping in the space of value differences. This mapping is the *Euler mapping (EE)*:

$$\widetilde{\mathbf{v}} \;=\; \Gamma_{EE}\left(\widetilde{\mathbf{v}}\right) \tag{15}$$

where $\widetilde{\mathbf{v}} = \{v_a(\mathbf{x}_t) : a \in \mathcal{A} \setminus \{0\}, \mathbf{x}_t \in \mathcal{X}\}$.

The vector of value differences $\widetilde{\mathbf{v}}$ does not provide the complete solution to the dynamic programming problem. We still need to obtain the values for baseline choice 0, $v_0(\mathbf{x}_t)$, or equivalently, the value function $V^\sigma(\mathbf{x}_t)$. Given the vector $\widetilde{\mathbf{v}}_a$ for $a \neq 0$, we can obtain the vector of values $\mathbf{V}^\sigma$ by simply solving a system of linear equations. Specifically, given equation (4), we can write the following system of equations for the vector of value differences $\widetilde{\mathbf{v}}_a$:

$$\widetilde{\mathbf{v}}_a \;=\; \widetilde{\pi}_a + \beta\left[\mathbf{F}_a - \mathbf{F}_0\right]\mathbf{V}^\sigma \tag{16}$$

Under two-period finite dependence, matrix $\mathbf{F}_a - \mathbf{F}_0$ is non-singular (Aguirregabiria and Magesan (2013)), so this linear system implies a unique solution for the vector of values $\mathbf{V}^\sigma$.

EXAMPLE. In a **multi-armed bandit model**, the state vector is given by $\mathbf{x}_t = (y_t, \mathbf{z}_t)$ with $y_t = a_{t-1}$. In this setting, the Euler fixed-point mapping described generally in equation (14) takes the following specific form:

$$\widetilde{v}_a(y_t, \mathbf{z}_t) \;=\; c_a(y_t, \mathbf{z}_t) + \beta \sum_{\mathbf{z}_{t+1}} \Big[S([0, \widetilde{\mathbf{v}}(a, \mathbf{z}_{t+1})]) - S([0, \widetilde{\mathbf{v}}(0, \mathbf{z}_{t+1})])\Big]\, f_z\left(\mathbf{z}_{t+1} \mid \mathbf{z}_t\right) \tag{17}$$

Function $c_a(y_t, \mathbf{z}_t)$ is given by $\widetilde{\pi}_a(y_t, \mathbf{z}_t) + \beta \sum_{\mathbf{z}_{t+1}} \Big[\pi_0(a, \mathbf{z}_{t+1}) - \pi_0(0, \mathbf{z}_{t+1})\Big] f_z\left(\mathbf{z}_{t+1} \mid \mathbf{z}_t\right)$. The associated system of linear equations in (16) becomes:

$$\widetilde{\mathbf{v}}_a \;=\; \widetilde{\pi}_a + \beta\left[(\mathbf{I}_a - \mathbf{I}_0) \otimes \mathbf{F}_z\right]\mathbf{V}^\sigma \tag{18}$$

where, for any $a \in \mathcal{A}$, $\mathbf{I}_a$ is a $(J+1) \times (J+1)$ matrix with a column of ones in the $(a+1)$-th position and zeros elsewhere. ■

PROPOSITION 1. *Consider a discrete choice MDP characterized by equations (3) and (4), and assume the model satisfies the two-period finite dependence property. Let $\widetilde{\mathbf{v}}_{EE}$ denote a fixed-point of the Euler operator Γ_{EE}, and let $\mathbf{V}^\sigma_{EE}$ be the associated value vector that solves the system of linear equations (16). Likewise, let $\mathbf{V}^\sigma_{VF}$ be a fixed-point of the value function operator Γ_{VF}, and let $\widetilde{\mathbf{v}}_{VF}$ be the corresponding vector of value differences. Then: $\widetilde{\mathbf{v}}_{EE} = \widetilde{\mathbf{v}}_{VF}$ and $\mathbf{V}^\sigma_{EE} = \mathbf{V}^\sigma_{VF}$.* ■

Proof of Proposition 1. By construction, $\mathbf{V}^\sigma_{EE}$ satisfies the Bellman equation and is therefore a fixed-point of the value function operator Γ_{VF}. Since Γ_{VF} is a contraction mapping with a unique fixed-point, it follows that $\mathbf{V}^{\sigma}_{EE} = \mathbf{V}^{\sigma}_{VF}$. By definition, there exists a unique vector of value differences $\widetilde{\mathbf{v}}_{EE}$ associated with $\mathbf{V}^{\sigma}_{EE}$, and similarly a unique vector $\widetilde{\mathbf{v}}_{VF}$ associated with $\mathbf{V}^{\sigma}_{VF}$. Hence, $\widetilde{\mathbf{v}}_{EE} = \widetilde{\mathbf{v}}_{VF}$. ■

2 Euler operator is a strictly stronger contraction

Proposition 2, the main result of this paper, establishes that the Euler mapping is a contraction in the supremum norm, with a Lipschitz constant strictly smaller than the discount factor β. This implies that the Euler iteration converges more rapidly than standard value function iteration, offering both theoretical and computational advantages.

Proposition 2 applies Williams-Daly-Zachary (WDZ) Theorem and Lemmas 1 and 2 below.

WDZ THEOREM. *Let $\Lambda(\widetilde{\mathbf{v}}) \equiv \{\Lambda_a(\widetilde{\mathbf{v}}) : a \in \mathcal{A}\}$ be the Conditional Choice Probability (CCP) function, with:*

$$\Lambda_a(\widetilde{\mathbf{v}}) = \int 1\{\widetilde{v}_a + \widetilde{\varepsilon}_a \geq \widetilde{v}_j + \widetilde{\varepsilon}_j \quad \forall j \in \mathcal{A}\}\, g(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon} \tag{19}$$

Then, for any $a \in \mathcal{A} \setminus \{0\}$:

$$\frac{\partial S\left([0, \widetilde{\mathbf{v}}]\right)}{\partial \widetilde{v}_a} \;=\; \Lambda_a(\widetilde{\mathbf{v}}) \qquad \blacksquare \tag{20}$$

Proof of WDZ Theorem: Define the function $m(\widetilde{\mathbf{v}} + \widetilde{\boldsymbol{\varepsilon}}) = \max_{a \in \mathcal{A}} [\widetilde{v}_a + \widetilde{\varepsilon}_a]$. It is clear that:

$$\frac{\partial m(\widetilde{\mathbf{v}} + \widetilde{\boldsymbol{\varepsilon}})}{\partial \widetilde{v}_a} \;=\; 1\{\widetilde{v}_a + \varepsilon_a \geq \widetilde{v}_j + \widetilde{\varepsilon}_j \; \forall j \in \mathcal{A}\}. \tag{21}$$

By definition of the social surplus and $m(\widetilde{\mathbf{v}} + \widetilde{\boldsymbol{\varepsilon}})$ functions, we have that $S\left([0, \widetilde{\mathbf{v}}]\right) = \int m(\widetilde{\mathbf{v}} + \widetilde{\boldsymbol{\varepsilon}})\, g(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon}$, and $\partial S\left([0, \widetilde{\mathbf{v}}]\right) / \partial \widetilde{v}_a$ is equal to $\int \partial m(\widetilde{\mathbf{v}} + \widetilde{\varepsilon}) / \partial \widetilde{v}_a\, g(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon}$. Therefore,

$$\frac{\partial S\left([0, \widetilde{\mathbf{v}}]\right)}{\partial \widetilde{v}_a} \;=\; \int 1\{\widetilde{v}_a + \widetilde{\varepsilon}_a \geq \widetilde{v}_j + \widetilde{\varepsilon}_j \; \forall j \in \mathcal{A}\}\, g(\boldsymbol{\varepsilon})\, d\boldsymbol{\varepsilon} \;=\; \Lambda_a(\widetilde{\mathbf{v}}) \qquad \blacksquare \tag{22}$$

LEMMA 1. *In a multi-armed MDP:*

a. The image of the Euler mapping implies the following restrictions. For any $\widetilde{\mathbf{v}} \in \mathbb{R}^{J|\mathcal{X}|}$ and any $(a, y, \mathbf{z}) \in \mathcal{A} \setminus \{0\} \times \mathcal{A} \times \mathcal{Z}$:

$$\Gamma_{EE}(\widetilde{\mathbf{v}})(a, y, \mathbf{z}) - \Gamma_{EE}(\widetilde{\mathbf{v}})(a, 0, \mathbf{z}) \;=\; \Delta_a(y, \mathbf{z}) \tag{23}$$

where $\Delta_a(y, \mathbf{z}) \equiv c_a(y, \mathbf{z}) - c_a(0, \mathbf{z})$ *is a model primitive that does not depend on* $\widetilde{\mathbf{v}}$.

b. *This property implies that we can define the Euler operator in the vector space of value differences with* $y = 0$:

$$\mathcal{V}^R \equiv \left\{ \widetilde{\mathbf{v}}(0) = (\widetilde{v}_a(0, \mathbf{z}) : (a, \mathbf{z}) \in \mathcal{A} \setminus \{0\} \times \mathcal{Z}) \in \mathbb{R}^{J|\mathcal{Z}|} \right\} \tag{24}$$

c. *The Euler operator in the space of vector* $\widetilde{\mathbf{v}}(0)$ *is:*

$$\Gamma_{EE}(\widetilde{\mathbf{v}}(0))(a, \mathbf{z}) \;=\; c_a(0, \mathbf{z}) + \beta \sum_{\mathbf{z}'} \Big[S([0, \widetilde{\mathbf{v}}(0, \mathbf{z}') + \boldsymbol{\Delta}(a, \mathbf{z}')]) - S([0, \widetilde{\mathbf{v}}(0, \mathbf{z}')]) \Big] f_z(\mathbf{z}' \mid \mathbf{z}) \tag{25}$$

with $\boldsymbol{\Delta}(y, \mathbf{z}) \equiv (\Delta_a(y, \mathbf{z}) : a \in \mathcal{A} \setminus \{0\})$. ■

Proof of Lemma 1: Property [a] follows directly from the structure of the Euler mapping on equation (17). In particular, note that the continuation value is independent of the state variable y_t. Properties [b] and [c] are straightforward implications of Property [a]. ■

LEMMA 2. *Define the difference surplus function:*

$$\begin{aligned} \widetilde{S}_y(\widetilde{\mathbf{v}}(0, \mathbf{z})) \;&\equiv\; S([0, \widetilde{\mathbf{v}}(y, \mathbf{z})]) - S([0, \widetilde{\mathbf{v}}(0, \mathbf{z})]) \\ &=\; S([0, \widetilde{\mathbf{v}}(0, \mathbf{z}) + \Delta(y, \mathbf{z})]) - S([0, \widetilde{\mathbf{v}}(0, \mathbf{z})]) \end{aligned} \tag{26}$$

Let constant M *represent the supremum norm of the function defined by the partial derivative* $\partial \widetilde{S}_y(\widetilde{\mathbf{v}}(0, \mathbf{z})) / \partial \widetilde{v}_a(0, \mathbf{z})$:

$$M \;\equiv\; \left\| \frac{\partial \widetilde{S}_y(\widetilde{\mathbf{v}}(0, \mathbf{z}))}{\partial \widetilde{v}_a(0, \mathbf{z})} \right\|_{\infty} \;=\; sup_{a,y,\mathbf{z},\widetilde{\mathbf{v}}} \left| \frac{\partial \widetilde{S}_y(\widetilde{\mathbf{v}}(0, \mathbf{z}))}{\partial \widetilde{v}_a(0, \mathbf{z})} \right| \tag{27}$$

Then, $M < 1$. ■

Proof of Lemma 2: By the Williams-Daly-Zachary Theorem, we have:

$$\begin{aligned} \frac{\partial \widetilde{S}_y(\widetilde{\mathbf{v}}(0, \mathbf{z}))}{\partial \widetilde{v}_a(0, \mathbf{z})} \;&=\; \frac{\partial S([0, \widetilde{\mathbf{v}}(0, \mathbf{z}) + \Delta(y, \mathbf{z})])}{\partial \widetilde{v}_a(0, \mathbf{z})} - \frac{\partial S([0, \widetilde{\mathbf{v}}(0, \mathbf{z})])}{\partial \widetilde{v}_a(0, \mathbf{z})} \\ &=\; \Lambda_a\left(\widetilde{\mathbf{v}}(y, \mathbf{z})\right) - \Lambda_a\left(\widetilde{\mathbf{v}}(0, \mathbf{z})\right) \end{aligned} \tag{28}$$

where $\Lambda_a\left(\widetilde{\mathbf{v}}(y,\mathbf{z})\right)$ are CCPs strictly greater than zero and smaller than one. Therefore, the absolute value of this derivative is strictly smaller than one. It follows that $M<1$. ■

PROPOSITION 2. *Consider a multi-armed MDP. The Euler mapping Γ_{EE} is a contraction in the complete metric space $(\mathcal{V}^R, \|.\|_\infty)$, and its Lipschitz constant δ is strictly smaller than β. That is, there is a constant $\delta\in(0,1)$ with $\delta<\beta$ such that for any pair $\widetilde{\mathbf{v}}(0)$ and $\widetilde{\mathbf{w}}(0)$ in $\mathcal{V}^R$,*

$$\|\Gamma_{EE}(\widetilde{\mathbf{v}}(0))-\Gamma_{EE}(\widetilde{\mathbf{w}}(0))\|_\infty \;\leq\; \delta\;\|\widetilde{\mathbf{v}}(0)-\widetilde{\mathbf{w}}(0)\|_\infty \qquad \blacksquare \tag{29}$$

Proof of Proposition 2: By definition of mappings Γ_{EE} and $\widetilde{S}_y$, we have:

$$\Gamma_{EE}(\widetilde{\mathbf{v}}(0))(a,\mathbf{z}) \;=\; c(a,0,\mathbf{z})+\beta\sum_{\mathbf{z}'}\widetilde{S}_a(\widetilde{\mathbf{v}}(0,\mathbf{z}'))\;f_z(\mathbf{z}'|\mathbf{z}) \tag{30}$$

Therefore, for any $\widetilde{\mathbf{v}}(0)$ and $\widetilde{\mathbf{w}}(0)$ in $\mathcal{V}^R$:

$$\Gamma_{EE}(\widetilde{\mathbf{v}}(0))(a,\mathbf{z})-\Gamma_{EE}(\widetilde{\mathbf{w}}(0))(a,\mathbf{z}) \;=\; \beta\sum_{\mathbf{z}'}\left[\widetilde{S}_a(\widetilde{\mathbf{v}}(0,\mathbf{z}'))-\widetilde{S}_a(\widetilde{\mathbf{w}}(0,\mathbf{z}'))\right]f_z(\mathbf{z}'|\mathbf{z}) \tag{31}$$

Applying the Mean Value Theorem to function $\widetilde{S}_a$ between $\widetilde{\mathbf{v}}(0,\mathbf{z}')$ and $\widetilde{\mathbf{w}}(0,\mathbf{z}')$, we have that there is a vector of values $\widetilde{\mathbf{v}}^*(0,\mathbf{z}')$ such that $\widetilde{S}_a(\widetilde{\mathbf{v}}(0,\mathbf{z}'))-\widetilde{S}_a(\widetilde{\mathbf{w}}(0,\mathbf{z}'))=\nabla\widetilde{S}_a(\widetilde{\mathbf{v}}^*(0,\mathbf{z}'))$ $[\widetilde{\mathbf{v}}(0,\mathbf{z}')-\widetilde{\mathbf{w}}(0,\mathbf{z}')]$, where $\nabla\widetilde{S}_a$ represents the gradient vector. Thus, we have:

$$\Gamma_{EE}(\widetilde{\mathbf{v}}(0))(a,\mathbf{z})-\Gamma_{EE}(\widetilde{\mathbf{w}}(0))(a,\mathbf{z}) \;=\; \beta\sum_{\mathbf{z}'}\nabla\widetilde{S}_a(\widetilde{\mathbf{v}}^*(0,\mathbf{z}'))\,[\widetilde{\mathbf{v}}(0,\mathbf{z}')-\widetilde{\mathbf{w}}(0,\mathbf{z}')]\,f_z(\mathbf{z}'|\mathbf{z}) \tag{32}$$

Taking into account this equation, we have:

$$\begin{aligned}
|\Gamma_{EE}(\widetilde{\mathbf{v}}(0))(a,\mathbf{z})-\Gamma_{EE}(\widetilde{\mathbf{w}}(0))(a,\mathbf{z})| \;&\leq\; \beta\sum_{\mathbf{z}'}\left|\left|\nabla\widetilde{S}_a(\widetilde{\mathbf{v}}^*(0,\mathbf{z}'))\right|\right|\;||\widetilde{\mathbf{v}}(0,\mathbf{z}')-\widetilde{\mathbf{w}}(0,\mathbf{z}')||\,f_z(\mathbf{z}'|\mathbf{z})\\
&\leq\; \beta M\sum_{\mathbf{z}'}||\widetilde{\mathbf{v}}(0,\mathbf{z}')-\widetilde{\mathbf{w}}(0,\mathbf{z}')||\,f_z(\mathbf{z}'|\mathbf{z})\\
&\leq\; \beta M\sup_{\mathbf{z}'}||\widetilde{\mathbf{v}}(0,\mathbf{z}')-\widetilde{\mathbf{w}}(0,\mathbf{z}')|| \;=\; \beta M\,||\widetilde{\mathbf{v}}(0)-\widetilde{\mathbf{w}}(0)||
\end{aligned} \tag{33}$$

The first inequality follows from the triangle inequality. The second uses Lemma 2. The third exploits the fact that the maximum over $\mathbf{z}'$ is greater than or equal to the average. This completes the proof that Γ_{EE} is a contraction with modulus $\delta=\beta\,M<\beta$. ■

3 Solution algorithms

We present the solution algorithms using vector-form notation. Let $|\mathcal{Z}|$ denote the dimension of the space of exogenous state variables. All vectors described below are of dimension $|\mathcal{Z}| \times 1$. The vector $\mathbf{V}(y)$ represents the value function when the endogenous state is $y_t = y$. Similarly, $\boldsymbol{\pi}_a(y)$, $\widetilde{\mathbf{v}}_a(y)$, and $\mathbf{P}_a(y)$ denote the vectors of one-period payoffs, value differences, and conditional choice probabilities, respectively, given current action $a_t = a$ and endogenous state $y_t = y$. Finally, $\mathbf{F}_z$ is a $|\mathcal{Z}| \times |\mathcal{Z}|$ matrix representing the transition probabilities of the exogenous state variables.

3.1 Value function (VF) iterations

The value function operator is a fixed-point mapping in the space of the vector of values $\mathbf{V} = \{\mathbf{V}(y) : y \in \mathcal{A}\}$. It is defined as $\Gamma_{VF}(\mathbf{V}) = \{\Gamma_{VF}(\mathbf{V})(y) : y \in \mathcal{A}\}$ with

$$\Gamma_{VF}(\mathbf{V})(y) \quad = \quad \int \max_{a \in A} \left\{ \boldsymbol{\pi}_a(y) + \beta \; \mathbf{F}_z \; \mathbf{V}(a) + \varepsilon_{a,t} \right\} \, g(\boldsymbol{\varepsilon}_t) \; d\boldsymbol{\varepsilon}_t \tag{34}$$

The algorithm starts with an initial vector of values $\mathbf{V}_0$. At every iteration $n \geq 1$ it updates the vector using $\mathbf{V}_n = \Gamma_{VF}(\mathbf{V}_{n-1})$.

The computational complexity per iteration arises from the number of multiplications required to compute the matrix products $\mathbf{F}_z \; \mathbf{V}_{n-1}(a)$ across all $J+1$ choice alternatives. Specifically, for each triple $(a, z, z') \in \mathcal{A} \times \mathcal{Z}^2$, we must evaluate the product $f_z(z'|z) * V(a, \mathbf{z}')$. Therefore, the computational complexity per-iteration is on the order of $(J+1) \; |\mathcal{Z}|^2$.

The contraction rate of this mapping, as measured by its Lipschitz constant, is equal to the discount factor β (see Puterman, 1994, and Rust, 1996).

3.2 Relative value function (RVF) iterations

Let $\mathbf{x}_0 \in \mathcal{X}$ be an arbitrary value of the vector of state variables. The *relative value function operator* is a fixed-point mapping in the space of the vector of values $\mathbf{V} = \{\mathbf{V}(y) : y \in \mathcal{A}\}$, that is defined as $\Gamma_{RVF}(\mathbf{V}) = \{\Gamma_{RVF}(\mathbf{V})(y) : y \in \mathcal{A}\}$ with

$$\Gamma_{RVF}(\mathbf{V})(y) \quad = \quad \int \max_{a \in A} \left\{ \boldsymbol{\pi}_a(y) + \beta \, \mathbf{F}_z \, \left[\mathbf{V}(a) - V(\mathbf{x}_0) \; \mathbf{1}_{|\mathcal{Z}|} \right] + \varepsilon_t(j) \right\} \, g(\boldsymbol{\varepsilon}_t) \; d\boldsymbol{\varepsilon}_t \tag{35}$$

where $\mathbf{1}_{|\mathcal{Z}|}$ is a column vector of ones. Given an initial $\mathbf{V}_0$, this vector is updated at every iteration $n \geq 1$ using $\mathbf{V}_n = \Gamma_{RVF}(\mathbf{V}_{n-1})$.

The computational complexity per iteration of the RVF algorithm is the same as that of value function iteration. It is driven by the number of multiplications required to compute the matrix product $\mathbf{F}_z \left[\mathbf{V}(j) - V(\mathbf{x}_0)\, \mathbf{1}_{|\mathcal{Z}|}\right]$. As a result, the complexity is of order $(J+1)\,|\mathcal{Z}|^2$.

The key difference between the two algorithms lies in their contraction properties. The Lipschitz constant of the RVF operator is $\beta\ \rho(\mathbf{F}_z)$, where where $\rho(\mathbf{F}_z) \in (0,1)$ denotes the spectral radius of the transition matrix $\mathbf{F}_z$ (Morton and Wecker, 1977; Puterman, 1994, section 6.6; Bray, 2019, Proposition 3).

When the stochastic process governing the exogenous state variables exhibits strong persistence, the spectral radius $\rho(\mathbf{F}_z)$ is close to one, and the contraction properties of the RVF and VF operators are similar. In contrast, when the exogenous process is less persistent, $\rho(\mathbf{F}_z)$ is significantly smaller than one, and the RVF algorithm converges faster than value function iteration.

3.3 Policy function (PF), or Newton-Kantorovich, iterations

The policy function operator (Puterman and Brumelle, 1979) is a fixed-point mapping in the space of the vector of *policies* or conditional choice probabilities $\mathbf{P} = \{\mathbf{P}_a(y) : (a,y) \in \mathcal{A} \setminus \{0\} \times \mathcal{A}\}$. It is defined as $\Gamma_{PF}(\mathbf{P}) = \{\Gamma_{PF}(\mathbf{P})(a,y) : (a,y) \in \mathcal{A} \setminus \{0\} \times \mathcal{A}\}$ with,

$$\Gamma_{PF}(\mathbf{P})(a,y) \ = \ \int 1\left\{\boldsymbol{\pi}_a(y) + \beta\, \mathbf{F}_z\, \mathbf{W}_a^{\mathbf{P}} + \varepsilon_{a,t} \geq \boldsymbol{\pi}_j(y) + \beta\, \mathbf{F}_z\, \mathbf{W}_j^{\mathbf{P}} + \varepsilon_{j,t},\ \forall j\right\}\ g(\boldsymbol{\varepsilon}_t)\ d\boldsymbol{\varepsilon}_t \quad (36)$$

where $\mathbf{W}_y^{\mathbf{P}}$ is an $|\mathcal{Z}| \times 1$ vector that contains the present discounted values of future payoffs conditional on every possible value of $\mathbf{z}_t$ and on $y_t = y$, and assumes that the agent's future behavior follows the vector of choice probabilities $\mathbf{P}$. The vectors of present values $\{\mathbf{W}_y^{\mathbf{P}} : y \in \mathcal{A}\}$ are obtained solving the system of linear equations: for any $y \in \mathcal{A}$,

$$\mathbf{W}_y^{\mathbf{P}} \ = \ \sum_{a=0}^{J} \mathbf{P}_a(y) * \left[\boldsymbol{\pi}_a(y) + \mathbf{e}_a(\mathbf{P}(y)) + \beta\, \mathbf{F}_z\ \mathbf{W}_a^{\mathbf{P}}\right], \quad (37)$$

where $*$ is the element-by-element product, and $\mathbf{e}_a(\mathbf{P}(y))$ is the vector of conditional expectations of $\varepsilon_{a,t}$ conditional on alternative a being the optimal choice.

The linear operator described in the system (37), which delivers the vectors $\mathbf{W}_a^{\mathbf{P}}$ for a given vector of CCPs $\mathbf{P}$, is denoted the *valuation operator*. The operator in the right-hand side of equation (36) that returns a new vector of CCPs given the vectors of valuations $\mathbf{W}^{\mathbf{P}}$ is denoted the *policy improvement operator*. By definition, the policy function operator in (36) is the composition of the valuation and the policy improvement operators.

The PF algorithm starts with an initial $\mathbf{P}_0$, and at every iteration $n \geq 1$ updates the vector using $\mathbf{P}_n = \Gamma_{PF}(\mathbf{P}_{n-1})$. The computational complexity of one iteration in the PF operator is given by the complexity of solving the system of linear equations (37). This complexity is determined by the number of objects to be solved in this system, and so is of the order $O(|\mathcal{X}|^3)$. Therefore, one PF iteration is more costly than one VF or RVF iteration, and the difference increases with the dimension of the state space.

However, the PF operator is a stronger contraction than VF and RVF operators, and therefore it requires a smaller number of iterations to achieve convergence.

3.4 Euler operator (EE) iterations

The EE operator is a fixed-point mapping in space of value differences for $y = 0$, $\widetilde{\mathbf{v}}(0) = \{\widetilde{\mathbf{v}}_a(0) : a \in \mathcal{A} \setminus \{0\}\}$. It is defined as $\Gamma_{EE}(\widetilde{\mathbf{v}}(0)) = \{\Gamma_{EE}(\widetilde{\mathbf{v}}(0))(a) : a \in \mathcal{A} \setminus \{0\}\}$ with,

$$\Gamma_{EE}(\widetilde{\mathbf{v}}(0))(a) \; = \; \mathbf{c}_a(0) \; + \; \beta \; \mathbf{F}_z \Big(S([0, \widetilde{\mathbf{v}}(0) + \mathbf{\Delta}(a)]) - S([0, \widetilde{\mathbf{v}}(0)]) \Big) \tag{38}$$

where $\mathbf{c}_a(0) \equiv \widetilde{\boldsymbol{\pi}}_a(0) + \beta \, \mathbf{F}_z \Big(\boldsymbol{\pi}_0(a) - \boldsymbol{\pi}_0(0) \Big)$.

The algorithm begins with an initial vector $\widetilde{\mathbf{v}}_0(0) \in \mathbb{R}^{J|\mathcal{Z}|}$, and at each iteration $n \geq 1$, it updates the vector using the rule $\widetilde{\mathbf{v}}_n(0) = \Gamma_{EE}(\widetilde{\mathbf{v}}_{n-1}(0))$. The computational complexity of each iteration is of order $J|\mathcal{Z}|^2$, driven by the evaluation of the matrix products $\mathbf{F}_z \, [S(\widetilde{\mathbf{v}}(a)) - S(\widetilde{\mathbf{v}}(0))]$. This involves computing the product $f_z(z'|z) \cdot (S(\widetilde{\mathbf{v}}(a, z')) - S(\widetilde{\mathbf{v}}(0, z')))$ for each triple $(a, z, z') \in (\mathcal{A} \setminus \{0\}) \times \mathcal{Z}^2$. Importantly, this computation is not required for one of the alternatives (namely, choice 0), which reduces the total number of operations. Thus, the overall computational complexity per iteration is $J|\mathcal{Z}|^2$, compared to $(J + 1)|\mathcal{Z}|^2$ for value function iteration (VF) and recursive value function iteration (RVF).

The Euler iteration algorithm offers two key computational advantages over VF and RVF methods. First, its per-iteration cost is lower, since it computes expected continuation values for all but one alternative, while VF and RVF compute them for all alternatives. This difference becomes more pronounced as the dimensionality of the state space increases. Second, and more importantly, the Euler operator is a stronger contraction, which leads to faster convergence in terms of iteration count. These advantages are illustrated in our Monte Carlo experiments.

References

Aguirregabiria, V., and A. Magesan (2013): "Euler equations for the estimation of dynamic discrete choice structural models," in *Structural Econometric Models (Advances in Econometrics)*, vol. 31, pp. 3–44. Emerald Group Publishing Limited.

Arcidiacono, P., and R. Miller (2011): "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity," *Econometrica*, 79(6), 1823–1867.

——— (2019): "Nonstationary dynamic models with finite dependence," *Quantitative Economics*, 10(3), 853–890.

Bray, R. L. (2019): "Strong convergence and dynamic economic models," *Quantitative Economics*, 10(1), 43–65.

Morton, T. E., and W. E. Wecker (1977): "Discounting, ergodicity and convergence for Markov decision processes," *Management Science*, 23(8), 890–900.

Puterman, M. (1994): *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons. New Jersey.

Puterman, M., and S. Brumelle (1979): "On the convergence of policy iteration in stationary dynamic programming," *Mathematics of Operations Research*, 4(1), 60–69.

Rust, J. (1994): "Structural estimation of Markov decision processes," *Handbook of Econometrics*, 4, 3081–3143.

——— (1996): "Numerical dynamic programming in economics," *Handbook of Computational Economics*, 1, 619–729.