ISOMORPHIC DYNAMIC PROGRAMS

JOHN STACHURSKI AND JUNNAN ZHANG

Abstract. We study relationships between dynamic programs by applying conjugacy methods from dynamical systems theory. When two dynamic programs are connected by an order isomorphism, we show that optimality properties transmit from one formulation to the other. We apply these results to Epstein–Zin preferences with time preference shocks, obtaining a sharp characterization of when optimality holds. We also show that multiplicative Kreps–Porteus preferences and risk-sensitive preferences are isomorphic, so that well-known results for the latter carry over to the former. Finally, we demonstrate how isomorphic transformations can improve the numerical accuracy of value function approximations, with gains of two orders of magnitude in a multisector real business cycle model.

1. Introduction

Dynamic programming is a major field of optimization with applications across a wide spectrum of scientific domains, including economics and finance (see, e.g., Bellman (1957), Stokey and Lucas (1989), Puterman (2005), Hernández-Lerma and Lasserre (2012), Bäuerle and Rieder (2011), Bertsekas (2012)). In recent years, variations of the standard model have flourished. This is particularly true in economics and finance, as applied researchers extend dynamic models to get closer to the data. Some of these variations add features to the standard framework, including Epstein–Zin preferences, risk-sensitive preferences, adversarial agents, ambiguity, and hyperbolic discounting.[1] Other variations take an existing model and rearrange the structure and timing. These variations include the Q-factor and exponential risk-sensitive Q-factor

Date: May 21, 2026.

The authors thank Yuchao Li, Shu Hu, and Jingni Yang for valuable comments and suggestions. Financial support from Schmidt Futures is gratefully acknowledged. The second author acknowledges support from NSFC Project 72303190.

[1]See, for example, Epstein and Zin (1989), Cagetti et al. (2002), Hansen and Sargent (2011), Bäuerle and Jaśkiewicz (2018), Fedus et al. (2019), Marinacci and Montrucchio (2019), Gao et al. (2021), or de Castro et al. (2022).

formulations from Q-learning, and the expected value function models and integrated value function models from structural estimation.[2]

Given this proliferation of variations and modifications, one significant issue is that each new formulation typically requires its own bespoke analysis to establish desirable optimality properties (e.g., existence of optimal policies, Bellman's principle of optimality, or convergence of Howard policy iteration). In this paper, we address this issue from a new direction, by borrowing ideas from the field of dynamical systems. Since each policy operator defines a discrete-time dynamical system, tools from that field become directly applicable. In particular, we draw on notions of conjugacy. In dynamical systems and chaos theory, topological conjugacy is used to identify families of dynamical systems that have "equivalent" or related dynamics (see, e.g., Smale (1967) or Arnold (2012)). These ideas have proved enormously useful for categorizing different types of systems and allowing researchers to make deductions about a particular system from their knowledge of other related systems.[3]

Specifically, we transfer these ideas to dynamic programming by studying order-theoretic conjugacy relationships between the policy operators that define a dynamic program. We call two dynamic programs linked in this way *isomorphic*; under an order-reversing link, we call them *anti-isomorphic*. Our main result shows that verifying this single relationship is enough to transfer an entire package of optimality properties from one program to the other: existence and uniqueness of the value function, Bellman's principle of optimality, existence of optimal policies, and convergence of Howard policy iteration. Solving one program therefore delivers a complete solution to the other, including recovery of the value function and optimal policies.

The benefits of this approach are both theoretical and numerical. On the theoretical side, dynamic programs that are not easily analyzed can often be connected, via an isomorphism, to a program whose optimality properties are already established or more readily verified; the isomorphism then allows those properties to carry over to the original problem. We illustrate this in two settings that are not covered by standard contraction-based methods. First, for a dynamic program with Epstein–Zin utility and time preference shocks, we obtain an exact characterization of when

[2]See, for example, Kochenderfer et al. (2022), Rust (1994), Fei et al. (2021), or Kristensen et al. (2021).

[3]For example, Kennedy and Stockman (2008), Gardini et al. (2009), and Raines and Stockman (2012) characterize chaotic dynamical systems in economic models utilizing the shift homeomorphism. In a similar vein, Flynn and Sastry (2022) establish topological conjugacy between equilibrium economic dynamics and dynamical systems with known chaotic behaviors. Also see Battaglini (2021) and Deng et al. (2022).

optimality holds. Second, we show that multiplicative Kreps–Porteus preferences are isomorphic to risk-sensitive preferences, so that the well-developed contraction theory for the latter yields optimality results for the former. On the numerical side, within an isomorphism class one is free to work with whichever representative is best suited to computation. Exploiting this flexibility in the multisector real business cycle model of Long and Plosser (1983), we obtain solutions that are, on average, two orders of magnitude more accurate than solving the original problem, by choosing a representative within the isomorphism class whose value function has lower curvature.

This paper builds on the abstract dynamic programming framework of Sargent and Stachurski (2025). In that paper, individual dynamic programs are represented as families of policy operators acting on a partially ordered space. When dynamic programs are viewed in this way, studying conjugacy relationships between related families of policy operators becomes a natural idea, since each policy operator identifies a discrete time dynamical system. The main deviation from traditional dynamical systems methodology is that we use order-theoretic conjugacy relationships, rather than topological ones, since the key objects in optimization are order-theoretic and preserved by this relation.

This paper contributes to a large and growing literature that studies the optimality properties of dynamic programs with nonstandard features, such as recursive utility (Marinacci and Montrucchio, 2010; Bloise et al., 2024; Rincón-Zapatero, 2024), risk sensitive preferences (Bäuerle and Jaśkiewicz, 2018), hyperbolic discounting (Bäuerle et al., 2021; Balbus et al., 2022), and Q-factors (Ma et al., 2022). For example, our work relates to the literature on existence and uniqueness in dynamic programs with recursive preferences, such as Bloise and Vailakis (2018) and Bloise et al. (2024). Our approach complements this literature in two ways. First, rather than focusing on specific models, we propose a general framework that, under appropriate transformations, can simplify the analysis of such dynamic programs. Second, as an application of our theory, we provide exact necessary and sufficient conditions for optimality in dynamic programs with Epstein–Zin preferences and state dependent discounting.

While not the primary focus of this paper, our work also connects to the literature on computational methods for solving dynamic economic models. These include traditional methods such as projection and perturbation (Foerster et al., 2016; Bayer and Luetticke, 2020), endogenous grid method (Carroll, 2006; Barillas and Fernández-Villaverde, 2007), adaptive sparse grids (Brumm and Scheidegger, 2017), and deep learning (Azinovic et al., 2022; Maliar et al., 2021). A notable feature of our framework is that it can complement existing computation methods: any suitable method can be applied to an isomorphic dynamic program to improve numerical performance compared to solving the original problem.

The structure of the paper is as follows. Section 2 introduces order conjugacy and discusses its properties. Sections 3–4 recall some definitions and optimality results for "abstract" dynamic programs, which are well-suited to our dynamical systems perspective. Section 5 introduces isomorphisms between abstract dynamic programs and shows how isomorphisms preserve optimality properties. Section 6 provides applications and Section 7 concludes.

2. Preliminaries

This section introduces the key order-theoretic concepts needed for our analysis of abstract dynamic programs. Additional definitions of posets, dynamical systems, and related background are given in Appendix A. Throughout, if A and B are sets, then the symbol B^A indicates the set of all maps from A to B.

A *partially ordered set* (poset) is a pair $V = (V, \preceq)$ where $\preceq$ is reflexive, antisymmetric, and transitive. Given posets $(V, \preceq)$ and $(W, \trianglelefteq)$, a function $F\colon V \to W$ is called *order-preserving* if $v \preceq w$ implies $Fv \trianglelefteq Fw$, and *order-reversing* if $v \preceq w$ implies $Fw \trianglelefteq Fv$. A bijection $F\colon V \to \hat{V}$ is an *order isomorphism* if both F and F^{-1} are order-preserving, and an *order anti-isomorphism* if both F and F^{-1} are order-reversing.

Let (V, S) be a dynamical system (a set V with a self-map S) where V is a poset and S has a unique fixed point $\bar{v}$. Following Sargent and Stachurski (2025), (V, S) is called *upward stable* if $v \preceq S\,v$ implies $v \preceq \bar{v}$, *downward stable* if $S\,v \preceq v$ implies $\bar{v} \preceq v$, and *order stable* if both hold.

Figure 1 gives an illustration of an order stable map S on $V = [0, 1]$. All points mapped up by S lie below its unique fixed point, while all points mapped down by S lie above its fixed point.

Intuitively, order stability says that the fixed point $\bar{v}$ of S separates V into two regions: elements that S maps upward all lie below $\bar{v}$, and elements that S maps downward all lie above $\bar{v}$. This is an order-theoretic analogue of global stability. (A dynamical system (V, S) on a metric space is called *globally stable* if S has a unique fixed point $\bar{v}$ and $S^k v \to \bar{v}$ for all $v \in V$.) Whereas global stability requires a topology and ensures convergence of trajectories, order stability requires only a partial order and is not as strong. In fact, order stability is implied by global stability when S is order-preserving and the space V has both a topology *and* a partial order:

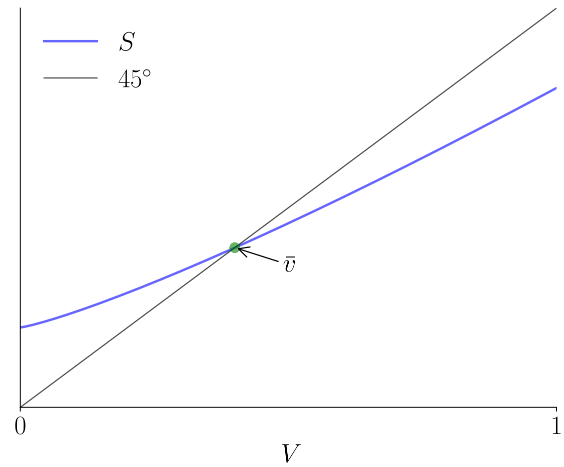

Figure 1. An order stable map S on $[0, 1]$

Example 2.1. Let $(V, \preceq, \rho)$ be a partially ordered space (i.e., a metric space (V, ρ) where $\preceq$ is closed under limits). If (V, S) is globally stable and S is order-preserving, then (V, S) is order stable. To see this, let $\bar{v}$ be the unique fixed point of S in V. Upward stability holds because if $v \in V$ and $v \preceq S\,v$, then, iterating on this inequality and using the fact that S is order-preserving, we have $v \preceq S^k\,v$ for all $k \in \mathbb{N}$. Applying global stability and taking the limit gives $v \preceq \bar{v}$. Hence upward stability holds. The proof of downward stability is similar.

Two dynamical systems (V, S) and $(\hat{V}, \hat{S})$ on posets are called *order conjugate* under F when they are conjugate (i.e., $F \circ S = \hat{S} \circ F$ for some bijection F) and F is an order isomorphism. Order conjugacy is an equivalence relation that preserves order stability: if (V, S) and $(\hat{V}, \hat{S})$ are order conjugate, then one is order stable if and only if the other is (Lemma A.4 in Appendix A.4).

3. Abstract Dynamic Programs

We follow the abstract dynamic programming framework of Sargent and Stachurski (2025), which was partly inspired by Bertsekas (2022). For us, the benefit of the framework in Sargent and Stachurski (2025) is that dynamic programs are represented as families of policy operators. This abstract representation allows us to apply the notion of order conjugacy to study relationships between dynamic programs.

3.1. **Definition and Examples.** Sargent and Stachurski (2025) define an *abstract dynamic program* (ADP) to be a pair $\mathscr{A} = (V, \{T_\sigma\}_{\sigma \in \Sigma})$, where

(i) $V = (V, \preceq)$ is a partially ordered set and
(ii) $\{T_\sigma\}_{\sigma \in \Sigma}$ is a family of order-preserving self-maps on V, indexed by $\sigma \in \Sigma$.

Elements of the index set Σ are referred to as *policies* and elements of $\{T_\sigma\}_{\sigma \in \Sigma}$ are called *policy operators*. When Σ is understood, we often write $\{T_\sigma\}_{\sigma \in \Sigma}$ as $\{T_\sigma\}$. In all applications, the significance of each policy operator T_σ is that its fixed point, denoted below by v_σ, represents the lifetime value (or cost) of following policy σ.

Example 3.1 (MDPs)**.** Consider a *Markov decision process* (MDP; see, e.g., Puterman (2005)) where the aim is to maximize $\mathbb{E} \sum_{t \geqslant 0} \beta^t r(X_t, A_t)$ when X_t takes values in finite set X (the state space), A_t takes values in finite set A (the action space), Γ is a nonempty correspondence from X to A, the set $\mathsf{G} := \{(x, a) \in \mathsf{X} \times \mathsf{A} : a \in \Gamma(x)\}$ denotes feasible state-action pairs, r is a reward function defined on G, $\beta \in (0, 1)$ is a discount factor, and $P \colon \mathsf{G} \times \mathsf{X} \to [0, 1]$ provides transition probabilities. The Bellman equation for this problem is

$$v(x) = \max_{a \in \Gamma(x)} \left\{ r(x, a) + \beta \sum_{x'} v(x') P(x, a, x') \right\} \qquad (x \in \mathsf{X}). \tag{1}$$

The set of feasible policies is $\Sigma := \{\sigma \in \mathsf{A}^{\mathsf{X}} : \sigma(x) \in \Gamma(x) \text{ for all } x \in \mathsf{X}\}$. We combine $\mathbb{R}^{\mathsf{X}}$ (the set of all real-valued functions on X) with the pointwise partial order $\leqslant$ and, for $\sigma \in \Sigma$ and $v \in \mathbb{R}^{\mathsf{X}}$, define the MDP policy operator

$$(T_\sigma \, v)(x) = r(x, \sigma(x)) + \beta \sum_{x'} v(x') P(x, \sigma(x), x') \qquad (x \in \mathsf{X}). \tag{2}$$

The pair $(\mathbb{R}^{\mathsf{X}}, \{T_\sigma\})$ is an ADP.

Example 3.2 (Risk-sensitive MDPs)**.** Risk-sensitive MDPs (Howard and Matheson, 1972) modify the MDP model in Example 3.1 so that the policy operators take the form

$$(T_\sigma^\theta \, v)(x) = r(x, \sigma(x)) + \frac{\beta}{\theta} \ln \left[\sum_{x'} \exp(\theta v(x')) P(x, \sigma(x), x') \right]$$

where θ is some fixed value in $\Theta := \mathbb{R} \setminus \{0\}$. The pair $(\mathbb{R}^{\mathsf{X}}, \{T_\sigma^\theta\})$ is an ADP.

Further examples of ADPs, including Q-factor and risk-sensitive Q-factor formulations from the reinforcement learning literature, are given in Appendix B.

3.2. **Lifetime Values.** The objective of dynamic programming is to maximize or minimize lifetime value. Following Sargent and Stachurski (2025), we identify the lifetime value of any given policy σ as the unique fixed point of T_σ, whenever it exists. When it does exist, we denote this fixed point by v_σ and call it the *σ-value function*.

Example 3.3. Consider the MDP setting of Example 3.1. Let r_σ and P_σ be defined by

$$r_\sigma(x) := r(x, \sigma(x)) \quad \text{and} \quad (P_\sigma v)(x) := \sum_{x'} v(x') P(x, \sigma(x), x'). \tag{3}$$

The lifetime value of policy σ given $X_0 = x$ is $v_\sigma(x) = \mathbb{E} \sum_{t \geqslant 0} \beta^t r(X_t, \sigma(X_t))$, where $(X_t)_{t \geqslant 0}$ is a Markov chain generated by P_σ with initial condition $X_0 = x \in \mathsf{X}$. Pointwise on X, we can express v_σ as

$$v_\sigma = \sum_{t \geqslant 0} (\beta P_\sigma)^t r_\sigma = (I - \beta P_\sigma)^{-1} r_\sigma \tag{4}$$

(see, e.g., Puterman (2005), Theorem 6.1.1, or Kochenderfer et al. (2022), Section 7.2). Equivalently, v_σ is the unique solution to the equation $v = r_\sigma + \beta P_\sigma v$. Inspecting the definition of T_σ in (2), we see this is also equivalent to the statement that v_σ is the unique fixed point of T_σ.

Similarly, for the risk-sensitive MDP of Example 3.2, the operator T_σ^θ is a contraction on $\mathbb{R}^{\mathsf{X}}$ and its unique fixed point represents the lifetime value of following σ.

3.3. **Greedy Policies.** The core idea behind the theory of dynamic programming is that an optimal policy can be obtained by choosing actions that solve a two-period problem involving a Bellman equation (Bellman, 1957). The optimal actions in the two-period problem produce what are typically called "greedy policies." For example, in the context of Example 3.1, a policy σ satisfying

$$\sigma(x) \in \operatorname{arg\,max} \left\{ r(x, a) + \beta \sum_{x'} v(x') P(x, a, x') \right\} \qquad \text{for all } x \in \mathsf{X}. \tag{5}$$

is called a "greedy policy with respect to v." See, for example, Kochenderfer et al. (2022), Section 7.3.

Following Sargent and Stachurski (2025), we generalize this idea to the abstract setting. In particular, given an ADP $(V, \{T_\sigma\})$ and an element $v \in V$, a policy σ in Σ is called

- *v-min-greedy* if $T_\sigma v \preceq T_\tau v$ for all $\tau \in \Sigma$, and

- *v-max-greedy* if $T_\tau\, v \preceq T_\sigma\, v$ for all $\tau \in \Sigma$.

For example, in the context of Example 3.1, with $\leqslant$ as the pointwise partial order, a policy σ obeying (5) satisfies $T_\tau\, v \leqslant T_\sigma\, v$ for all $\tau \in \Sigma$, and hence is v-max-greedy. A v-min-greedy policy can be constructed by replacing arg max in (5) with arg min.

3.4. **Bellman Operators.** For a generic ADP $(V, \{T_\sigma\})$, we respectively define the *Bellman min-operator* and the *Bellman max-operator* via

$$T_\triangledown\, v := \bigwedge_\sigma T_\sigma\, v \quad \text{and} \quad T_\vartriangle\, v := \bigvee_\sigma T_\sigma\, v \tag{6}$$

whenever the infimum (resp., supremum) exists. We say that $v \in V$ satisfies the *Bellman max-equation* (resp., the *Bellman min-equation*) if it is a fixed point of $T_\vartriangle$ (resp., $T_\triangledown$). Notice that $\sigma \in \Sigma$ is

(i) v-max-greedy if and only if $T_\sigma\, v = T_\vartriangle\, v$, and
(ii) v-min-greedy if and only if $T_\sigma\, v = T_\triangledown\, v$.

To illustrate, consider the MDP setting of Example 3.1. Traditionally, the Bellman operator for this model is given by

$$(Tv)(x) = \max_{a \in \Gamma(x)} \left\{ r(x,a) + \beta \sum_{x'} v(x') P(x,a,x') \right\} \qquad (x \in \mathsf{X}). \tag{7}$$

(See, e.g., Puterman (2005).) For the ADP $(\mathbb{R}^{\mathsf{X}}, \{T_\sigma\})$ generated by this MDP, this Bellman operator exactly coincides with the Bellman max-operator $T_\vartriangle$ in (6). Replacing max with min in (7) produces the Bellman min-operator from (6).

3.5. **Properties of ADPs.** To obtain optimality results, we need to place structure on ADPs. Here and below, we call an ADP $\mathscr{A} = (V, \{T_\sigma\})$

- *well-posed* if T_σ has a unique fixed point v_σ in V for each $\sigma \in \Sigma$,
- *order stable* if (V, T_σ) is order stable for each $\sigma \in \Sigma$,
- *max-stable* if $\mathscr{A}$ is order stable, each $v \in V$ has at least one max-greedy policy, and $T_\vartriangle$ has at least one fixed point in V, and
- *min-stable* if $\mathscr{A}$ is order stable, each $v \in V$ has at least one min-greedy policy, and $T_\triangledown$ has at least one fixed point in V.

Well-posedness is a minimum regularity condition for ADPs. Without it, we cannot be sure that policies have well-defined lifetime values. Well-defined lifetime values are essential because maximizing (or minimizing) lifetime value over the set of all policies is the objective of dynamic programming.

Order stability is a natural regularity assumption for ADPs. To understand this, suppose $\mathscr{A}$ is well-posed and consider upward stability for an arbitrary policy operator T_σ with fixed point v_σ. If $v \preceq T_\sigma v$, then following policy σ for one period offers an improvement in value. Since the problem is stationary, this suggests that following the policy forever will also be an improvement. Thus, we expect $v \preceq v_\sigma$, in which case upward stability holds. Intuition for downward stability is similar.

Example 3.4. Consider the MDP setting of Example 3.1 and fix $\sigma \in \Sigma$. Recall that the policy operator T_σ has unique fixed point $v_\sigma := (I - \beta P_\sigma)^{-1} r_\sigma$. If $T_\sigma v \geqslant v$, then $r_\sigma + \beta P_\sigma v \geqslant v$ and hence $(I - \beta P_\sigma) v \leqslant r_\sigma$. Since $(I - \beta P_\sigma)^{-1}$ is positive, we get $v \leqslant v_\sigma$. Hence (V, T_σ) is upward stable. A similar proof shows that (V, T_σ) is downward stable, and therefore order stable.

Example 3.5. Consider the ADP $(\mathbb{R}^{\mathsf{X}}, \{T_\sigma\})$ from the MDP setting of Example 3.1. For this case $T_\vartriangle$ is given by (7). Since $T_\vartriangle$ is a contraction on $\mathbb{R}^{\mathsf{X}}$, it has a unique fixed point in $\mathbb{R}^{\mathsf{X}}$. Since max-greedy policies always exist, and since $(\mathbb{R}^{\mathsf{X}}, \{T_\sigma\})$ is order stable by Example 3.4, we see that $(\mathbb{R}^{\mathsf{X}}, \{T_\sigma\})$ is max-stable.

Regarding the definition of max-stability (resp., min-stability), the Bellman min- and max-operators are often contraction maps and existence of a fixed point is easily verified (see, e.g., Denardo (1967) or Chapter 2 of Bertsekas (2022)). Here is another useful condition, which covers the case where state and action spaces are finite.

Proposition 3.1. *Let $\mathscr{A}$ be order stable and suppose the set of policies is finite. In this setting,*

(i) *if each $v \in V$ has at least one max-greedy policy, then $\mathscr{A}$ is max-stable, and*
(ii) *if each $v \in V$ has at least one min-greedy policy, then $\mathscr{A}$ is min-stable.*

Proposition 3.1 is proved in the appendix (page 23).

Example 3.6. The risk-sensitive MDP from Example 3.2 is max-stable. Evidently T^θ_σ is order-preserving on $V = \mathbb{R}^{\mathsf{X}}$. Moreover, (V, T^θ_σ) is globally stable (see, e.g., Bäuerle and Jaśkiewicz (2018)) and hence order stable, by Example 2.1. This shows that $(V, \{T^\theta_\sigma\})$ is an order stable ADP. Given $v \in V$, we can construct a v-max-greedy policy σ by setting

$$\sigma(x) \in \arg\max \left\{ r(x, a) + \frac{\beta}{\theta} \ln \left[\sum_{x'} \exp(\theta v(x')) P(x, a, x') \right] \right\}$$

for all $x \in \mathsf{X}$. As the policy set Σ is finite (since X and the choice sets are all finite), Proposition 3.1 implies that $(V, \{T_\sigma^\theta\})$ is max-stable.

4. Optimality

In this section we state conditions for optimality in our abstract setting. These conditions are closely related to those in Sargent and Stachurski (2025). Later, in Section 5, we will study how optimality properties are preserved under transformations.

4.1. Max-Optimality.

Let $\mathscr{A}$ be a well-posed ADP. We let $V_\Sigma := \{v_\sigma\}_{\sigma \in \Sigma}$ denote the set of σ-value functions. If V_Σ has a greatest element, then we denote it by $v_\vartriangle$ and call it the *max-value function*. A policy $\sigma \in \Sigma$ is called *max-optimal* for $\mathscr{A}$ if v_σ is a greatest element of V_Σ; that is, if $v_\vartriangle$ exists and $v_\sigma = v_\vartriangle$.

We define a map $H_\vartriangle$ from V to $\{v_\sigma\}$ via $H_\vartriangle\, v = v_\sigma$ where σ is v-max-greedy. Iterating with $H_\vartriangle$ is an abstraction of Howard policy iteration.[4] In what follows, we call $H_\vartriangle$ the *Howard max-operator* generated by the ADP.

In the following result, we take $\mathscr{A}$ to be an ADP with Bellman operator $T_\vartriangle$.

Theorem 4.1 (Max-optimality). *If $\mathscr{A}$ is max-stable, then*

(i) *the max-value function $v_\vartriangle$ exists in V,*
(ii) *$v_\vartriangle$ is the unique solution to the Bellman max-equation in V,*
(iii) *a policy is max-optimal if and only if it is $v_\vartriangle$-max-greedy.*
(iv) *at least one max-optimal policy exists.*

If, in addition, Σ is finite, then Howard max-policy iteration converges to $v_\vartriangle$ in finitely many steps.

The last statement means that, for all $v \in V$, there exists a $K \in \mathbb{N}$ such that $k \geqslant K$ implies $H_\vartriangle^k v = v_\vartriangle$. The proof of Theorem 4.1 is given in the appendix.

Analogous results hold for minimization; see Appendix C.3 for the definitions of min-optimality, the Howard min-operator, dual ADPs, and the corresponding min-optimality theorem.

[4]For $H_\vartriangle$ to be well-defined, we must always select the same v-greedy policy when the operator is applied to v. We can use the axiom of choice to assign to each v a designated v-greedy policy, although, in applications, a simple rule usually suffices. For example, if Σ is finite, we can enumerate the policy set Σ and choose the first v-greedy policy.

5. Isomorphic ADPs

In this section we introduce isomorphic relationships between ADPs and explore their implications for optimality. True to their name, isomorphic relationships are symmetric, transitive and reflexive. We show that isomorphic ADPs have identical optimality properties.

5.1. Definition and Properties.

Let $\mathscr{A} = (V, \{T_\sigma\})$ and $\hat{\mathscr{A}} = (\hat{V}, \{\hat{T}_\sigma\})$ be two ADPs. We call $\mathscr{A}$ and $\hat{\mathscr{A}}$ *isomorphic* under F if these two ADPs have the same policy set Σ and F is an order isomorphism from V to $\hat{V}$ such that

$$F \circ T_\sigma = \hat{T}_\sigma \circ F \quad \text{on } V \text{ for all } \sigma \in \Sigma. \tag{8}$$

In other words, (V, T_σ) and $(\hat{V}, \hat{T}_\sigma)$ are order conjugate under F for all $\sigma \in \Sigma$.[5]

An illustration of isomorphic ADPs built from Q-factor formulations is given in Appendix B. We also provide a further illustration involving risk-sensitive and Kreps–Porteus preferences in Section 6.3.

Remark 5.1. For ADPs $\mathscr{A}, \hat{\mathscr{A}}$, let $\mathscr{A} \sim \hat{\mathscr{A}}$ indicate that $\mathscr{A}$ and $\hat{\mathscr{A}}$ are isomorphic. It is elementary to show that the relation $\sim$ is reflexive, symmetric and transitive. Hence $\sim$ is an equivalence relation on the set of all ADPs.

5.2. Isomorphisms and Optimality.

We seek a connection between value functions and optimality properties of isomorphic ADPs. The theory below provides this relationship. For all of this section, we take $\mathscr{A} = (V, \{T_\sigma\})$ and $\hat{\mathscr{A}} = (\hat{V}, \{\hat{T}_\sigma\})$ to be two ADPs with the same policy set. When they exist, we let

- v_σ (resp., $\hat{v}_\sigma$) be the unique fixed point of T_σ (resp., $\hat{T}_\sigma$)
- $T_\vartriangle$ (resp., $\hat{T}_\vartriangle$) be the Bellman max-operator of $\mathscr{A}$ (resp., $\hat{\mathscr{A}}$)
- $T_\triangledown$ (resp., $\hat{T}_\triangledown$) be the Bellman min-operator of $\mathscr{A}$ (resp., $\hat{\mathscr{A}}$)
- $v_\vartriangle$ (resp., $\hat{v}_\vartriangle$) be the max-value function of $\mathscr{A}$ (resp., $\hat{\mathscr{A}}$)
- $v_\triangledown$ (resp., $\hat{v}_\triangledown$) be the min-value function of $\mathscr{A}$ (resp., $\hat{\mathscr{A}}$)

Isomorphic ADPs share the same regularity properties:

Theorem 5.1. *If $\mathscr{A}$ and $\hat{\mathscr{A}}$ are isomorphic under F, then*

(i) *$\mathscr{A}$ is well-posed if and only if $\hat{\mathscr{A}}$ is well-posed.*

[5] While the definition requires that the two ADPs have the same policy set Σ, it suffices that the policy sets can be put in one-to-one correspondence with each other.

(ii) *$\mathscr{A}$ is order stable if and only if $\hat{\mathscr{A}}$ is order stable.*

(iii) *$\mathscr{A}$ is max-stable if and only if $\hat{\mathscr{A}}$ is max-stable. In this case,*

$$F \circ T_\Delta = \hat{T}_\Delta \circ F \quad \text{and} \quad \hat{v}_\Delta = F v_\Delta. \tag{9}$$

Moreover, $\mathscr{A}$ and $\hat{\mathscr{A}}$ have the same max-optimal policies.

(iv) *$\mathscr{A}$ is min-stable if and only if $\hat{\mathscr{A}}$ is min-stable. In this case,*

$$F \circ T_\nabla = \hat{T}_\nabla \circ F \quad \text{and} \quad \hat{v}_\nabla = F v_\nabla. \tag{10}$$

Moreover, $\mathscr{A}$ and $\hat{\mathscr{A}}$ have the same min-optimal policies.

The proof (page 24) uses condition (8) to show that F commutes with the Bellman operators, and then combines Theorem 4.1 with properties of order conjugate systems.

5.3. **Anti-Isomorphic ADPs.** Let $\mathscr{A} = (V, \{T_\sigma\})$ and $\hat{\mathscr{A}} = (\hat{V}, \{\hat{T}_\sigma\})$ be two ADPs. We call $\mathscr{A}$ and $\hat{\mathscr{A}}$ *anti-isomorphic* under F if they have the same policy set Σ and, in addition, there exists an order anti-isomorphism F from V to $\hat{V}$ such that (8) holds (see Example B.3).

If $\mathscr{A} \overset{a}{\sim} \hat{\mathscr{A}}$ indicates that $\mathscr{A}$ and $\hat{\mathscr{A}}$ are anti-isomorphic, then $\overset{a}{\sim}$ is symmetric and transitive but, in general, not reflexive.

Here is an optimality result for anti-isomorphic ADPs that parallels Theorem 5.1.

Theorem 5.2. *If $\mathscr{A}$ and $\hat{\mathscr{A}}$ are anti-isomorphic under F, then*

(i) *$\mathscr{A}$ is well-posed if and only if $\hat{\mathscr{A}}$ is well-posed.*

(ii) *$\mathscr{A}$ is order stable if and only if $\hat{\mathscr{A}}$ is order stable.*

(iii) *$\mathscr{A}$ is max-stable if and only if $\hat{\mathscr{A}}$ is min-stable. In this case,*

$$F \circ T_\Delta = \hat{T}_\nabla \circ F \quad \text{and} \quad \hat{v}_\nabla = F v_\Delta. \tag{11}$$

Moreover, $\sigma \in \Sigma$ is max-optimal for $\mathscr{A}$ if and only if σ is min-optimal for $\hat{\mathscr{A}}$.

The proof of Theorem 5.2 is given in the appendix (page 24).

6. Applications

In this section we show how isomorphic relationships can simplify or illuminate dynamic programming problems.

6.1. **Modified Epstein–Zin Equations.** Consider an Epstein–Zin version of the MDP in Example 3.1 (see, e.g., Epstein and Zin (1989) or Weil (1990)), in which a Bellman max-equation takes the form

$$v(x) = \max_{a \in \Gamma(x)} \left\{ r(x,a)^{\alpha} + \beta(x) \left(\sum_{x'} v(x')^{\gamma} P(x,a,x') \right)^{\alpha/\gamma} \right\}^{1/\alpha}.$$

Following much of the recent literature, we have allowed the rate of time preference β to depend on the state (see, e.g., Albuquerque et al. (2016), Schorfheide et al. (2018), de Groot et al. (2018), Gomez-Cram and Yaron (2020)). The function β maps X to $\mathbb{R}_+$ and γ and α are nonzero parameters. Other details are as in Example 3.1. Using the symbols r_σ and P_σ from (3), the policy operator T_σ can be written as

$$T_\sigma v = \left\{ r_\sigma^{\alpha} + \beta \left(P_\sigma v^{\gamma} \right)^{\alpha/\gamma} \right\}^{1/\alpha}, \tag{12}$$

where powers are taken pointwise. Since γ and α can be negative, we assume that r is positive. We also suppose that P_σ is irreducible for all $\sigma \in \Sigma$. Under these assumptions, T_σ is an order-preserving self-map on $(V, \leqslant)$, the set of all strictly positive functions on X paired with the pointwise partial order, and $\mathscr{A} := (V, \{T_\sigma\})$ is an ADP.

Now set

$$\theta := \frac{\gamma}{\alpha} \quad \text{and} \quad \hat{T}_\sigma v := \left\{ r_\sigma^{\alpha} + \beta \left(P_\sigma v \right)^{1/\theta} \right\}^{\theta}. \tag{13}$$

The pair $\hat{\mathscr{A}} = (V, \{\hat{T}_\sigma\})$ is also an ADP.

Lemma 6.1. *The following relationships hold:*

(i) *If $\gamma > 0$, then $\mathscr{A}$ and $\hat{\mathscr{A}}$ are isomorphic.*
(ii) *If $\gamma < 0$, then $\mathscr{A}$ and $\hat{\mathscr{A}}$ are anti-isomorphic.*

Proof. Let $F \colon V \to V$ be the bijective map $Fv = v^{\gamma}$. Fixing σ and applying (12) yields

$$F\, T_\sigma v = (T_\sigma v)^{\gamma} = \left\{ r_\sigma^{\alpha} + \beta \left(P_\sigma v^{\gamma} \right)^{\alpha/\gamma} \right\}^{\gamma/\alpha} = \left\{ r_\sigma^{\alpha} + \beta \left(P_\sigma v^{\gamma} \right)^{1/\theta} \right\}^{\theta}.$$

Inspection of (13) shows that $\hat{T}_\sigma F v = \hat{T}_\sigma v^{\gamma}$ is identical to the last expression in the display above. Hence $F \circ T_\sigma = \hat{T}_\sigma \circ F$ on V. If $\gamma > 0$, then F is order-preserving, so $\mathscr{A}$ and $\hat{\mathscr{A}}$ are isomorphic. If $\gamma < 0$, then F is order-reversing, so $\mathscr{A}$ and $\hat{\mathscr{A}}$ are anti-isomorphic. □

Lemma 6.1 gives us a way to solve for max-optimal policies of $\mathscr{A}$ by studying $\hat{\mathscr{A}}$ (and applying either Theorem 5.1 or Theorem 5.2). This is convenient because $\hat{\mathscr{A}}$ is easier to analyze. The next section illustrates.

6.2. **Characterizing Optimality.** Suppose that β depends on x through a purely exogenous state component (as in, say, de Groot et al. (2018) and Schorfheide et al. (2018)). Specifically, $\mathsf{X} = \mathsf{Y} \times \mathsf{Z}$ and $x = (y, z)$, where

$$P(x, a, x') = R(y, a, y')Q(z, z') \quad \text{and} \quad \beta(x) = \beta(z).$$

Here $R(y, a, \cdot)$ is a distribution over y for each feasible (y, a) pair and Q is a stochastic matrix over Z. For each $z \in \mathsf{Z}$, let $(Z_t(z))_{t \geqslant 0}$ be a Markov chain on Z generated by Q and starting at z. Define

$$\mathcal{E}(\beta, Q, \theta) := \lim_{k \to \infty} \left\{ \sup_{z \in \mathsf{Z}} \mathbb{E} \prod_{t=0}^{k-1} \beta(Z_t(z))^{\theta} \right\}^{1/k}. \tag{14}$$

We can now state the following exact result.

Theorem 6.2. *The Epstein–Zin ADP $\mathscr{A}$ is max-stable if and only if*

$$\mathcal{E}(\beta, Q, \theta)^{1/\theta} < 1 \tag{15}$$

Hence, under (15), all of the optimality results in Theorem 4.1 apply. Conversely, if (15) fails, then $\mathscr{A}$ is not well-posed and optimality is undefined.

The sufficiency of (15) for optimality properties is related to an earlier result in Stachurski and Zhang (2021). It is also connected to Bloise et al. (2024), who establish sufficient conditions for optimality via the spectral radius of a monotone sublinear operator. The converse result is new. Unlike the existing literature, the proof of Theorem 6.2 proceeds by studying the simpler ADP $\hat{\mathscr{A}}$ introduced in Section 6.1, which we show is either isomorphic or anti-isomorphic, depending on the sign of γ. Then the theorem follows from Proposition 3.1 together with a result on irreducible positive operators. See Section C.4 of the appendix.

6.3. **Kreps–Porteus vs Risk Sensitive Preferences.** As a further application, we show that two widely used preference specifications, the risk-sensitive preferences and a multiplicative variation of Kreps–Porteus preferences, are connected by an isomorphism. This allows optimality results from the well-studied risk-sensitive framework to be transferred to the multiplicative Kreps–Porteus model.

We continue with the finite MDP framework of Example 3.1, using r_σ and P_σ as in (3). Recall the risk-sensitive MDP in Example 3.2, where the policy operators take the form

$$(T_\sigma^\theta \, v)(x) = r_\sigma(x) + \frac{\beta}{\theta} \ln \left[\sum_{x'} \exp(\theta v(x')) P(x, \sigma(x), x') \right] \tag{16}$$

for $v \in \mathbb{R}^{\mathsf{X}}$ and $\theta \in \mathbb{R} \setminus \{0\}$. We show in Example 3.6 that the ADP $(\mathbb{R}^{\mathsf{X}}, \{T_\sigma^\theta\})$ is max-stable and hence all of the optimality results in Theorem 4.1 apply.

An alternative formulation is to replace the entropic certainty equivalent with Kreps–Porteus expectations, leading to the policy operator

$$(K_\sigma v)(x) = r_\sigma(x) + \beta \{(P_\sigma v^\nu)(x)\}^{1/\nu} \tag{17}$$

for $v \in (0, \infty)^{\mathsf{X}}$ and $\nu \in \mathbb{R} \setminus \{0\}$, where r is strictly positive. The Kreps–Porteus ADP $((0, \infty)^{\mathsf{X}}, \{K_\sigma\})$ is harder to analyze directly because K_σ is not generally a contraction. There is, however, a multiplicative variation on the Kreps–Porteus ADP that is simple to analyze. The model is obtained by setting

$$(M_\sigma v)(x) = r_\sigma(x) \cdot \{(P_\sigma v^\nu)(x)\}^{\beta/\nu} \tag{18}$$

where $v \in (0, \infty)^{\mathsf{X}}$, r is strictly positive, and $\beta \in [0, 1)$. We call $\hat{\mathscr{A}} := ((0, \infty)^{\mathsf{X}}, \{M_\sigma\})$ the multiplicative Kreps–Porteus (MKP) ADP.

It turns out that the MKP ADP and the risk-sensitive ADP are isomorphic. To see this, we take logs of the Bellman max-equation associated with the MKP ADP, obtaining

$$\ln v(x) = \max_{a \in \Gamma(x)} \left\{ \ln r(x, a) + \frac{\beta}{\nu} \ln \left[\sum_{x'} v(x')^\nu P(x, a, x') \right] \right\}.$$

Setting $\hat{v} = \ln v$ and $\hat{r} = \ln r$ yields

$$\hat{v}(x) = \max_{a \in \Gamma(x)} \left\{ \hat{r}(x, a) + \frac{\beta}{\nu} \ln \left[\sum_{x'} \exp(\nu \, \hat{v}(x')) P(x, a, x') \right] \right\}.$$

This is exactly the Bellman equation for (16), after replacing r with $\hat{r}$ and θ with ν.

We can formalize this observation. The MKP ADP is $\hat{\mathscr{A}} = ((0, \infty)^{\mathsf{X}}, \{M_\sigma\})$. Let $F \colon (0, \infty)^{\mathsf{X}} \to \mathbb{R}^{\mathsf{X}}$ be the pointwise logarithm, $Fv = \ln v$, which is an order isomorphism with inverse $F^{-1} = \exp$. Consider the risk-sensitive ADP $\mathscr{A} := (\mathbb{R}^{\mathsf{X}}, \{T_\sigma^\nu\})$ with reward $\hat{r} = \ln r$ and risk-sensitivity parameter $\theta = \nu$.

It remains to verify condition (8), i.e., that $M_\sigma = F^{-1} \circ T_\sigma^\nu \circ F$ on $(0, \infty)^{\mathsf{X}}$. Setting $\hat{v} = Fv = \ln v$ for $v \in (0, \infty)^{\mathsf{X}}$, we have

$$\begin{aligned}
(F^{-1} \circ T_\sigma^\nu \circ F)(v)(x) &= \exp\left[\ln r_\sigma(x) + \frac{\beta}{\nu} \ln \left[\sum_{x'} \exp(\nu \, \hat{v}(x')) P(x, \sigma(x), x') \right] \right] \\
&= r_\sigma(x) \cdot \left\{ \sum_{x'} \exp(\nu \, \ln v(x')) P(x, \sigma(x), x') \right\}^{\beta/\nu} \\
&= r_\sigma(x) \cdot \{(P_\sigma v^\nu)(x)\}^{\beta/\nu} = (M_\sigma v)(x).
\end{aligned}$$

Since F is an order isomorphism, the MKP ADP $\hat{\mathscr{A}}$ and the risk-sensitive ADP $\mathscr{A}$ are isomorphic. By Theorem 5.1, they share all optimality properties: $\hat{\mathscr{A}}$ is max-stable if and only if $\mathscr{A}$ is max-stable, the value functions are related by $\hat{v}_\Delta = Fv_\Delta$, and the two ADPs have the same optimal policies.

This result provides a practical benefit. Since the risk-sensitive MDP is an additive model with well-developed theory (Bäuerle and Jaśkiewicz, 2018), the isomorphism allows us to immediately establish optimality results for the multiplicative Kreps–Porteus model without the need for separate analysis. The same isomorphism extends naturally to continuous state space settings, where sums are replaced by expectations and $\mathbb{R}^{\mathsf{X}}$ by bounded measurable functions.

6.4. **Curvature and Approximation.** Implementing dynamic programs with continuous state spaces on computers requires approximation of value functions, policy functions, or both. The need for high quality approximations of these functions is amplified by the fact that, in dynamic programming, most computational algorithms use some form of iteration (such as value or Howard policy iteration), and approximation errors tend to compound across iterations, potentially leading to poor final approximations or failures of convergence (see, e.g., Farahmand et al. (2010)).

In general, functions are harder to approximate when they involve high degrees of curvature. Intuitively, isomorphic relationships can improve accuracy of approximations and stabilize iteration by transforming dynamic programs in order to reduce curvature or other complexities in target functions. We explore this idea in the current section, using the multisector real business cycle model of Long and Plosser (1983) as a test case.

In their model, the Bellman max-equation has the form

$$v(y) = \max_{c,X} \left\{ u(c) + \beta \int v(f(\lambda, X)) \varphi(\mathrm{d}\lambda) \right\}, \tag{19}$$

subject to

$$0 \leqslant c_i, X_{ij}, y_i \quad \text{and} \quad c_j + \sum_{i=1}^{n} X_{ij} = y_j \quad \text{for all } i, j \text{ in } \{1, \ldots, n\}. \tag{20}$$

The function f is a Borel measurable production function taking values in $\mathbb{R}^n_+$, $c = (c_j)$ is an n-vector of consumption quantities across n goods, $y = (y_j)$ is an n-vector of final outputs, $X = (x_{ij})$ is an $n \times n$ matrix of commodity inputs, and φ is a distribution over $\mathbb{R}^n_+$. (The labor-leisure decision is missing from (19) because we adopt a special case of the parameterization in Long and Plosser (1983) that allows us to assume inelastic labor supply.)

Let the state space be $\mathsf{Y} = \mathbb{R}^n_+$ and let $V = m\mathbb{R}^{\mathsf{Y}}$, the space of (extended) real-valued Borel measurable functions on Y. The policy operator is given by

$$(T_\sigma v)(y) = u(\sigma_c(y)) + \beta \int v\left[f(\lambda, \sigma_X(y))\right] \varphi(d\lambda) \qquad (y \in \mathsf{Y})$$

where $v \in V$ and $\sigma = (\sigma_c, \sigma_X)$ is a Borel measurable feasible policy such that $(c, X) = \sigma(y)$ satisfies the constraints in (20). Then the model represented by the pair $(V, \{T_\sigma\})$ is an ADP. In many settings, such as the one we will study below, this program is well-posed and max-stable. In view of Theorem 5.1, any isomorphic program shares the same optimality and convergence properties of the original problem. Therefore, we can equivalently solve the transformed Bellman equation

$$w(y) = \max_{c,X} F\left[u(c) + \beta \int (F^{-1} \circ w)(f(\lambda, X))\varphi(\mathrm{d}\lambda)\right] \tag{21}$$

for any strictly increasing function $F \colon \mathbb{R} \to \mathbb{R}$, and recover the value function of the original problem via $v = F^{-1}w$. (We understand F as both a strictly increasing function from $\mathbb{R}$ to itself and as an order isomorphism on V sending v into $y \mapsto F(v(y))$.) This is advantageous because, for suitably chosen F, solving (21) significantly improves numerical accuracy.

To illustrate this point, we follow Long and Plosser (1983) in assuming that utility obeys $U(c) = \sum_{i=1}^n \theta_i \ln c_i$ and production has the multisector Cobb-Douglas form $f(y, \lambda) = \lambda \prod_{j=1}^n x_{ij}^{a_{ij}}$. Under these assumptions, an analytical solution to (19) exists, allowing us to compare approximate numerical solutions with exact solutions. The analytical solution has the form $v_0(y) = \sum_{i=1}^n \gamma_i \ln y_i + K$, where γ_i and K depend on the parameters a_{ij} and θ_i. This function has relatively high curvature near zero, with both slope and curvature approaching infinity as output converges to the origin. These features make it challenging to approximate with a high degree of accuracy. In contrast, if we consider an isomorphic transformation (21) with $F(x) = e^{mx+b}$, then the transformed value function is given by

$$(Fv_0)(y) = e^{mK+b} \prod_{i=1}^n y_i^{m\gamma_i}. \tag{22}$$

This function has lower curvature and finite derivatives at zero. These properties facilitate approximation and lead to improved accuracy.

To demonstrate, we solve (19) and (21) for the two-sector case under several parameterizations using value function iteration where the value functions are approximated

Table 1. Accuracy Gain from Solving the Transformed Problem

Parameterization	$A = ([0.2, 0.7], [0.6, 0.1])$	$A = ([0.1, 0.7], [0.3, 0.1])$
	$\theta_2 = 1.0$	$\theta_2 = 1.0$
$\theta_1 = 0.5$	104.6	68.8
$\theta_1 = 0.6$	153.6	65.3
$\theta_1 = 0.7$	143.6	63.1
$\theta_1 = 0.8$	109.7	62.6
$\theta_1 = 0.9$	173.3	178.9
$\theta_1 = 1.0$	178.7	122.3

by linear interpolation on a common grid. To ensure a fair comparison, we perform a fixed number of iterations ($N = 500$) for both the original and transformed problems. The maximization operations in (19) and (21) are carried out using the Adam optimizer (Kingma and Ba, 2014) for efficient computation. In each maximization operation, we run the optimizer for 2000 iterations following a cosine decay learning rate schedule with a 500-step warmup. The initial, peak, and final learning rates of the schedule are 1×10^{-6}, 1×10^{-4}, and 1×10^{-7}, respectively, chosen through trial and error. We then compute the sup-norm errors for the two problems, $\|v - v_0\|$ and $\|F^{-1}w - v_0\|$. Table 1 shows the improvement in accuracy from solving the transformed problem, measured by $\|v - v_0\| / \|F^{-1} \circ w - v_0\|$.[6]

Although the transformed value function w is not exactly linear for the two-sector case (see (22)), we can make it approximately linear in at least one dimension by choosing F appropriately (e.g., setting $m = 1/\gamma_1$). Table 1 shows that this improved approximation of the value function leads to solutions that are, on average, 100 times more accurate across a wide range of parameterizations.

7. Conclusion

We studied isomorphic and anti-isomorphic relationships between dynamic programs and showed how optimality properties transmit across these relationships. We applied these ideas to Epstein–Zin preferences with time preference shocks, to the

[6]For each parameterization, we choose the parameters of the exponential transformation $F(x) = e^{mx+b}$ so that Fv_0 is approximately linear in at least one dimension. In particular, in the second column, $m = [0.35, 0.3, 0.3, 0.25, 0.25, 0.25]$ and $b = [75, 60, 65, 60, 60, 75]$; in the third column, $m = [0.9, 0.8, 0.7, 0.7, 0.6, 0.5]$ and $b = [90, 85, 80, 85, 80, 75]$. In all cases, the state space is $[1 \times 10^{-7}, 20]^2$ with 500 grid points along each dimension. The other parameter values are listed in the table.

connection between multiplicative Kreps–Porteus and risk-sensitive preferences, and to improving numerical accuracy in multisector real business cycle models.

Our research focused on discrete time dynamic programs, which we paired with discrete time concepts of topological and order conjugacy. Continuous time dynamic programs are also important and the concept of topological conjugacy has a well-defined analogy in continuous time dynamics. This suggests that many of our ideas can be carried over to continuous time systems. We leave this task for future research.

Appendix A. Order-Theoretic Preliminaries

This appendix collects the order-theoretic definitions, dynamical systems background, and auxiliary results that underpin the main text.

A.1. **Posets.** Let $V = (V, \preceq)$ be a partially ordered set, also called a poset. When $A \subset V$, the symbol $\bigvee A$ refers to the supremum of A in V and $\bigwedge A$ is the infimum (see, e.g., Davey and Priestley (2002)). V is called *bounded* if V has a least and greatest element. A sequence (v_n) in $V = (V, \preceq)$ is called *increasing* if $v_n \preceq v_{n+1}$ for all $n \in \mathbb{N}$. If (v_n) is increasing and $\bigvee_n v_n = v$ for some $v \in V$, then we write $v_n \uparrow v$. The set V is called *countably chain complete* if V is bounded and every increasing sequence in V has a supremum in V. A self-map S on V is called *order continuous* on V if $Sv_n \uparrow Sv$ whenever $v_n \uparrow v$.

A.2. **Dynamical Systems.** A *dynamical system* is a pair (V, S) where V is a set and S is a self-map on V. Given $v \in V$, the sequence $(S^n v)_{n \geqslant 1}$ is called the *trajectory* of v under S. When V has a topology, we can study the convergence of trajectories. In particular, in this setting, a system (V, S) is called *globally stable* when S has a unique fixed point $\bar{v}$ in V and $\lim_{k \to \infty} S^k v = \bar{v}$ for all $v \in V$.

Two dynamical systems (V, S) and $(\hat{V}, \hat{S})$ are called *conjugate* under F if there exists a bijection F from V to $\hat{V}$ such that $F \circ S = \hat{S} \circ F$ on V. In this setting, if v is a fixed point of S, then $\hat{S}Fv = FSv = Fv$, so Fv is a fixed point of $\hat{S}$. With slightly more effort, one can show the following.

Lemma A.1. *If (V, S) and $(\hat{V}, \hat{S})$ are conjugate under F and S has a unique fixed point $v \in V$, then Fv is the unique fixed point of $\hat{S}$ in $\hat{V}$.*

For more details see, for example, Sternberg (2014) or Layek (2015).

Dynamical systems (V, S) and $(\hat{V}, \hat{S})$ are called *topologically conjugate* under F if both V and $\hat{V}$ have topologies, the pairs (V, S) and $(\hat{V}, \hat{S})$ are conjugate under F and, in addition, F is a homeomorphism (i.e., continuous bijection with continuous inverse) from V to $\hat{V}$. In this setting we can extend Lemma A.1 as follows.

Lemma A.2. *If* (V, S) *and* $(\hat{V}, \hat{S})$ *are topologically conjugate under* F*, then* (V, S) *is globally stable if and only if* $(\hat{V}, \hat{S})$ *is globally stable.*

Lemma A.2 is standard and routinely used in dynamical systems theory. (There are similar results concerning local stability for topologically conjugate systems, as well as additional results concerning other kinds of asymptotic behavior. See, for example, Sternberg (2014) or Layek (2015).)

A.3. **Order Stability.** Let (V, S) be a dynamical system where V is partially ordered by $\preceq$ and S has a unique fixed point $\bar{v}$ in V. In this environment, Sargent and Stachurski (2025) call (V, S)

(i) *upward stable* if $v \in V$ and $v \preceq Sv$ implies $v \preceq \bar{v}$,
(ii) *downward stable* if $v \in V$ and $Sv \preceq v$ implies $\bar{v} \preceq v$, and
(iii) *order stable* if upward and downward stability both hold.

In this paper, we discuss both maximization and minimization. To link these ideas in a poset environment, we use the notion of order duals. In particular, given poset $V = (V, \preceq)$, let $(V, \preceq^{\partial})$ be the order dual, so that, for $u, v \in V$, we have $u \preceq^{\partial} v$ if and only if $v \preceq u$. For convenience, we sometimes denote $(V, \preceq^{\partial})$ by V^{∂}.

Lemma A.3. (V, S) *is order stable if and only if* (V^{∂}, S) *is order stable.*

Proof of Lemma A.3. Let (V, S) be order stable. By definition, S has a unique fixed point $\bar{v} \in V$. We claim that (V^{∂}, S) is upward and downward stable. Regarding upward stability, suppose $v \in V$ and $v \preceq^{\partial} Sv$. Then $Sv \preceq v$ and hence $\bar{v} \preceq v$, by downward stability of (V, S). But then $v \preceq^{\partial} \bar{v}$, so (V^{∂}, S) is upward stable. The proof of downward stability is similar. Hence order stability of (V, S) is sufficient for order stability of (V^{∂}, S). Necessity follows from sufficiency, since the dual of (V^{∂}, S) is (V, S). □

A.4. **Order Conjugacy.** Our next step is to replace the concept of topological conjugacy from Section A.2 with a parallel notion for dynamical systems on posets. To this end, we recall that a bijective map $F : V \to \hat{V}$ is called an *order isomorphism* if both F and its inverse F^{-1} are order preserving. (We call $F : V \to \hat{V}$ an *order anti-isomorphism* if both F and F^{-1} are order reversing.) Let (V, S) and $(\hat{V}, \hat{S})$ be two dynamical systems where V and $\hat{V}$ are partially ordered. We call (V, S) and $(\hat{V}, \hat{S})$ *order conjugate* under F when (V, S) and $(\hat{V}, \hat{S})$ are conjugate under F and, in addition,

F is an order isomorphism. It is easy to verify that order conjugacy is an equivalence relation on the set of dynamical systems over partially ordered sets.

Lemma A.4. *If (V, S) and $(\hat{V}, \hat{S})$ are order conjugate under F, then (V, S) is order stable if and only if $(\hat{V}, \hat{S})$ is order stable.*

Proof of Lemma A.4. Let (V, S) and $(\hat{V}, \hat{S})$ be order conjugate under F, and suppose that (V, S) is order stable. By Lemma A.1, the map $\hat{S}$ has a unique fixed point in $\hat{V}$. Let $\hat{w}$ be an element of $\hat{V}$ satisfying $\hat{S}\hat{w} \preceq \hat{w}$. Let v and $\hat{v} := Fv$ be the fixed points of S and $\hat{S}$, respectively. Then $F^{-1}\hat{S}\hat{w} \preceq F^{-1}\hat{w}$ and hence $SF^{-1}\hat{w} \preceq F^{-1}\hat{w}$. But then $v \preceq F^{-1}\hat{w}$, by downward stability of (V, S). Applying F gives $\hat{v} \preceq \hat{w}$. Hence $(\hat{V}, \hat{S})$ is downward stable. Similarly, if $\hat{w}$ is an element of $\hat{V}$ satisfying $\hat{w} \preceq \hat{S}\hat{w}$, then $F^{-1}\hat{w} \preceq F^{-1}\hat{S}\hat{w} = SF^{-1}\hat{w}$. By upward stability of (V, S), we have $F^{-1}\hat{w} \preceq v$. Applying F gives $\hat{w} \preceq \hat{v}$, so $(\hat{V}, \hat{S})$ is upward stable. Together, these results show that $(\hat{V}, \hat{S})$ is order stable. The converse implication follows from symmetry. □

Appendix B. Additional Examples

The following examples extend the ADP framework of Section 3 to Q-factor formulations from the reinforcement learning literature. All primitives are as in Example 3.1.

Example B.1 (Q-factors)**.** The Q-learning literature studies the Q-factor Bellman equation (see, e.g., Kochenderfer et al. (2022), Section 17.2), which is given by

$$f(x, a) = r(x, a) + \beta \sum_{x'} \max_{a' \in \Gamma(x')} f(x', a') P(x, a, x') \qquad ((x, a) \in \mathsf{G}). \tag{23}$$

Here $f \in \mathbb{R}^{\mathsf{G}}$ is called the *Q-factor*. The policy operators over Q-factors take the form

$$(Q_\sigma f)(x, a) = r(x, a) + \beta \sum_{x'} f(x', \sigma(x')) P(x, a, x'), \tag{24}$$

where $f \in \mathbb{R}^{\mathsf{G}}$ and $\sigma \in \Sigma$. If we pair $\mathbb{R}^{\mathsf{G}}$ (the set of all real-valued functions on G) with the pointwise partial order $\leqslant$ and Q_σ as in (24), then $(\mathbb{R}^{\mathsf{G}}, \{Q_\sigma\})$ is an ADP.

Example B.2 (Risk-sensitive Q-factors)**.** It has become popular in reinforcement learning and related fields to extend the Q-factor approach from Example B.1 to risk-sensitive decision processes (see, e.g., Fei et al. (2021)). The corresponding Q-factor Bellman equation is given by

$$f(x, a) = r(x, a) + \frac{\beta}{\theta} \ln \left\{ \sum_{x'} \exp \left[\theta \max_{a' \in \Gamma(x')} f(x', a') \right] P(x, a, x') \right\} \qquad ((x, a) \in \mathsf{G}). \tag{25}$$

The policy operators over risk-sensitive Q-factors take the form

$$(Q_\sigma^\theta f)(x,a) = r(x,a) + \frac{\beta}{\theta} \ln \left[\sum_{x'} \exp\left[\theta f(x', \sigma(x'))\right] P(x,a,x') \right] \tag{26}$$

where $f \in \mathbb{R}^\mathsf{G}$ and $\sigma \in \Sigma$. The pair $(\mathbb{R}^\mathsf{G}, \{Q_\sigma^\theta\})$ is an ADP.

Example B.3 (Exponential risk-sensitive Q-factors)**.** Fei et al. (2021) work with an "exponential" risk-sensitive Q-factor Bellman equation where the corresponding policy operators have the form

$$(M_\sigma h)(x,a) = \exp\left\{ \theta r(x,a) + \beta \ln \left[\sum_{x'} h(x', \sigma(x')) P(x,a,x') \right] \right\}$$

Here $h \in (0,\infty)^\mathsf{G}$ and M_σ maps $(0,\infty)^\mathsf{G}$ into itself. All primitives are as in the risk-sensitive Q-factor ADP $\mathscr{A} := (\mathbb{R}^\mathsf{G}, \{Q_\sigma\})$ in Example B.2. Since each M_σ is order preserving (under the usual pointwise order), the pair $\hat{\mathscr{A}} := ((0,\infty)^\mathsf{G}, \{M_\sigma\})$ is also an ADP. If we take F to be the bijection from $\mathbb{R}^\mathsf{G}$ to $(0,\infty)^\mathsf{G}$ defined by $(Fh)(x,a) = \exp(\theta h(x,a))$, then, for Q_σ defined in (26), $h \in \mathbb{R}^\mathsf{G}$ and $(x,a) \in \mathsf{G}$,

$$(FQ_\sigma h)(x,a) = \exp\left\{ \theta r(x,a) + \beta \ln \left[\sum_{x'} \exp\left[\theta h(x', \sigma(x'))\right] P(x,a,x') \right] \right\}.$$

This is equal to $(M_\sigma F h)(x,a)$, which shows that $F \circ Q_\sigma = M_\sigma \circ F$ on $\mathbb{R}^\mathsf{G}$. As a consequence, $\mathscr{A}$ and $\hat{\mathscr{A}}$ are isomorphic when $\theta > 0$ and anti-isomorphic when $\theta < 0$.

Appendix C. Remaining Proofs

C.1. **Proof of Optimality Results.** Let $\mathscr{A} = (V, \{T_\sigma\})$ be an ADP. When max-greedy policies exist, we let $T_\vartriangle$ be the Bellman max-operator and $H_\vartriangle$ be the Howard max-operator. As above, we denote the greatest element of V_Σ by $v_\vartriangle$ whenever it exists.

Lemma C.1. *If every $v \in V$ has at least one max-greedy policy, then the following statements are true:*

(i) *$H_\vartriangle$ obeys $v_\sigma \preceq H_\vartriangle v_\sigma$ for all $\sigma \in \Sigma$.*
(ii) *If $\sigma \in \Sigma$ and $T v_\sigma = v_\sigma$, then $v_\vartriangle$ exists in V and $v_\sigma = v_\vartriangle$.*
(iii) *If $v \in V$ and $H_\vartriangle v = v$, then $v = v_\vartriangle$ and $T_\vartriangle v_\vartriangle = v_\vartriangle$.*
(iv) *If $v \in V$ and Σ is finite, then $v_\vartriangle$ exists, $H_\vartriangle v_\vartriangle = v_\vartriangle$ and $(H_\vartriangle^k v)_{k \geqslant 0}$ converges to $v_\vartriangle$ in finitely many steps.*

Proof. As for (i), fix $\sigma \in \Sigma$ and let τ be such that $H_\Delta v_\sigma = v_\tau$. Since τ is v_σ-greedy, we have $v_\sigma = T_\sigma v_\sigma \preceq T_\Delta v_\sigma = T_\tau v_\sigma$. Upward stability of T_τ gives $v_\sigma \preceq v_\tau = H_\Delta v_\sigma$.

As for (ii), suppose $\sigma \in \Sigma$ and $T_\Delta v_\sigma = v_\sigma$. Fix $\tau \in \Sigma$ and note that $v_\sigma = T_\Delta v_\sigma \succeq T_\tau v_\sigma$. Downward stability of T_τ implies $v_\sigma \succeq v_\tau$. Since $\tau \in \Sigma$ was arbitrary, $v_\sigma = v_\Delta$.

As for (iii), fix $v \in V$ with $H_\Delta v = v$ and let σ be such that $H_\Delta v = v_\sigma$. Then $v_\sigma = v$, and, since σ is v-max-greedy, $T_\sigma v = T_\Delta v$. But then $T_\sigma v_\sigma = T_\Delta v_\sigma$, and, since $v_\sigma = T_\sigma v_\sigma$, we have $v_\sigma = T_\Delta v_\sigma$. Part (ii) now implies $v = v_\sigma = v_\Delta$. This proves the first claim. Regarding the second, substituting $v_\sigma = v_\Delta$ into $v_\sigma = T_\Delta v_\sigma$ yields $v_\Delta = T_\Delta v_\Delta$.

For (iv), it suffices to show that $H_\Delta v_\Delta = v_\Delta$ and there exists a $K \in \mathbb{N}$ such that $H_\Delta^K v = v_\Delta$. To this end, let $v_k = H_\Delta^k v$ and note that $v_k \in V_\Sigma$ for all $k \geqslant 1$. Part (i) implies that $v_{k+1} \succeq v_k$ for all $k \in \mathbb{N}$. Since the sequence (v_k) is contained in the finite set V_Σ, it must be that $v_{K+1} = v_K$ for some $K \in \mathbb{N}$ (since otherwise V_Σ contains an infinite sequence of distinct points). But then $H_\Delta v_K = v_{K+1} = v_K$, so v_K is a fixed point of H_Δ. Part (iii) now implies that $v_K = v_\Delta$. □

Proof of Proposition 3.1. If $\mathscr{A}$ is an ADP such that max-greedy policies exist and Σ is finite, then, by (iii)–(iv) of Lemma C.1, the point v_Δ is a fixed point of T_Δ. This proves the max version of Proposition 3.1. The proof of the min version is analogous. □

Lemma C.2. *If $\mathscr{A}$ is max-stable, then the following statements hold.*

(i) *V_Σ has a greatest element v_Δ and*
(ii) *v_Δ is the unique fixed point of T_Δ in V.*
(iii) *a policy is max-optimal if and only if it is v_Δ-max-greedy.*
(iv) *at least one optimal policy exists.*

Proof. As for parts (i)–(ii), we observe that, by max-stability, T_Δ has a fixed point $\bar{v}$ in V. By existence of max-greedy policies, we can find a $\sigma \in \Sigma$ such that $\bar{v} = T_\Delta \bar{v} = T_\sigma \bar{v}$. But T_σ has a unique fixed point in V, equal to v_σ, so $\bar{v} = v_\sigma$. Moreover, if τ is any policy, then $T_\tau \bar{v} \preceq T_\Delta \bar{v} = \bar{v}$ and hence, by downward stability, $v_\tau \preceq \bar{v}$. These facts imply that $v_\Delta := \bar{v}$ is the greatest element of V_Σ and a fixed point of T_Δ. Since greatest elements are unique, v_Δ is the only fixed point of T_Δ in V.

For (iii), parts (i)–(ii) give $v_\Delta \in V$ and $T_\Delta v_\Delta = v_\Delta$. Now recall that σ is optimal if and only if $v_\sigma = v_\Delta$. Since v_σ is the unique fixed point of T_σ, this is equivalent to $T_\sigma v_\Delta = v_\Delta$. Since $T_\Delta v_\Delta = v_\Delta$, the last statement is equivalent to $T_\sigma v_\Delta = T_\Delta v_\Delta$, which is, in turn equivalent to the statement that σ is v_Δ-greedy.

Part (iv) follows from part (iii) and existence of a v_Δ-greedy policy. □

Proof of Theorem 4.1. Parts (i)–(iv) of Theorem 4.1 follow from Lemma C.2. The last claim follows from Lemma C.1. □

C.2. **Proofs of Isomorphism Results.**

Proof of Theorem 5.1. Claims (i)–(ii) follow directly from Lemma A.4. Regarding (iii), suppose $\mathscr{A}$ is max-stable. We claim that, for $\hat{\mathscr{A}}$, max-greedy policies always exist. To see this, fix $\hat{v} \in \hat{V}$. Since $\mathscr{A}$ is max-stable, we can choose σ to be $F^{-1}\hat{v}$-max-greedy, so that $T_\tau F^{-1}\hat{v} \preceq T_\sigma F^{-1}\hat{v}$ for all $\tau \in \Sigma$. Then $F^{-1}\hat{T}_\tau \hat{v} \preceq F^{-1}\hat{T}_\sigma \hat{v}$ and hence $\hat{T}_\tau \hat{v} \preceq \hat{T}_\sigma \hat{v}$ for all $\tau \in \Sigma$. In particular, σ is $\hat{v}$-max-greedy.

Continuing to assume that $\mathscr{A}$ is max-stable, we now prove (9). For given $v \in V$, applying the order conjugacy (8) yields

$$T_\vartriangle v = \bigvee_\sigma T_\sigma v = \bigvee_\sigma F^{-1}\hat{T}_\sigma F v = F^{-1} \bigvee_\sigma \hat{T}_\sigma F v = F^{-1}\hat{T}_\vartriangle F v,$$

which is equivalent to $F \circ T_\vartriangle = \hat{T}_\vartriangle \circ F$ from (9), and implies that $(V, T_\vartriangle)$ and $(\hat{V}, \hat{T}_\vartriangle)$ are order conjugate under F. By max-stability of $\mathscr{A}$ and Theorem 4.1, the operator $T_\vartriangle$ has unique fixed point $v_\vartriangle$ in V. Lemma A.4 then implies that $\hat{T}_\vartriangle$ has unique fixed point $Fv_\vartriangle$ in $\hat{V}$. This completes the proof that $\hat{\mathscr{A}}$ is max-stable. By max-stability of $\hat{\mathscr{A}}$, the unique fixed point of $\hat{T}_\vartriangle$ in $\hat{V}$ is $\hat{v}_\vartriangle$, so $Fv_\vartriangle = \hat{v}_\vartriangle$ and both claims in (9) are verified. Finally, note that σ is max-optimal for $\mathscr{A}$ if and only if $T_\sigma v_\vartriangle = v_\vartriangle$, which, by the bijection property of F, is also equivalent to $F T_\sigma v_\vartriangle = \hat{v}_\vartriangle$. Using $F \circ T_\sigma = \hat{T}_\sigma \circ F$, we can write this as $\hat{T}_\sigma \hat{v}_\vartriangle = \hat{v}_\vartriangle$, which is equivalent to the statement that σ is max-optimal for $\hat{\mathscr{A}}$. We have now confirmed all the claims in (iii).

The proof of (iv) is identical after replacing max with min and $\bigvee$ with $\bigwedge$. (Alternatively, the proof can be derived from (iii) and duality.) □

Proof of Theorem 5.2. Let $\mathscr{A}$ and $\hat{\mathscr{A}}$ be anti-isomorphic, so that $\mathscr{A}$ is isomorphic to $\hat{\mathscr{A}}^\partial$. If $\mathscr{A}$ is well-posed, then, by Theorem 5.1, $\hat{\mathscr{A}}^\partial$ is well-posed, so $\hat{T}_\sigma$ has a unique fixed point in $\hat{V}$ for all $\sigma \in \Sigma$. This implies that $\hat{\mathscr{A}}$ is likewise well-posed, completing the proof of (i). Similarly, if $\mathscr{A}$ is order stable, then, by Theorem 5.1, $\hat{\mathscr{A}}^\partial$ is order stable, in which case $\hat{\mathscr{A}}$ is order stable, by Lemma A.3. This proves (ii).

Now suppose $\mathscr{A}$ is max-stable. Then, by $\mathscr{A} \sim \hat{\mathscr{A}}^\partial$ and Theorem 5.1, $\hat{\mathscr{A}}^\partial$ is max-stable with $F \circ T_\vartriangle = \hat{T}^\partial_\vartriangle \circ F$ and $\hat{v}^\partial_\vartriangle = F v_\vartriangle$. As with our discussion of duality in Appendix C.3, this is equivalent to $F \circ T_\vartriangle = \hat{T}_\triangledown \circ F$ and $\hat{v}_\triangledown = F v_\vartriangle$, which proves (11).

Finally, Theorem 5.1 tells us that $\mathscr{A}$ and $\hat{\mathscr{A}}^\partial$ have the same max-optimal policies. Applying Lemma C.3, we see that the max-optimal policies of $\mathscr{A}$ are the same as the min-optimal policies of $\hat{\mathscr{A}}$. □

C.3. **Min-Optimality.** In the abstract setting of Section 4, minimization results are readily recovered from maximization results by order duality.

We call a policy $\sigma \in \Sigma$ *min-optimal* for $\mathscr{A}$ if v_σ is a least element of V_Σ. When V_Σ has a least element we denote it by $v_\triangledown$ and call it the *min-value function*. We define $H_\triangledown$ from V to $\{v_\sigma\}$ via $H_\triangledown \, v = v_\sigma$ where σ is v-min-greedy and call $H_\triangledown$ the *Howard min-operator* generated by $\mathscr{A}$.

Below, if $\mathscr{A} := (V, \{T_\sigma\})$ is an ADP then its *dual* $\mathscr{A}^\partial$ is the ADP $(V^\partial, \{T_\sigma\})$ where the partial order $\preceq$ on V is replaced with its dual $\preceq^\partial$. In this setting, we let $T^\partial_\vartriangle$ be the Bellman max-operator for $\mathscr{A}^\partial$, $v^\partial_\vartriangle$ be the max-value function for $\mathscr{A}^\partial$, and so on.

Lemma C.3. *$\mathscr{A}$ is min-stable if and only if $\mathscr{A}^\partial$ is max-stable, in which case $T_\triangledown = T^\partial_\vartriangle$ and $H_\triangledown = H^\partial_\vartriangle$. A policy σ is max-optimal for $\mathscr{A}$ if and only if σ is min-optimal for $\mathscr{A}^\partial$.*

Proof. Let $\mathscr{A}$ be min-stable. Then $\mathscr{A}^\partial$ is order stable, by Lemma A.3. Now fix $v \in V$ and suppose that σ is min-greedy for $\mathscr{A}$, so that $T_\sigma \, v \preceq T_\tau \, v$ for all $\tau \in \Sigma$. Then $T_\sigma \, v \succeq^\partial T_\tau \, v$ for all $\tau \in \Sigma$, so σ is v-max-greedy for $\mathscr{A}^\partial$ and $T^\partial_\vartriangle v = T_\sigma v = T_\triangledown v$. We have proved that $\mathscr{A}^\partial$ is max-stable and $T^\partial_\vartriangle = T_\triangledown$. The remaining steps follow easily from the definitions. □

Results analogous to Theorem 4.1 hold for minimization.

Theorem C.4 (Min-optimality)**.** *If $\mathscr{A}$ is min-stable, then*

(i) *the min-value function $v_\triangledown$ exists in V,*
(ii) *$v_\triangledown$ is the unique solution to the Bellman min-equation in V,*
(iii) *a policy is min-optimal if and only if it is $v_\triangledown$-min-greedy.*
(iv) *at least one min-optimal policy exists.*

If, in addition, Σ is finite, then Howard min-policy iteration converges to $v_\triangledown$ in finitely many steps.

Proof of Theorem C.4. Let $\mathscr{A}$ be min-stable. By Lemma C.3, the dual $\mathscr{A}^\partial$ is max-stable. Hence, by Theorem 4.1, $v^\partial_\vartriangle$ exists in V. But then $v_\triangledown$ exists in V and is equal to $v^\partial_\vartriangle$, since $v_\triangledown = \bigwedge_\sigma v_\sigma = \bigvee^\partial_\sigma v_\sigma = v^\partial_\vartriangle$. Also, by Theorem 4.1, $v^\partial_\vartriangle$ is the unique solution to $T^\partial_\vartriangle v^\partial_\vartriangle = v^\partial_\vartriangle$. Applying Lemma C.3, we see that $T_\triangledown \, v_\triangledown = v_\triangledown$. The remaining claims follow by analogous arguments. □

C.4. **Proofs of Epstein–Zin Optimality Results.** This section offers a proof of Theorem 6.2. To do so, we establish the following.

(C1) If $\mathcal{E}(\beta, Q, \theta)^{1/\theta} < 1$ and $\gamma < 0$, then $\hat{\mathscr{A}}$ is min-stable.
(C2) If $\mathcal{E}(\beta, Q, \theta)^{1/\theta} < 1$ and $\gamma > 0$, then $\hat{\mathscr{A}}$ is max-stable.
(C3) If $\mathcal{E}(\beta, Q, \theta)^{1/\theta} \geqslant 1$, then $\hat{\mathscr{A}}$ is not well-posed.

Together these facts establish Theorem 6.2. Indeed, if (C1) holds, then, since $\mathscr{A}$ and $\hat{\mathscr{A}}$ are anti-isomorphic (see Lemma 6.1), it follows that $\mathscr{A}$ is max-stable (Theorem 5.2). If (C2) holds, then, since $\mathscr{A}$ and $\hat{\mathscr{A}}$ are isomorphic (see Lemma 6.1), it follows that $\mathscr{A}$ is max-stable (Theorem 5.1). Finally, if (C3) holds, then $\mathscr{A}$ is also not well-posed (by Theorem 5.1 or Theorem 5.2, depending on whether $\gamma > 0$ or $\gamma < 0$.) When $\mathscr{A}$ is not well-posed, recursive utility does not exist, so the dynamic program is undefined.

In what follows, given $\sigma \in \Sigma$ we set

$$A_\sigma(x, x') := \beta(x)^\theta P(x, \sigma(x), x') \qquad (x, x' \in \mathsf{X}).$$

Also, for any linear operator B, the symbol $\rho(B)$ represents the spectral radius.

Lemma C.5. *For all $\sigma \in \Sigma$, we have $\rho(A_\sigma) = \mathcal{E}(\beta, Q, \theta)$.*

Proof. Fix $z \in \mathsf{Z}$ and let $\mathbb{1}$ be a vector of ones. An inductive argument shows that

$$(A_\sigma^k \mathbb{1})(x) = (A_\sigma^k \mathbb{1})(z) = \mathbb{E} \prod_{t=0}^{k-1} \beta(Z_t(z))^\theta. \tag{27}$$

Combining (27) with Theorem 9.1 of Krasnosel'skii et al. (1972), we have

$$\rho(A_\sigma) = \lim_{k \to \infty} \left\{ \sup_z (A_\sigma^k \mathbb{1})(z) \right\}^{1/k} = \lim_{k \to \infty} \left\{ \sup_{z \in \mathsf{Z}} \mathbb{E} \prod_{t=0}^{k-1} \beta(Z_t(z))^\theta \right\}^{1/k},$$

as was to be shown. □

Lemma C.6. *The ADP $\hat{\mathscr{A}}$ is order stable if and only if $\mathcal{E}(\beta, Q, \theta)^{1/\theta} < 1$. Moreover, if this condition fails, then $\hat{\mathscr{A}}$ is not well posed.*

Proof. Fix $\sigma \in \Sigma$ and let V and $\hat{T}_\sigma$ be as defined in Section 6.1. By Theorem 3.1 of Stachurski et al. (2025),

(i) $\rho(A_\sigma)^{1/\theta} < 1 \implies (V, \hat{T}_\sigma)$ is globally stable on V, and
(ii) $\rho(A_\sigma)^{1/\theta} \geqslant 1 \implies \hat{T}_\sigma$ has no fixed point in V.

We saw in Lemma C.5 that $\rho(A_\sigma) = \mathcal{E}(\beta, Q, \theta)$, so $\mathcal{E}(\beta, Q, \theta)^{1/\theta} < 1$ if and only if (i) holds. In this case, $(V, \hat{T}_\sigma)$ is globally stable and hence order stable (by Example 2.1). Therefore $\hat{\mathscr{A}}$ is order stable.

If, on the other hand, $\mathcal{E}(\beta, Q, \theta)^{1/\theta} \geqslant 1$, then (ii) holds and $\hat{\mathscr{A}}$ is not well-posed (and therefore not order stable). □

Now we return to (C1)–(C3) above. Assume the conditions in (C1). Then $\hat{\mathscr{A}}$ is order stable by Lemma C.6. Also, for $v \in V$, we construct a v-min-greedy policy σ by taking

$$\sigma(x) \in \arg\min \left\{ r(x,a)^\alpha + \beta(x) \left[\sum_{x' \in \mathsf{X}} v(x') P(x,a,x') \right]^{1/\theta} \right\}^\theta$$

for all $x \in \mathsf{X}$. Since the policy set is finite, Proposition 3.1 implies that $\hat{\mathscr{A}}$ is min-stable. Hence (C1) holds. The proof of (C2) is analogous. Finally, (C3) follows directly from Lemma C.6.

References

Albuquerque, R., Eichenbaum, M., Luo, V. X., and Rebelo, S. (2016). Valuation risk and asset pricing. *The Journal of Finance*, 71(6):2861–2904.

Arnold, V. I. (2012). *Geometrical methods in the theory of ordinary differential equations*, volume 250. Springer Science & Business Media.

Azinovic, M., Gaegauf, L., and Scheidegger, S. (2022). Deep equilibrium nets. *International Economic Review*, 63(4):1471–1525.

Balbus, Ł., Reffett, K., and Woźny, Ł. (2022). Time-consistent equilibria in dynamic models with recursive payoffs and behavioral discounting. *Journal of Economic Theory*, page 105493.

Barillas, F. and Fernández-Villaverde, J. (2007). A generalization of the endogenous grid method. *Journal of Economic Dynamics and Control*, 31(8):2698–2712.

Battaglini, M. (2021). Chaos and unpredictability in dynamic social problems. Technical report, National Bureau of Economic Research.

Bäuerle, N. and Jaśkiewicz, A. (2018). Stochastic optimal growth model with risk sensitive preferences. *Journal of Economic Theory*, 173:181–200.

Bäuerle, N., Jaśkiewicz, A., and Nowak, A. S. (2021). Stochastic dynamic programming with non-linear discounting. *Applied Mathematics & Optimization*, 84(3):2819–2848.

Bäuerle, N. and Rieder, U. (2011). *Markov decision processes with applications to finance*. Springer Science & Business Media.

Bayer, C. and Luetticke, R. (2020). Solving discrete time heterogeneous agent models with aggregate risk and many idiosyncratic states by perturbation. *Quantitative Economics*, 11(4):1253–1288.

Bellman, R. (1957). *Dynamic programming*. American Association for the Advancement of Science.

Bertsekas, D. (2012). *Dynamic programming and optimal control*, volume 1. Athena Scientific.

Bertsekas, D. P. (2022). *Abstract dynamic programming*. Athena Scientific, 3 edition.

Bloise, G., Le Van, C., and Vailakis, Y. (2024). Do not blame Bellman: It is Koopmans' fault. *Econometrica*, 92(1):111–140.

Bloise, G. and Vailakis, Y. (2018). Convex dynamic programming with (bounded) recursive utility. *Journal of Economic Theory*, 173:118–141.

Borovička, J. and Stachurski, J. (2020). Necessary and sufficient conditions for existence and uniqueness of recursive utilities. *The Journal of Finance*.

Brumm, J. and Scheidegger, S. (2017). Using adaptive sparse grids to solve high-dimensional dynamic models. *Econometrica*, 85(5):1575–1612.

Cagetti, M., Hansen, L. P., Sargent, T. J., and Williams, N. (2002). Robustness and pricing with uncertain growth. *The Review of Financial Studies*, 15(2):363–404.

Carroll, C. D. (2006). The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics letters*, 91(3):312–320.

Christensen, T. M. (2022). Existence and uniqueness of recursive utilities without boundedness. *Journal of Economic Theory*, page 105413.

Davey, B. A. and Priestley, H. A. (2002). *Introduction to lattices and order*. Cambridge University Press.

de Castro, L., Galvao, A. F., and Nunes, D. (2022). Dynamic economics with quantile preferences. *Available at SSRN 4108230*.

de Groot, O., Richter, A. W., and Throckmorton, N. A. (2018). Uncertainty shocks in a model of effective demand: Comment. *Econometrica*, 86(4):1513–1526.

Denardo, E. V. (1967). Contraction mappings in the theory underlying dynamic programming. *Siam Review*, 9(2):165–177.

Deng, L., Khan, M. A., and Mitra, T. (2022). Continuous unimodal maps in economic dynamics: On easily verifiable conditions for topological chaos. *Journal of Economic Theory*, 201:105446.

Epstein, L. G. and Zin, S. E. (1989). Risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969.

Farahmand, A.-M., Munos, R., and Szepesvári, C. (2010). Error propagation for approximate policy and value iteration. In Lafferty, J. D., Williams, C. K. I., Shawe-Taylor, J., Zemel, R. S., and Culotta, A., editors, *Advances in Neural Information Processing Systems 23*, pages 568–576.

Fedus, W., Gelada, C., Bengio, Y., Bellemare, M. G., and Larochelle, H. (2019). Hyperbolic discounting and learning over multiple horizons.

Fei, Y., Yang, Z., Chen, Y., and Wang, Z. (2021). Exponential Bellman equation and improved regret bounds for risk-sensitive reinforcement learning. *Advances in Neural Information Processing Systems*, 34:20436–20446.

Flynn, J. P. and Sastry, K. (2022). *The macroeconomics of narratives*. SSRN.

Foerster, A., Rubio-Ramírez, J. F., Waggoner, D. F., and Zha, T. (2016). Perturbation methods for markov-switching dynamic stochastic general equilibrium models. *Quantitative economics*, 7(2):637–669.

Gao, Y., Lui, K. Y. C., and Hernandez-Leal, P. (2021). Robust risk-sensitive reinforcement learning agents for trading markets. Technical report, arXiv preprint arXiv:2107.08083.

Gardini, L., Hommes, C., Tramontana, F., and De Vilder, R. (2009). Forward and backward dynamics in implicitly defined overlapping generations models. *Journal of Economic Behavior & Organization*, 71(2):110–129.

Gomez-Cram, R. and Yaron, A. (2020). How Important Are Inflation Expectations for the Nominal Yield Curve? *The Review of Financial Studies*, 34(2):985–1045.

Hansen, L. P. and Sargent, T. J. (2011). *Robustness*. Princeton university press.

Hansen, L. P. and Scheinkman, J. A. (2012). Recursive utility in a Markov environment with stochastic growth. *Proceedings of the National Academy of Sciences*, 109(30):11967–11972.

Hernández-Lerma, O. and Lasserre, J. B. (2012). *Discrete-time Markov control processes: basic optimality criteria*, volume 30. Springer Science & Business Media.

Howard, R. A. and Matheson, J. E. (1972). Risk-sensitive Markov decision processes. *Management Science*, 18(7):356–369.

Judd, K. L. (1992). Projection methods for solving aggregate growth models. *Journal of Economic theory*, 58(2):410–452.

Judd, K. L. (1996). Approximation, perturbation, and projection methods in economic analysis. *Handbook of computational economics*, 1:509–585.

Kennedy, J. A. and Stockman, D. R. (2008). Chaotic equilibria in models with backward dynamics. *Journal of Economic Dynamics and Control*, 32(3):939–955.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kochenderfer, M. J., Wheeler, T. A., and Wray, K. H. (2022). *Algorithms for decision making*. The MIT Press.

Krasnosel'skii, M. A., Vainikko, G. M., Zabreiko, P. P., Rutitskii, Y. B., and Stetsenko, V. Y. (1972). *Approximate Solution of Operator Equations*. Springer Netherlands.

Kristensen, D., Mogensen, P. K., Moon, J. M., and Schjerning, B. (2021). Solving dynamic discrete choice models using smoothing and sieve methods. *Journal of Econometrics*, 223(2):328–360.

Layek, G. (2015). *An Introduction to Dynamical Systems and Chaos*. Springer.

Long, J. B. and Plosser, C. I. (1983). Real business cycles. *Journal of Political Economy*, 91(1):39–69.

Ma, Q., Stachurski, J., and Toda, A. A. (2022). Unbounded dynamic programming via the q-transform. *Journal of Mathematical Economics*, 100:102652.

Maliar, L., Maliar, S., and Winant, P. (2021). Deep learning for solving dynamic economic models. *Journal of Monetary Economics*, 122:76–101.

Marinacci, M. and Montrucchio, L. (2010). Unique solutions for stochastic recursive utilities. *Journal of Economic Theory*, 145(5):1776–1804.

Marinacci, M. and Montrucchio, L. (2019). Unique tarski fixed points. *Mathematics of Operations Research*, 44(4):1174–1191.

Puterman, M. L. (2005). *Markov decision processes: discrete stochastic dynamic programming*. Wiley Interscience.

Raines, B. and Stockman, D. (2012). Fixed points imply chaos for a class of differential inclusions that arise in economic models. *Transactions of the American Mathematical Society*, 364(5):2479–2492.

Rincón-Zapatero, J. P. (2024). Existence and uniqueness of solutions to the bellman equation in stochastic dynamic programming. *Theoretical Economics*, 19(3):1223–1260.

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

Sargent, T. J. and Stachurski, J. (2025). Dynamic programming on partially ordered sets. *SIAM Journal on Control and Optimization*, in press.

Schorfheide, F., Song, D., and Yaron, A. (2018). Identifying long-run risks: A Bayesian mixed-frequency approach. *Econometrica*, 86(2):617–654.

Smale, S. (1967). Differentiable dynamical systems. *Bulletin of the American Mathematical Society*, 73(6):747–817.

Stachurski, J., Wilms, O., and Zhang, J. (2025). Unique solutions to power-transformed affine systems. *Journal of Mathematical Analysis and Applications*, 550(1):129515.

Stachurski, J. and Zhang, J. (2021). Dynamic programming with state-dependent discounting. *Journal of Economic Theory*, 192:105190.

Sternberg, S. (2014). *Dynamical Systems*. Dover Publications.

Stokey, N. L. and Lucas, R. E. (1989). *Recursive methods in dynamic economics*. Harvard University Press.

Weil, P. (1990). Nonexpected utility in macroeconomics. *The Quarterly Journal of Economics*, 105(1):29–42.

National Graduate Institute for Policy Studies

Email address: john.stachurski@gmail.com

Center for Macroeconomic Research and Wang Yanan Institute for Studies in Economics at Xiamen University, China

Email address: zhangjunnan1224@gmail.com