Dynamic Programming with Recursive Preferences: Optimality and Applications[1]

Guanlong Ren and John Stachurski

Research School of Economics, Australian National University

December 17, 2018

ABSTRACT. This paper provides an alternative approach to the theory of dynamic programming, designed to accommodate the recursive preference specifications commonly used in modern economic analysis while still supporting traditional additively separable rewards. The approach exploits the theory of monotone convex operators, which turns out to be well suited to dynamic maximization. The intuition is that convexity is preserved under maximization, so convexity properties found in preferences extend naturally to the Bellman operator.

JEL Classifications: C61, E00
Keywords: Dynamic programming, optimality

1. INTRODUCTION

Through analysis of observational data and the design of suitable experiments, economists have constructed progressively more realistic representations of agents and their choices. For intertemporal decisions, this typically involves a departure from the additively separable benchmark. A familiar example is the recursive preference framework of Epstein and Zin (1989), which has become central to the quantitative asset pricing literature, while also finding widespread use

[1]Thanks to be added. guanlong.ren@anu.edu.au, john.stachurski@anu.edu.au

in applications ranging from optimal taxation to fiscal policy and business cycles.[2] Others include the robust control model of Hansen and Sargent (2008) and the recursive smooth ambiguity model of Ju and Miao (2012), which successfully generate features of the data that fail to arise under more traditional preferences.[3]

At the same time, the underlying theory of optimal choice in the presence of such specifications remains less well understood than the classic, additively separable framework of Bellman (1957) and Blackwell (1965). While early attempts to treat recursive preferences in a dynamic programming framework continued to use the contraction mapping arguments that had been successful for additively separable models (see, e.g., Lucas and Stokey (1984)), it was soon realized that the Bellman operators generated by the most common recursive preference specifications are not supremum norm contractions.[4]

This realization drove a second wave of theoretical analysis, built instead around a certain forms of monotonicity and concavity exhibited by many intertemporal preferences. For example, Marinacci and Montrucchio (2010) exploit monotonicity and concavity to obtain a range of deep results on existence and uniqueness of recursive utilities. Le Van et al. (2008) adapted the theory of monotone concave operators, as pioneered by Krasnoselskii (1964) and coauthors, to dynamic programming problems. Marinacci and Montrucchio (2017) and Bloise and Vailakis (2018) further extended these ideas. Like contraction maps, under

[2]A sample of the literature can be found in Bansal and Yaron (2004), Hansen and Sargent (2008), Farhi and Werning (2008), Hansen and Scheinkman (2012), Kaplan and Violante (2014), Gottardi et al. (2015), Bhandari et al. (2017), Basu and Bundick (2017), Schorfheide et al. (2018), Petrosky-Nadeau et al. (2018) and Karantounias (2018).

[3]Other valuable contributions to the literature on recursive smooth ambiguity preferences include Klibanoff et al. (2009) and Hayashi and Miao (2011). An overview of recursive preferences can be found in Backus et al. (2004).

[4]In addition to Lucas and Stokey (1984), related work can be found in Boyd (1990), Durán (2003), Le Van and Vailakis (2005) and Rincón-Zapatero and Rodríguez-Palmero (2007). It was Marinacci and Montrucchio (2010) who emphasized that sup norm contractivity fails for many economically reasonable aggregators, such as Thompson aggregators.

certain regularity conditions, monotone concave operators have unique, globally attracting fixed points—a highly attractive property in the context of dynamic programming.[5]

However, while we now have a good understanding of existence and uniqueness of recursive utilities—that is, results showing that the preference specifications are well defined at a fixed consumption path or policy—our understanding of optimality in the context of recursive preferences is far less complete. Despite important contributions, foundations have been lacking for some of the most popular specifications for applied work, including certain commonly used parameterizations of the Epstein–Zin specification, many robust control models, or the ambiguity sensitive preferences discussed above. For example, it has not previously been determined whether value function iteration is valid in a number of important cases, or whether Bellman's principle of optimality holds true.

In this paper we resolve many of these outstanding problems by developing a set of sufficient conditions for abstract dynamic programs—including both additively separable and recursive preference models—that provide global convergence of the Bellman operator to the value function and optimality of the associated policies. These conditions are applied to a range of recursive preference specifications popular in applied settings, including standard Epstein–Zin models with the constant elasticity of substitution aggregators, risk-sensitive and robust control models, and ambiguity sensitive preferences such as those proposed by Ju and Miao (2012).

Our approach builds on the monotone concave approach but with one significant difference: the relevant operators are convex. Put differently, we use monotone *convex* operators to study the maximization problems associated with dynamic programming. The main benefit is that, unlike concavity, convexity is

[5]This property was used the show existence of Markov equilibria in the presence of distortions in Datta et al. (2002), Morand and Reffett (2003) and several related papers. Additional results on existence and uniqueness of recursive utility via monotone concave operators can be found in Becker and Rincón-Zapatero (2017). Alternative approaches that admit unbounded aggregators are studied in Rincón-Zapatero and Rodríguez-Palmero (2007) and Martins-da Rocha and Vailakis (2013).

preserved under the taking of pointwise suprema. Hence convexity pairs naturally with maximization. Moreover, under suitable conditions, monotone convex operators enjoy all the stability properties possessed by monotone concave operators.

At the same time, we find that the theory of monotone concave operators is ideal for minimization problems. This is because concavity is preserved by minimization, in the sense that the infimum of a family of concave functions is concave. Thus, any concavity inherent in the dynamic program flows naturally into the Bellman operator.

As the last piece of this puzzle, we note that simple continuous transformations can be used to transform inherently concave problems into convex problems and vice versa. Through these transformations, one can shift between convex maximization problems and concave minimization problems on a case by case basis. In particular, we show how preference specifications that have been recognized as concave can be modified so that they exhibit convexity rather than concavity. We apply all of the ideas described above to a variety of well-known models, including Epstein–Zin preferences, risk sensitive preferences and recursive smooth ambiguity. In each case we show that value function iteration converges to the value function and that Bellman's principle of optimality is valid.

As one extension of our ideas, we also consider an Epstein–Zin recursive preference model with unbounded rewards. Problems with unbounded rewards are difficult to treat in a systematic way. Important innovations tailored to economic applications can be found in Durán (2003), Le Van and Vailakis (2005), Rincón-Zapatero and Rodríguez-Palmero (2007), Martins-da Rocha and Vailakis (2010) and Bäuerle and Jaśkiewicz (2018). Perhaps the most common approach to treating unbounded rewards in the broader field of dynamic programming has been one involving contraction mapping arguments in a setting of weighted supremum norms (see, e.g., Bertsekas (2013)). Here we show that similar ideas can be applied when contractivity fails or is difficult to obtain. In particular, we show how one of the results from the preceding sections can be adapted to accommodate unbounded rewards.

One of the most closely related studies to ours is Bloise and Vailakis (2018), who analyze dynamic programming problems with bounded recursive utility. By exploiting the theory of monotone concave operators, they prove significant optimality results. While their paper is carefully constructed and serves as an inspiration for our work, it does not address several popular specifications in the applied literature. For example, their condition $W(c, 0) > 0$ on the aggregator fails in the CES setting with elasticity of intertemporal substitution less than 1 (a parametrization that is not uncommon—see, e.g., Basu and Bundick (2017)). In addition, models of ambiguity aversion are not tackled in their paper.[6]

Another related paper is Guo and He (2017), who consider an extension to the Epstein-Zin recursive utility model that includes utility measures for investment gains and losses. As a part of their study, they obtain results for existence, uniqueness and successive approximations of the solution to the Bellman equation for a portfolio selection problem with gain-loss utility. While they provide sharp results under the assumption that the state space is finite and the exogenous state process is irreducible, our results are significantly broader, applying as they do to a large range of settings outside of portfolio selection problem within the framework of Epstein and Zin (1989).

Also related is Balbus (2016), who considers a class of non-negative recursive utilities with certain types of nonlinear aggregators and certainty equivalent operators, and studies the corresponding dynamic programming problem. His results for existence, uniqueness and convergence of solutions to recursive utilities and to the Bellman equation rest upon the theory of monotone α-concave operators. While these results are valuable in some instances, the requisite α-concave property does not hold for a number of popular specifications of recursive utility, such as those of Epstein and Zin (1989) when the elasticity of intertemporal substitution differs from unity.[7]

[6]Also valuable for dynamic programming with recursive preferences is Marinacci and Montrucchio (2017), which develops new methods for determining when isotone maps have unique fixed points. Their treatment of dynamic programming in the context of recursive preferences also assumes that $W(c, 0) > 0$.

[7]See, e.g., page 8 of Balbus (2016), as well as the analogous discussion in Le Van et al. (2008).

Ozaki and Streufert (1996) provide a comprehensive study of the recovery of recursive preferences and the corresponding dynamic programming problem under a non-Markovian environment. While useful for studying dynamic programming for non-additive stochastic objectives in a much more general setting, their assumptions also exclude some popular specifications. For example, with Epstein–Zin preferences and elasticity of intertemporal substitution greater than one (as found in, say, Kaplan and Violante (2014)), the conditions of Theorem D fail, since the parameters related to the variable discount factor $\bar{\delta}$ and δ are infinite.[8]

We note that our theoretical framework departs from the separate specification of aggregator and certainty equivalent that has been popular in the economic literature since Kreps and Porteus (1978). Instead we adopt the abstract dynamic programming framework developed and collated by Bertsekas (2013). In abstract dynamic programming, the most cohesive sufficient conditions are still driven by contractions or semi-contractive properties (see, e.g., Bertsekas (2013), chapters 2–3). The monotone-convex and monotone-concave results set out below offer an alternative branch of cohesive and broadly applicable methods.

The remainder of the paper is structured as follows: Section 2 contains our main results. Section 3 has applications. Section 4 considers an extension (unbounded rewards). Section 5 concludes. Apart from some simple arguments, all proofs are deferred to the appendix.

2. General Results

Let X and A be separable metric spaces, called the *state* and *action space* respectively. Let $\mathbb{R}^{\mathsf{X}}$ represent all functions from X to $\mathbb{R}$ and let $\|\cdot\|$ denote the supremum norm on the bounded functions in $\mathbb{R}^{\mathsf{X}}$. For f and g in $\mathbb{R}^{\mathsf{X}}$, the statement $f \leqslant g$ means $f(x) \leqslant g(x)$ for all $x \in \mathsf{X}$. Let Γ be a nonempty correspondence from X to A, referred to below as the *feasible correspondence*. We understand $\Gamma(x)$

[8]Recently, based on the biconvergence technique, Bich et al. (2018) establish existence, uniqueness and computation of the solution to the Bellman equation for deterministic dynamic programming problems under certain types of continuity properties imposed on temporal aggregators.

as representing all actions available to the controller in state x. The correspondence Γ in turn defines the set of *feasible state-action pairs*

$$\mathbb{G} := \{(x, a) \in \mathsf{X} \times \mathsf{A} : a \in \Gamma(x)\}.$$

Let

- w_1 and w_2 be bounded continuous functions in $\mathbb{R}^{\mathsf{X}}$ satisfying $w_1 \leqslant w_2$,
- $\mathscr{V}$ be all Borel measurable functions v in $\mathbb{R}^{\mathsf{X}}$ satisfying $w_1 \leqslant v \leqslant w_2$, and
- $\mathscr{C}$ be the continuous functions in $\mathscr{V}$.

Both $\mathscr{V}$ and $\mathscr{C}$ are understood as classes of candidate value functions. The functions w_1 and w_2 serve as lower and upper bounds for lifetime value respectively. Their role will be clarified below.

Current and future payoffs are subsumed into a *state-action aggregator* Q, which maps a feasible state-action pair (x, a) and function v in $\mathscr{V}$ into a real value $Q(x, a, v)$. The interpretation of $Q(x, a, v)$ is total lifetime rewards, contingent on current action a, current state x and the use of v to evaluate future states. In other words, $Q(x, a, v)$ corresponds to the right hand side of the Bellman equation when v represents the value function.

The central role of convexity and concavity was discussed in the introduction. To implement the corresponding restrictions, we call Q *value-convex* if

$$Q(x, a, \lambda v + (1 - \lambda)w) \leqslant \lambda Q(x, a, v) + (1 - \lambda)Q(x, a, w)$$

for each $(x, a) \in \mathbb{G}$, $\lambda \in [0, 1]$ and v, w in $\mathscr{V}$. Similarly, Q will be called *value-concave* when the reverse inequality holds (i.e., when $-Q$ is value-convex). One of these restrictions will be imposed on each problem we consider.

We also impose some basic properties that will be assumed in every case:

Assumption 2.1. The following conditions hold:

(a) The feasible correspondence Γ is compact valued and continuous.
(b) The map $(x, a) \mapsto Q(x, a, v)$ is Borel measurable on $\mathbb{G}$ whenever $v \in \mathscr{V}$ and continuous on $\mathbb{G}$ whenever $v \in \mathscr{C}$.

(c) The state-action aggregator satisfies

$$v \leqslant v' \implies Q(x, a, v) \leqslant Q(x, a, v') \quad \text{for all } (x, a) \in \mathbb{G}. \tag{1}$$

(d) The functions w_1 and w_2 satisfy

$$w_1(x) \leqslant Q(x, a, w_1) \quad \text{and} \quad Q(x, a, w_2) \leqslant w_2(x) \tag{2}$$

for all (x, a) in $\mathbb{G}$.

The primary role of conditions (a) and (b) is to obtain existence of solutions. If the state and action space are discrete (finite or countably infinite) then we adopt the discrete topology, in which case the continuity requirements in (a) and (b) are satisfied automatically, while the compactness requirement on Γ is satisfied if $\Gamma(x)$ is finite for each x.

Condition (c) imposes the natural requirement that higher continuation values increase lifetime values, while condition (d) is a consistency requirement that allows w_1 and w_2 to act as lower and upper bounds for lifetime value. The conditions in assumption 2.1 are held to be true throughout the remainder of the paper.

Let Σ be a family of maps from X to A, referred to below as the set of all *feasible policies*, such that each $\sigma \in \Sigma$ is Borel measurable and satisfies $\sigma(x) \in \Gamma(x)$ for all $x \in \mathsf{X}$.

Lemma 2.1. *The map $w(x) := Q(x, \sigma(x), v)$ is an element of $\mathscr{V}$ for all $v \in \mathscr{V}$.*

Proof. Borel measurability of $(x, a) \mapsto Q(x, a, v)$ and σ imply that w is Borel measurable on X. Moreover, since $w_1 \leqslant v$, the inequalities in (1) and (2) imply $w_1(x) \leqslant Q(x, \sigma(x), w_1) \leqslant Q(x, \sigma(x), v)$ for all x. In particular, $w_1 \leqslant w$. A similar argument gives $w \leqslant w_2$, so $w \in \mathscr{V}$. $\square$

Given $\sigma \in \Sigma$, a function $v_\sigma \in \mathscr{V}$ that satisfies

$$v_\sigma(x) = Q(x, \sigma(x), v_\sigma) \quad \text{for all } x \in \mathsf{X} \tag{3}$$

is called a *σ-value function*. The value $v_\sigma(x)$ can be interpreted as the lifetime value of following policy σ. Its existence and uniqueness is discussed below.

2.1. **Maximization.** We begin by studying maximization of value. Our key assumption is that the state-action aggregator satisfies value-convexity and possesses a strict upper solution:

Assumption 2.2 (Convex Program)**.** The following conditions are satisfied:

(a) Q is value-convex.
(b) There exists an $\varepsilon > 0$ such that $Q(x, a, w_2) \leqslant w_2(x) - \varepsilon$ for all $(x, a) \in \mathbb{G}$.

Note that part (b) is a strengthening of one of the conditions in (2).

Proposition 2.2. *If assumption 2.2 holds, then, for each σ in Σ, the set $\mathscr{V}$ contains exactly one σ-value function v_σ.*

Proposition 2.2 assures us that the value v_σ of a given policy σ is well defined. From this foundation we can introduce optimality concerning a maximization decision problem. In particular, in the present setting, a policy $\sigma^* \in \Sigma$ is called *optimal* if

$$v_{\sigma^*}(x) \geqslant v_\sigma(x) \quad \text{for all } \sigma \in \Sigma \text{ and all } x \in \mathsf{X}.$$

The *maximum value function* associated with this planning problem is the map v^* defined at $x \in \mathsf{X}$ by

$$v^*(x) = \sup_{\sigma \in \Sigma} v_\sigma(x). \tag{4}$$

One can show from conditions (c) and (d) of assumption 2.1 that v^* is well defined as a real valued function on X and satisfies $w_1 \leqslant v^* \leqslant w_2$.

A function $v \in \mathscr{V}$ is said to satisfy the *Bellman equation* if

$$v(x) = \max_{a \in \Gamma(x)} Q(x, a, v) \quad \text{for all } x \in \mathsf{X}. \tag{5}$$

The *Bellman operator* T associated with our abstract dynamic program is a map sending v in $\mathscr{C}$ into

$$Tv(x) = \max_{a \in \Gamma(x)} Q(x, a, v). \tag{6}$$

Since v is in $\mathscr{C}$, existence of the maximum is guaranteed by assumption 2.1. It follows from Berge's theorem of the maximum that Tv is an element of $\mathscr{C}$. Evidently solutions to the Bellman equation in $\mathscr{C}$ exactly coincide with fixed points of T.

The convex program conditions lead to the following central result:

Theorem 2.3. *If assumption 2.2 holds, then*

(a) *The Bellman equation has exactly one solution in $\mathscr{C}$ and that solution is v^*.*
(b) *If v is in $\mathscr{C}$, then $T^n v \to v^*$ uniformly on X as $n \to \infty$.*
(c) *A policy σ in Σ is optimal if and only if*

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} Q(x, a, v^*) \;\; \textit{for all } x \in \mathsf{X}.$$

(d) *At least one optimal policy exists.*

The fixed point and convergence results for T in theorem 2.3 rely on a fixed point theorem for isotone convex operators due to Du (1989), reprinted in section 2.1 of Zhang (2012). In those references, convergence is shown to be uniformly geometric, in the sense that there exist constants $\lambda \in (0, 1)$ and $K \in \mathbb{R}$ such that

$$\|T^n v - v^*\| \leqslant \lambda^n K \text{ for all } n \in \mathbb{N} \text{ and } v \in \mathscr{C}.$$

2.2. **Minimization.** Next we treat minimization. In this setting, the convexity and strict upper solution in assumption 2.2 are replaced by concavity and a strict lower solution.

In order to maintain consistency with other sources, we admit some overloading of terminology relative to section 2.1 on maximization. For example, the optimal policy will now reference a minimizing policy rather than a maximizing one, and the Bellman equation will shift from maximization to minimization. The relevant definition will be clear from context.

The next assumption is analogous to assumption 2.2, which was used for maximization.

Assumption 2.3 (Concave Program)**.** The following conditions are satisfied:

(a) Q is value-concave.
(b) There exists an $\varepsilon > 0$ such that $Q(x, a, w_1) \geqslant w_1(x) + \varepsilon$ for all $(x, a) \in \mathbb{G}$.

Note that part (b) is a strengthening of one of the conditions in (2).

Proposition 2.4. *If assumption 2.3 holds, then, for each σ in Σ, the set $\mathscr{V}$ contains exactly one σ-value function v_σ.*

Proposition 2.4 mimics proposition 2.2, assuring us that, in the present context, the cost v_σ of a given policy σ is well defined. A policy $\sigma^* \in \Sigma$ is then called *optimal* if

$$v_{\sigma^*}(x) \leqslant v_\sigma(x) \quad \text{for all } \sigma \in \Sigma \text{ and all } x \in \mathsf{X}.$$

The *minimum cost function* associated with this planning problem is the function v^* defined at $x \in \mathsf{X}$ by

$$v^*(x) = \inf_{\sigma \in \Sigma} v_\sigma(x). \tag{7}$$

A function $v \in \mathscr{V}$ is said to satisfy the *Bellman equation* if

$$v(x) = \min_{a \in \Gamma(x)} Q(x, a, v) \qquad \text{for all } x \in \mathsf{X}. \tag{8}$$

The *Bellman operator* S associated with our abstract dynamic program is a map sending v in $\mathscr{C}$ into

$$Sv(x) = \min_{a \in \Gamma(x)} Q(x, a, v). \tag{9}$$

Analogous to theorem 2.3, we have

Theorem 2.5. *If assumption 2.3 holds, then*

(a) The Bellman equation (8) has exactly one solution in $\mathscr{C}$ and that solution is the minimum cost function v^*.
(b) If v is in $\mathscr{C}$, then $S^n v \to v^*$ uniformly on X as $n \to \infty$.
(c) A policy σ in Σ is optimal if and only if

$$\sigma(x) \in \operatorname*{argmin}_{a \in \Gamma(x)} Q(x, a, v^*) \;\; \textit{for all } x \in \mathsf{X}.$$

(d) At least one optimal policy exists.

3. Applications

In this section we study a collection of applications, showing how the general results in section 2 can be used to solve the dynamic programming problems discussed in the introduction.

3.1. **An Additively Separable Decision Process.** Before treating more sophisticated preference specifications, it is worth noting that the results stated above can be applied in the standard additive separable case, alongside the more traditional Bellman–Blackwell contraction mapping approach to dynamic programming. To see this, consider the generic additively separable dynamic programming model of Stokey et al. (1989) with Bellman equation

$$v(s,z) = \max_{y \in \Gamma(s,z)} \left\{ F(s,y,z) + \beta \int v(y,z') P(z, \mathrm{d}z') \right\} \tag{10}$$

over $(s,z) \in \mathsf{S} \times \mathsf{Z}$. Here S and Z are compact metric spaces containing possible values for the endogenous and exogenous state variables, respectively. Let the transition function P on Z have the Feller property, let the feasible correspondence $\Gamma \colon \mathsf{S} \times \mathsf{Z} \to \mathsf{S}$ be compact valued and continuous, let $F \colon \mathbb{G} \to \mathbb{R}$ be continuous, and let β lie in $(0,1)$.

We translate this model to our environment by taking $x := (s,z)$ to be the state, $\mathsf{X} := \mathsf{S} \times \mathsf{Z}$ to be the state space, $a = y \in \mathsf{S}$ to be the action, and setting

$$Q((s,z), y, v) = F(s,y,z) + \beta \int v(y,z') P(z, \mathrm{d}z').$$

Since F is continuous on a compact set, there exists a finite constant M with $|F| \leqslant M$.[9] For the bracketing functions w_1 and w_2 we fix $\varepsilon > 0$ and adopt the constant functions

$$w_1 \equiv -\frac{M}{1-\beta} \quad \text{and} \quad w_2 \equiv \frac{M+\varepsilon}{1-\beta}.$$

The conditions of assumption 2.1 are all satisfied. Conditions (a) and (b) are true by assumption and condition (c) is trivial to verify. To see that condition (d) of assumption 2.1 holds, we note that w_1 and w_2 lie in $bc\mathsf{X}$. In addition, for any given $((s,z),y) \in \mathbb{G}$, we have

$$Q((s,z), y, w_1) = F(s,y,z) - \beta \frac{M}{1-\beta} \geqslant -M - \beta \frac{M}{1-\beta} = w_1(s,z).$$

Similarly,

$$Q((s,z), y, w_2) = F(s,y,z) + \beta \frac{M+\varepsilon}{1-\beta} \leqslant M + \beta \frac{M+\varepsilon}{1-\beta} = w_2(s,z) - \varepsilon.$$

[9] The domain $\mathbb{G}$ of F is compact in the product topology by Tychonoff's theorem.

The last inequality gives not only $Q(x, a, w_2) \leqslant w_2(x)$, as required for part (d) of assumption, but also the stronger condition in part (b) of assumption 2.2. Thus, to verify the requirements of theorem 2.3, we need only check the convexity condition in part (a) of assumption 2.2. But this is immediate from the linearity of expectations. Hence theorem 2.3 applies.

3.2. **Epstein-Zin Preferences.** Kreps and Porteus (1978) and Epstein and Zin (1989) propose an alternative specification of lifetime value that separates and independently parameterizes intertemporal elasticity of substitution and risk aversion. To be more precise, Epstein and Zin (1989) propose the following preferences that is defined recursively by the CES aggregator

$$U_t = \left[(1-\beta)C_t^{1-\rho} + \beta \left\{\mathcal{R}_t(U_{t+1})\right\}^{1-\rho}\right]^{\frac{1}{1-\rho}} \quad (0 < \rho \neq 1),$$

where $\{C_t\}$ is a consumption path, U_t is the utility value of the path onward from time t, and $\mathcal{R}_t$ is the Kreps-Porteus certainty equivalent operator

$$\mathcal{R}_t(U_{t+1}) = \left(\mathbb{E}_t U_{t+1}^{1-\gamma}\right)^{\frac{1}{1-\gamma}} \quad (0 < \gamma \neq 1).$$

Here, $\mathbb{E}_t$ stands for the conditional expectation with respect to the period t information. The value $1/\rho$ represents elasticity of intertemporal substitution (EIS) between the composite good and the certainty equivalent, while γ governs the level of relative risk aversion (RRA) with respect to atemporal gambles. The most empirically relevant case is $\rho < \gamma$, implying that the agent prefers early resolution of uncertainty. We focus on this case in what follows.

Under Epstein–Zin preferences, the generic additively separable Bellman equation in (10) becomes

$$v(s,z) = \max_{y \in \Gamma(s,z)} \left\{ r(s,y,z) + \beta \left[\int v(y,z')^{1-\gamma} P(z, \mathrm{d}z') \right]^{\frac{1-\rho}{1-\gamma}} \right\}^{\frac{1}{1-\rho}} \tag{11}$$

for each $(s,z) \in \mathsf{S} \times \mathsf{Z}$, where, here and below,

$$r(s,y,z) := (1-\beta) F(s,y,z)^{1-\rho}$$

We impose the same conditions on the primitives discussed in section 3.1. In particular, F is continuous, P is Feller, Γ is continuous and compact valued and both S and Z are compact. To ensure $F(s, y, z)^{1-\rho}$ is always well defined, we also assume that F is everywhere positive.

3.2.1. *The Case* $\rho < \gamma < 1$. As in Hansen and Scheinkman (2012), we begin with the continuous strictly increasing transformation $\hat{v} = v^{1-\gamma}$, which allows us to rewrite (11) as

$$\hat{v}(s, z) = \max_{y \in \Gamma(s,z)} \left\{ r(s, y, z) + \beta \left[\int \hat{v}(y, z') P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^{\theta} \tag{12}$$

where

$$\theta := \frac{1-\gamma}{1-\rho}.$$

Since this transformation is bijective, there is a one-to-one correspondence between v and $\hat{v}$, in the sense that v solves (11) if and only if $\hat{v}$ solves (12). Note that in the current setting we have $\theta \in (0, 1)$.

The state-action aggregator Q corresponding to (12) is

$$Q((s, z), y, v) = \left\{ r(s, y, z) + \beta \left[\int v(y, z') P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^{\theta}. \tag{13}$$

For the bracketing functions w_1 and w_2, we fix $\delta > 0$ and take the constant functions

$$w_1 := \left(\frac{m}{1-\beta} \right)^{\theta} \quad \text{and} \quad w_2 := \left(\frac{M+\delta}{1-\beta} \right)^{\theta},$$

where

$$m := \min_{((s,z),y) \in \mathbb{G}} r(s, y, z) \;\text{ and }\; M := \max_{((s,z),y) \in \mathbb{G}} r(s, y, z). \tag{14}$$

These values are finite and positive, since F is continuous and positive on a compact domain.[10] Being constant, w_1 and w_2 are continuous.

We now show that the conditions of assumptions 2.1 and 2.2 are all satisfied. Regarding assumption 2.1, condition (a) is true by assumption, while condition (b) follows immediately from the continuity imposed on F and the Feller property of P. Condition (c) is easy to verify, since, for any $b > 0$, the scalar map

$$\psi(t) := (b + \beta t^{1/\theta})^{\theta} \qquad (t \geqslant 0) \tag{15}$$

[10] In this case, positivity of F can be weakened to nonnegativity.

is monotone increasing. To check condition (d), observe that, for fixed $((s,z),y) \in \mathbb{G}$, we have

$$Q((s,z),y,w_1) = \left\{ r(s,y,z) + \beta \frac{m}{1-\beta} \right\}^\theta \geqslant \left\{ m + \beta \frac{m}{1-\beta} \right\}^\theta = w_1(s,z).$$

Similarly,

$$Q((s,z),y,w_2) = \left\{ r(s,y,z) + \beta \frac{M+\delta}{1-\beta} \right\}^\theta \leqslant \left\{ M + \beta \frac{M+\delta}{1-\beta} \right\}^\theta ,$$

or, with some rearranging,

$$Q((s,z),y,w_2) \leqslant \left\{ \frac{M+\delta}{1-\beta} - \delta \right\}^\theta < w_2(s,z). \tag{16}$$

Hence condition (d) of assumption 2.1 holds. In fact, (16) implies that our choice of w_2 also satisfies the uniformly strict inequality in (b) of assumption 2.2.[11]

It only remains to check value-convexity of Q. But this is implied by the convexity of ψ defined in (15), which holds whenever $0 < \theta \leqslant 1$, along with linearity of the integral. The conclusions of theorem 2.3 now follow.

3.2.2. *The Case* $\rho < 1 < \gamma$. To treat this case we again apply the continuous transformation $\hat{v} \equiv v^{1-\gamma}$ to the Bellman equation (11). But now $1-\gamma$ is negative, leading to the *minimization* problem

$$\hat{v}(s,z) = \min_{y \in \Gamma(s,z)} \left\{ r(s,y,z) + \beta \left[\int \hat{v}(y,z') P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^\theta \tag{17}$$

for each $(s,z) \in \mathsf{X}$. The state-action aggregator Q corresponding to (17) is still as defined in (13). Note that in the current setting we have $\theta < 0$.

As (17) is a minimization problem, we aim to apply theorem 2.5. For the bracketing functions w_1 and w_2, we take the constant functions

$$w_1 := \left(\frac{M+\delta}{1-\beta} \right)^\theta \quad \text{and} \quad w_2 := \left(\frac{m}{1-\beta} \right)^\theta ,$$

where δ is a positive constant and m and M are as defined in (14).

The conditions of assumptions 2.1 and 2.3 are all satisfied. Regarding assumption 2.1, the arguments verifying conditions (a) to (c) are identical to those in

[11]To be precise, condition (b) holds when $\varepsilon := [(M+\delta)/(1-\beta)]^\theta - [(M+\delta)/(1-\beta) - \delta]^\theta$.

section 3.2.1. To check condition (d), observe that, for fixed $((s,z),y) \in \mathbb{G}$, we have

$$Q((s,z),y,w_1) = \left\{ r(s,y,z) + \beta \frac{M+\delta}{1-\beta} \right\}^\theta \geqslant \left\{ M + \beta \frac{M+\delta}{1-\beta} \right\}^\theta ,$$

or, with some rearranging,

$$Q((s,z),y,w_1) \geqslant \left\{ \frac{M+\delta}{1-\beta} - \delta \right\}^\theta > w_1(s,z). \tag{18}$$

Similarly, for fixed $((s,z),y) \in \mathbb{G}$, we have

$$Q((s,z),y,w_2) = \left\{ r(s,y,z) + \beta \frac{m}{1-\beta} \right\}^\theta \leqslant \left\{ m + \beta \frac{m}{1-\beta} \right\}^\theta ,$$

and the last term is equal to $w_2(s,z)$. Hence, condition (d) of assumption 2.1 is verified. Furthermore, it is immediate from (18) that our choice of w_1 also satisfies the uniformly strict inequality in (b) of assumption 2.3.

It only remains to check the value-concavity of Q. But this follows directly from the concavity of the function ψ defined in (15), as implied by $\theta < 0$, along with linearity of the integral. We have now checked all conditions of theorem 2.5.

3.2.3. *The Case* $1 < \rho < \gamma$. We now turn to the model in the case where the coefficient of relative risk aversion is still strictly great than 1 but intertemporal elasticity of substitution is less than 1, as is commonly found in the literature.[12] As before, we apply the continuous transformation $\hat{v} \equiv v^{1-\gamma}$ to the Bellman equation (11) and, since $1-\gamma < 0$, the transformed counterpart leads us to the minimization problem as defined in (17). Note that $\theta > 1$ in the current setting.

As (17) is a minimization problem, we aim to apply theorem 2.5. For the bracketing functions w_1 and w_2, we take the constant functions

$$w_1 := \left(\frac{m-\delta}{1-\beta} \right)^\theta \quad \text{and} \quad w_2 := \left(\frac{M}{1-\beta} \right)^\theta ,$$

for some positive $\delta < m$, where m and M are as defined in (14).

[12] See, for example, Hall (1988), Farhi and Werning (2008) and Basu and Bundick (2017).

Assumptions 2.1 and 2.3 are again satisfied. Regarding assumption 2.1, the arguments of verifying conditions (a) to (c) are identical to those in 3.2.1. To check condition (d), observe that, for fixed $((s,z),y) \in \mathbb{G}$, we have

$$Q((s,z),y,w_1) = \left\{ r(s,y,z) + \beta \frac{m-\delta}{1-\beta} \right\}^{\theta} \geqslant \left\{ m + \beta \frac{m-\delta}{1-\beta} \right\}^{\theta},$$

or, with some rearranging

$$Q((s,z),y,w_1) \geqslant \left\{ \frac{m-\delta}{1-\beta} + \delta \right\}^{\theta} > w_1(s,z). \tag{19}$$

Similarly, for fixed $((s,z),y) \in \mathbb{G}$, we have

$$Q((s,z),y,w_2) = \left\{ r(s,y,z) + \beta \frac{M}{1-\beta} \right\}^{\theta} \leqslant \left\{ M + \beta \frac{M}{1-\beta} \right\}^{\theta} = w_2(s,z).$$

Hence condition (d) of assumption 2.1 holds. In fact (19) implies that our choice of w_1 also satisfies the uniformly strict inequality in (b) of assumption 2.3.

Value-concavity of Q is a direct consequence of the concavity of ψ, which holds again when $\theta > 1$, along with linearity of the integral. The conclusions of theorem 2.5 now follow.

3.3. **Risk Sensitive Preferences.** In this section, we consider an economy with a representative agent having the risk sensitive preferences, as in, say, Hansen and Sargent (2008), Gottardi et al. (2015), or Bäuerle and Jaśkiewicz (2018). The generic Bellman equation associated with risk sensitive preferences is

$$v(s,z) = \max_{y \in \Gamma(s,z)} \left\{ r(s,y,z) - \frac{\beta}{\theta} \ln \left[\int \exp\left(-\theta\, v(y,z')\right) P(z, \mathrm{d}z') \right] \right\} \tag{20}$$

for each $(s,z) \in \mathsf{S} \times \mathsf{Z}$. Here, $r \colon \mathbb{G} \to \mathbb{R}$ is a continuous one-period reward function. The parameter $\theta > 0$ captures the risk sensitivity, while other primitives are as discussed in section 3.1. In particular, P is Feller, Γ is continuous and compact valued and both S and Z are compact.

Applying the continuous bijective transformation $\hat{v} \equiv \exp(-\theta v)$ to v in the Bellman equation (20) leads to the minimization problem

$$\hat{v}(s,z) = \min_{y \in \Gamma(s,z)} \exp \left\{ -\theta \left\{ r(s,y,z) - \frac{\beta}{\theta} \ln \left[\int \hat{v}(y,z') P(z, \mathrm{d}z') \right] \right\} \right\}. \tag{21}$$

We translate (21) to our environment by taking $\mathsf{X} := \mathsf{S} \times \mathsf{Z}$ to be the state space, $a = y \in \mathsf{S}$ to be the action, and setting

$$Q((s,z),y,v) = \exp\left\{-\theta\left\{r(s,y,z) - \frac{\beta}{\theta}\ln\left[\int v(y,z')P(z,\mathrm{d}z')\right]\right\}\right\}. \qquad (22)$$

Since r is continuous, there exists a finite constant M with $|r| \leqslant M$. For the bracketing functions, we fix $\delta > 0$ and take the constant functions

$$w_1 := \exp\left[-\theta\left(\frac{M}{1-\beta} + \delta\right)\right] \quad \text{and} \quad w_2 := \exp\left[-\theta\left(\frac{-M}{1-\beta}\right)\right].$$

Assumptions 2.1 and 2.3 are all satisfied. Regarding assumption 2.1, the steps verifying conditions (a) and (b) are identical to those in section 3.2.1. Condition (c) clearly holds, since, for any $b \in \mathbb{R}$, the scalar map

$$\phi(t) := \exp\left[-\theta\left(b - \frac{\beta}{\theta}\ln t\right)\right] \qquad (t > 0) \qquad (23)$$

is monotone increasing. To check condition (d), observe that, for fixed $((s,z),y) \in \mathbb{G}$, we have

$$\begin{aligned} Q((s,z),y,w_1) &= \exp\left(-\theta\left\{r(s,y,z) + \beta\left(\frac{M}{1-\beta} + \delta\right)\right\}\right) \\ &\geqslant \exp\left(-\theta\left\{M + \beta\left(\frac{M}{1-\beta} + \delta\right)\right\}\right) \end{aligned}$$

or, with some rearranging,

$$Q((s,z),y,w_1) \geqslant \exp\left(-\theta\left\{\frac{M}{1-\beta} + \beta\delta\right\}\right) > w_1(s,z). \qquad (24)$$

Similarly, for fixed $((s,z),y) \in \mathbb{G}$, we have

$$\begin{aligned} Q((s,z),y,w_2) &= \exp\left(-\theta\left\{r(s,y,z) + \beta\left(\frac{-M}{1-\beta}\right)\right\}\right) \\ &\leqslant \exp\left(-\theta\left\{-M - \beta\frac{M}{1-\beta}\right\}\right), \end{aligned}$$

and the last term is equal to $w_2(s,z)$. Hence condition (d) of assumption 2.1 holds. In addition, it is obvious from (24) that our choice of w_1 also satisfies the uniformly strict inequality in part (b) of assumption 2.3.[13]

[13] To be precise, condition (b) of assumption 2.3 holds when we set $\varepsilon := \exp\{-\theta[M/(1-\beta) + \beta\delta]\} - \exp\{-\theta[M/(1-\beta) + \delta]\}$.

Finally, condition (a) of assumption 2.3, which is value-concavity of Q, follows from directly the concavity of the function ϕ defined in (23), along with linearity of the integral. The conclusions of theorem 2.5 now follow.

3.4. **Ambiguity.** Extending earlier work by Epstein and Zin (1989) and Klibanoff et al. (2009), Ju and Miao (2012) propose and study a recursive smooth ambiguity model where lifetime value satisfies

$$V_t(C) = \left[(1-\beta)C^{1-\rho} + \beta\{\mathcal{R}_t(V_{t+1}(C))\}^{1-\rho}\right]^{1/(1-\rho)} \tag{25}$$

with

$$\mathcal{R}_t(V_{t+1}(C)) = \left\{\mathbb{E}_{\mu_t}\left(\mathbb{E}_{\pi_{\theta,t}}\left[V_{t+1}^{1-\gamma}(C)\right]\right)^{(1-\eta)/(1-\gamma)}\right\}^{1/(1-\eta)}. \tag{26}$$

As before, ρ is the reciprocal of the EIS and γ governs risk aversion, while η satisfies $0 < \eta \neq 1$ and captures ambiguity aversion. If $\eta = \gamma$, the decision maker is ambiguity neutral and (25)–(26) reduces to the classical recursive utility model of Epstein and Zin (1989). The decision maker displays ambiguity aversion if and only if $\gamma < \eta$. We focus primarily on the case $0 < \rho < 1 < \gamma < \eta$, which is the most empirically relevant.[14]

As a generic formulation of the preferences of Ju and Miao (2012), we consider the Bellman equation

$$v(s,z) = \max_{y \in \Gamma(s,z)} \left\{ r(s,y,z) + \beta \left\{ \int \left[\int v(y,z')^{1-\gamma} \pi_\theta(z, \mathrm{d}z') \right]^{\frac{1-\eta}{1-\gamma}} \mu(z, \mathrm{d}\theta) \right\}^{\frac{1-\rho}{1-\eta}} \right\}^{\frac{1}{1-\rho}} \tag{27}$$

where $(s,z) \in \mathsf{S} \times \mathsf{Z}$. We assume both S and Z to be compact, Γ to be continuous and compact valued, F to be continuous and everywhere positive. The set Θ is a finite parameter space, each element of which is a vector of parameters in the specification of the exogenous state process. Given any $\theta \in \Theta$, the transition function π_θ on Z is assumed to have the Feller property. Given any $z \in \mathsf{Z}$, the distribution $\mu(z, \cdot)$ maps subsets of Θ to $[0,1]$ and evolves as a function of the exogenous state process. We suppose that μ is continuous in z for each $\theta \in \Theta$.

[14]The calibration used in Ju and Miao (2012) is $(\rho, \gamma, \eta) = (0.66, 2.0, 8.86)$. See p. 574.

3.4.1. *The Case $\rho \neq 1$.* Applying the continuous bijective transformation $\hat{v} \equiv v^{1-\eta}$ to v in the Bellman equation (27) leads to the minimization problem

$$\hat{v}(s,z) = \min_{y \in \Gamma(s,z)} \left\{ r(s,y,z) + \beta \left\{ \int \left[\int \hat{v}(y,z')^{\xi_1} \pi_\theta(z, \mathrm{d}z') \right]^{\frac{1}{\xi_1}} \mu(z, \mathrm{d}\theta) \right\}^{\frac{1}{\xi_2}} \right\}^{\xi_2} \tag{28}$$

for all $(s,z) \in \mathsf{S} \times \mathsf{Z}$, where, here and below,

$$\xi_1 := \frac{1-\gamma}{1-\eta} \quad \text{and} \quad \xi_2 := \frac{1-\eta}{1-\rho}.$$

Since this transformation is bijective, there is a one-to-one correspondence between v and $\hat{v}$, in the sense that v solves (27) if and only if $\hat{v}$ solves (28). Note that in the current setting we have $\xi_1 \in (0,1)$ and $\xi_2 < 0$.

We translate this model to our environment by taking $\mathsf{X} := \mathsf{S} \times \mathsf{Z}$ to be the state space, $a = y$ to be the action taking values in S, and setting the state-action aggregator Q to

$$Q((s,z),y,\hat{v}) = \left\{ r(s,y,z) + \beta \left\{ \int \left[\int \hat{v}(y,z')^{\xi_1} \pi_\theta(z, \mathrm{d}z') \right]^{\frac{1}{\xi_1}} \mu(z, \mathrm{d}\theta) \right\}^{\frac{1}{\xi_2}} \right\}^{\xi_2}. \tag{29}$$

As (28) is a minimization problem, we aim to apply theorem 2.5. For the bracketing functions w_1 and w_2, we fix $\delta > 0$ and take the constant functions

$$w_1 := \left(\frac{M+\delta}{1-\beta} \right)^{\xi_2} \quad \text{and} \quad w_2 := \left(\frac{m}{1-\beta} \right)^{\xi_2},$$

where the real numbers m and M are as defined in section 3.2.1.

Assumptions 2.1 and 2.3 are satisfied. Regarding assumption 2.1, condition (a) is true by assumption. Conditions (b) and (c) are proved in lemma 6.12 in the appendix. To verify condition (d), for fixed $((s,z),y) \in \mathbb{G}$, we have

$$\begin{aligned} Q((s,z),y,w_1) &= \left\{ r(s,y,z) + \beta \left\{ \int \left(\frac{M+\delta}{1-\beta} \right)^{\xi_2} \mu(z, \mathrm{d}\theta) \right\}^{1/\xi_2} \right\}^{\xi_2} \\ &= \left\{ r(s,y,z) + \beta \frac{M+\delta}{1-\beta} \right\}^{\xi_2} \geqslant \left\{ M + \beta \frac{M+\delta}{1-\beta} \right\}^{\xi_2}, \end{aligned}$$

where the first equality follows from directly the definition of Q and the fact that $[\int d(z')^{\xi_1} \pi_\theta(z, \mathrm{d}z')]^{1/\xi_1} = d$ for any non-negative constant function d. Furthermore, with some rearranging, we obtain

$$Q((s,z), y, w_1) \geqslant \left\{ \frac{M+\delta}{1-\beta} - \delta \right\}^{\xi_2} > w_1(s,z). \tag{30}$$

Similarly, for fixed $((s,z), y) \in \mathbb{G}$, we have

$$\begin{aligned} Q((s,z), y, w_2) &= \left\{ r(s,y,z) + \beta \left\{ \int \left(\frac{m}{1-\beta} \right)^{\xi_2} \mu(z, \mathrm{d}\theta) \right\}^{1/\xi_2} \right\}^{\xi_2} \\ &= \left\{ r(s,y,z) + \beta \frac{m}{1-\beta} \right\}^{\xi_2} \leqslant \left\{ m + \beta \frac{m}{1-\beta} \right\}^{\xi_2} = w_2(s,z). \end{aligned}$$

Hence condition (d) of assumption 2.1 indeed holds true. Moreover, it is clear from (30) that our choice of w_1 also satisfies the uniformly strict inequality in part (b) of assumption 2.3.

Condition (a) of assumption 2.3 (i.e., value-concavity of Q) is also satisfied, as shown in lemma 6.12 of the appendix. The conclusions of theorem 2.5 now follow.

3.4.2. *The Case* $\rho = 1$. In the limiting case with $\rho = 1$, the generic ambiguity recursion (26) becomes

$$\begin{aligned} U_t(C) =& (1-\beta) \ln C_t \\ &+ \frac{\beta}{1-\eta} \ln \left\{ \mathbb{E}_{\mu_t} \exp \left(\frac{1-\eta}{1-\gamma} \ln \left(\mathbb{E}_{\pi_{\theta,t}} \exp \left((1-\gamma) U_{t+1} \right) \right) \right) \right\}, \end{aligned} \tag{31}$$

where $U_t = \ln V_t$.[15] The generic Bellman equation in (27) becomes

$$\begin{aligned} v(s,z) = \max_{y \in \Gamma(s,z)} & \left\{ r(s,y,z) + \frac{\beta}{1-\eta} \times \right. \\ & \left. \times \ln \left[\int \exp \left(\frac{1}{\xi_1} \ln \left(\int \exp \left((1-\gamma) v(y, z') \right) \pi_\theta(z, \mathrm{d}z') \right) \right) \mu(z, \mathrm{d}\theta) \right] \right\}, \end{aligned} \tag{32}$$

[15] This specification connects with risk-sensitive control and robustness, as studied by Hansen and Sargent (2008). In particular, there are two risk-sensitivity adjustments in (31).

for each $(s, z) \in \mathsf{S} \times \mathsf{Z}$. The one-period return function r is still assumed to be continuous but no longer restricted to be positive, while other primitives are as discussed in section 3.4.1.

Applying the transformation $\hat{v} \equiv \exp[(1-\eta)v]$ to v in the Bellman equation (32) leads us to the minimization problem

$$\begin{aligned} \hat{v}(s,z) = \min_{y \in \Gamma(s,z)} \exp\Bigg((1-\eta)\Bigg\{ r(s,y,z) + \frac{\beta}{1-\eta} \times \\ \times \ln\left[\int \exp\left(\frac{1}{\xi_1} \ln\left(\int \exp\left(\xi_1 \ln \hat{v}(y,z')\right) \pi_\theta(z, \mathrm{d}z') \right) \right) \mu(z, \mathrm{d}\theta) \right] \Bigg\} \Bigg). \end{aligned} \tag{33}$$

With some rearranging, (33) can be written as

$$\begin{aligned} \hat{v}(s,z) = \min_{y \in \Gamma(s,z)} \exp\Bigg((1-\eta)\Bigg\{ r(s,y,z) + \frac{\beta}{1-\eta} \times \\ \times \ln\left[\int \left(\int \hat{v}(y,z')^{\xi_1} \pi_\theta(z, \mathrm{d}z') \right)^{1/\xi_1} \mu(z, \mathrm{d}\theta) \right] \Bigg\} \Bigg). \end{aligned} \tag{34}$$

Note that we still have $\xi_1 \in (0,1)$ and $\eta > 1$ in the current setting with ambiguity aversion. The state-action aggregator Q corresponding to (34) is

$$\begin{aligned} Q((s,z), y, \hat{v}) = \exp\Bigg((1-\eta)\Bigg\{ r(s,y,z) + \frac{\beta}{1-\eta} \times \\ \times \ln\left[\int \left(\int \hat{v}(y,z')^{\xi_1} \pi_\theta(z, \mathrm{d}z') \right)^{1/\xi_1} \mu(z, \mathrm{d}\theta) \right] \Bigg\} \Bigg). \end{aligned} \tag{35}$$

Since the return function r is continuous on a compact set, there exists a finite constant M such that $|r| \leqslant M$. Hence for the bracketing function w_1 and w_2, we fix $\delta > 0$ and take the constant functions

$$w_1 := \exp\left((1-\eta)\left(\frac{M}{1-\beta} + \delta \right) \right) \quad \text{and} \quad w_2 := \exp\left((1-\eta)\left(\frac{-M}{1-\beta} \right) \right).$$

As (34) is the minimization problem, we aim to apply theorem 2.5. Again, assumptions 2.1 and 2.3 are all satisfied.

Regarding assumption 2.1, condition (a) is trivial. Conditions (b) and (c) follow from lemma 6.14 in the appendix. To check condition (d), observe that, for fixed

$((s,z),y) \in \mathbb{G}$, we have

$$\begin{aligned} Q((s,z),y,w_1) &= \exp\left((1-\eta)\left\{r(s,y,z) + \beta\left(\frac{M}{1-\beta} + \delta\right)\right\}\right) \\ &\geqslant \exp\left((1-\eta)\left\{M + \beta\left(\frac{M}{1-\beta} + \delta\right)\right\}\right), \end{aligned}$$

or, with some rearranging,

$$Q((s,z),y,w_1) \geqslant \exp\left((1-\eta)\left\{\frac{M}{1-\beta} + \beta\delta\right\}\right) > w_1(s,z). \tag{36}$$

Similarly, for fixed $((s,z),y) \in \mathbb{G}$, we have

$$\begin{aligned} Q((s,z),y,w_2) &= \exp\left((1-\eta)\left\{r(s,y,z) + \beta\left(\frac{-M}{1-\beta}\right)\right\}\right) \\ &\leqslant \exp\left((1-\eta)\left\{-M - \beta\frac{M}{1-\beta}\right\}\right) = w_2(s,z). \end{aligned}$$

Hence condition (d) of assumption 2.1 holds.

In fact, it is immediate from (36) that our choice of w_1 also satisfies the uniformly strict inequality in part (b) of assumption 2.3. Regarding part (a) of assumption 2.3, value-concavity of Q is immediate from lemma 6.14. We have now checked all conditions of theorem 2.5, and hence the conclusions of that theorem now follow.

4. Unbounded Rewards

This section gives an example of how the methodology proposed above can be extended to the setting of unbounded rewards. The application we consider is the Epstein-Zin problem of section 3.2.3, where $1 < \rho < \gamma$, with Bellman equation

$$v(s,z) = \max_{y \in \Gamma(s,z)} \left\{ r(s,y,z) + \beta \left[\int v(y,z')^{1-\gamma} P(z, \mathrm{d}z') \right]^{\frac{1-\rho}{1-\gamma}} \right\}^{\frac{1}{1-\rho}}$$

for each $(s,z) \in \mathsf{S} \times \mathsf{Z}$. Dropping the compactness assumption, we allow S and Z to be arbitrary separable metric spaces containing possible values for the endogenous and exogenous state variables respectively. As before, Γ is compact valued and continuous, while $r\colon \mathbb{G} \to \mathbb{R}$ is continuous and β lies in $(0,1)$. Let θ be defined by $\theta = (1-\gamma)/(1-\rho)$, so that $\theta > 1$ in the current setting.

In addition, we make the following assumptions.

Assumption 4.1. There exist a continuous function $\kappa\colon \mathsf{S} \times \mathsf{Z} \to [1, \infty)$, positive constants M, L with $L \leqslant M$, and $c \in (0, 1/\beta^\theta)$ and $d \in [0, 1/\beta^\theta)$ satisfying the conditions

$$\sup_{y \in \Gamma(s,z)} r(s,y,z) \leqslant M\kappa(s,z) \quad \text{for all } (s,z) \in \mathsf{S} \times \mathsf{Z}, \tag{37}$$

$$\inf_{y \in \Gamma(s,z)} r(s,y,z) \geqslant L\kappa(s,z) \quad \text{for all } (s,z) \in \mathsf{S} \times \mathsf{Z}, \tag{38}$$

$$\sup_{y \in \Gamma(s,z)} \int \kappa(y,z')^\theta P(z, \mathrm{d}z') \leqslant c\kappa(s,z)^\theta \quad \text{for all } (s,z) \in \mathsf{S} \times \mathsf{Z}, \tag{39}$$

$$\inf_{y \in \Gamma(s,z)} \int \kappa(y,z') P(z, \mathrm{d}z') \geqslant d\kappa(s,z) \quad \text{for all } (s,z) \in \mathsf{S} \times \mathsf{Z}. \tag{40}$$

Moreover, the map $(y,z) \mapsto \int \kappa(y,z')^\theta P(z, \mathrm{d}z')$ is continuous on $\mathsf{S} \times \mathsf{Z}$.

Before progressing further, we make a basic notion that will be used in this section. Suppose ℓ is a real-valued continuous function defined on $\mathsf{S} \times \mathsf{Z}$ with $\ell(s,z) > 0$ for all $(s,z) \in \mathsf{S} \times \mathsf{Z}$. A function $f\colon \mathsf{S} \times \mathsf{Z} \to \mathbb{R}$ is called *ℓ-bounded* if $f(s,z)/\ell(s,z)$ is bounded as (s,z) ranges over $\mathsf{S} \times \mathsf{Z}$.

As in section 3.2.3, we apply the continuous transformation $\hat{v} \equiv v^{1-\gamma}$ to the Bellman equation. Since $1 - \gamma < 0$, the transformed counterpart leads us to a minimization problem with state-action aggregator

$$Q((s,z), y, \hat{v}) = \left\{ r(s,y,z) + \beta \left[\int \hat{v}(y,z') P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^\theta.$$

For the bracketing functions w_1 and w_2, we fix δ such that $0 < \delta < L$, and then adopt the functions

$$w_1 := (L - \delta)^\theta \cdot \kappa \quad \text{and} \quad w_2 := \left(\frac{M}{1 - \beta c^{1/\theta}} \right)^\theta \cdot \kappa^\theta.$$

Note that $\kappa \leqslant \kappa^\theta$, since $\theta > 1$ and $\kappa \geqslant 1$. Hence w_1 and w_2 are both κ^θ-bounded. In addition, the positivity and the continuity of κ directly imply the positivity and continuity of w_1 and w_2. Hence, such w_1 and w_2 are positive κ^θ-bounded continuous functions in $\mathbb{R}^{\mathsf{X}}$ with $w_1 \leqslant w_2$.

Let $\mathscr{V}$ and $\mathscr{C}$ be all Borel measurable functions v in $\mathbb{R}^{\mathsf{X}}$ satisfying $w_1 \leqslant v \leqslant w_2$, and be the continuous functions in $\mathscr{V}$ respectively. For fixed $\sigma \in \Sigma$, a function $\hat{v}_\sigma \in \mathscr{V}$ is called σ-value function if

$$\hat{v}_\sigma(s,z) = \left\{ r(s,\sigma(s,z),z) + \beta \left[\int \hat{v}_\sigma(\sigma(s,z),z') P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^\theta$$

for all $(s,z) \in \mathsf{S} \times \mathsf{Z}$. The following proposition states a result for its existence and uniqueness.

Proposition 4.1. *If assumption 4.1 holds, then, for each $\sigma \in \Sigma$, the set $\mathscr{V}$ contains exactly one σ-value function $\hat{v}_\sigma$.*

From this foundation, the minimum cost function $\hat{v}^*$ associated with this planning problem is defined at $(s,z) \in \mathsf{S} \times \mathsf{Z}$ by

$$\hat{v}^*(s,z) = \inf_{\sigma \in \Sigma} \hat{v}_\sigma(s,z).$$

A function $\hat{v} \in \mathscr{V}$ is said to satisfy the Bellman equation if

$$\hat{v}(s,z) = \min_{y \in \Gamma(s,z)} \left\{ r(s,y,z) + \beta \left[\int \hat{v}(y,z') P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^\theta .$$

In this connection, the corresponding Bellman operator S on $\mathscr{C}$ is defined through

$$S\hat{v}(s,z) = \inf_{y \in \Gamma(s,z)} \left\{ r(s,y,z) + \beta \left[\int \hat{v}(y,z') P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^\theta .$$

Analogous to the result in bounded case (see, e.g., section 3.2.3), we have

Theorem 4.2. *If assumptions 4.1 holds, then*

- *(a) The minimum cost function $\hat{v}^*$ lies in $\mathscr{C}$ and is the unique solution of the Bellman equation (17) in that set.*
- *(b) If $\hat{v}$ is in $\mathscr{C}$, then $S^n \hat{v} \to \hat{v}^*$ as $n \to \infty$.*
- *(c) A policy σ in Σ is optimal if and only if*

$$\sigma(s,z) \in \operatorname*{argmin}_{y \in \Gamma(s,z)} Q((s,z), y, \hat{v}^*) \text{ for all } (s,z) \in \mathsf{S} \times \mathsf{Z}.$$

(d) *At least one optimal policy exists.*

5. CONCLUSION

Recursive preference models have allowed economists to successfully replicate important empirical phenomena in a range of different settings. To date, most attempts to provide a theory of dynamic programming for recursive preferences models commensurate with that available for traditional additivity separable preferences has focused on exploiting concavity available in some classes of preferences. Here we instead used convexity, which pairs well with maximization, allowing us to provide conditions for optimality that are relatively simple, general enough to include many classes of preferences and strong in their conclusions.

Our work leaves several open questions. For example, while an application with unbounded rewards was provided in section 4, it would be more valuable still to obtain a systematic treatment of the unbounded case. This topic is left for future research.

6. APPENDIX

Let $m\mathsf{X}$ represent all Borel measurable functions in $\mathbb{R}^{\mathsf{X}}$ and let $c\mathsf{X}$ denote all continuous functions in $m\mathsf{X}$. Let $bm\mathsf{X}$ be the bounded functions in $m\mathsf{X}$ and let $bc\mathsf{X}$ be the continuous functions in $bm\mathsf{X}$. Let $m\mathsf{X}_+$ and $m\mathsf{X}_{++}$ be the non-negative and positive functions in $m\mathsf{X}$, respectively. Recall that a self map A on a convex subset M of $bm\mathsf{X}$ is called

- *asymptotically stable* on M if A has a unique fixed point v^* in M and $A^n v \to v^*$ as $n \to \infty$ whenever $v \in M$,
- *isotone* if $Av \leqslant Av'$ whenever $v, v' \in M$ with $v \leqslant v'$,
- *convex* if $A(\lambda v + (1-\lambda)v') \leqslant \lambda Av + (1-\lambda)Av'$ whenever $v, v' \in M$ and $0 \leqslant \lambda \leqslant 1$, and
- *concave* if $A(\lambda v + (1-\lambda)v') \geqslant \lambda Av + (1-\lambda)Av'$ whenever $v, v' \in M$ and $0 \leqslant \lambda \leqslant 1$.

For $f, g \in bm\mathsf{X}$, the statement $f \ll g$ means that there exists an $\varepsilon > 0$ such that $f(x) \leqslant g(x) - \varepsilon$ for all $x \in \mathsf{X}$.

For each $\sigma \in \Sigma$, we define the *σ-value operator* T_σ on $\mathscr{V}$ by

$$T_\sigma v(x) := Q(w, \sigma(x), v) \quad \text{for all } x \in \mathsf{X},\ v \in \mathscr{V}. \tag{41}$$

Stating that $v_\sigma \in \mathscr{V}$ solves (3) is equivalent to stating that v_σ is a fixed point of T_σ. By lemma 2.1, the operator T_σ is a well defined self-map on $\mathscr{V}$.

6.1. **Proofs for the Convex Case.**

Lemma 6.1. *If assumption 2.2 holds, then, for each $\sigma \in \Sigma$, the operator T_σ is asymptotically stable on $\mathscr{V}$.*

Proof of lemma 6.1. Fix $\sigma \in \Sigma$. We aim to apply theorem 3.1 of Du (1990). To this end, it is sufficient to show that

(i) T_σ is isotone and convex on $\mathscr{V}$.
(ii) $T_\sigma w_1 \geqslant w_1$ and $T_\sigma w_2 \ll w_2$.

Regarding condition (i), pick any $v, v' \in \mathscr{V}$ with $v \leqslant v'$. For fixed $x \in \mathsf{X}$, we have

$$T_\sigma v(x) = Q(x, \sigma(x), v) \leqslant Q(x, \sigma(x), v') = T_\sigma v'(x),$$

by (1). Hence, isotonicity of T_σ holds.

To see that T_σ is convex, fix $v, v' \in \mathscr{V}$ and $\lambda \in [0, 1]$. For any given $x \in \mathsf{X}$, we have

$$\begin{aligned} T_\sigma(\lambda v + (1-\lambda)v')(x) &= Q(x, \sigma(x), \lambda v + (1-\lambda)v') \\ &\leqslant \lambda Q(x, \sigma(x), v) + (1-\lambda) Q(x, \sigma(x), v') \\ &= \lambda T_\sigma v(x) + (1-\lambda) T_\sigma v'(x), \end{aligned}$$

where the inequality directly follows from part (a) of assumption 2.2. Since $x \in \mathsf{X}$ was arbitrary, the convexity of T_σ follows.

The first part of condition (ii) follows directly from (2), since, for each $x \in \mathsf{X}$,

$$T_\sigma w_1(x) = Q(x, \sigma(x), w_1) \geqslant w_1(x).$$

To see that the second part of condition (ii) is satisfied, it follows from part (b) of assumption 2.2 that

$$T_\sigma w_2(x) = Q(x, \sigma(x), w_2) \leqslant w_2(x) - \varepsilon$$

for each $x \in \mathsf{X}$ and for some $\varepsilon > 0$. Hence $w_2 \gg T_\sigma w_2$, as was to be shown. □

Proof of proposition 2.2. This follows directly from lemma 6.1. □

Given $v \in \mathscr{V}$, a policy σ in Σ will be called *v-maximal-greedy* if

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} Q(x, a, v) \quad \text{for all } x \in \mathsf{X}. \tag{42}$$

Lemma 6.2. *If $v \in \mathscr{C}$, then there exists at least one v-maximal-greedy policy.*

Proof. Fixing $v \in \mathscr{C}$ and using the compactness and continuity conditions in assumption 2.1, we can choose for each $x \in \mathsf{X}$ an action $\sigma(x) \in \Gamma(x)$ such that (42) holds. The map σ constructed in this manner can be chosen to be Borel measurable by theorem 18.19 of Aliprantis and Border (2006). □

Lemma 6.3. *If assumption 2.2 holds, then T is asymptotically stable on $\mathscr{C}$.*

Proof of lemma 6.3. In order to apply theorem 3.1 of Du (1990), it suffices to show that

(i) T is isotone and convex on $\mathscr{C}$.
(ii) $Tw_1 \geqslant w_1$ and $Tw_2 \ll w_2$.

The isotonicity of T on $\mathscr{C}$ is obvious, since, by the monotonicity restriction (1),

$$v \leqslant v' \implies \max_{a \in \Gamma(x)} Q(x, a, v) \leqslant \max_{a \in \Gamma(x)} Q(x, a, v) \quad \text{for all } x \in \mathsf{X}.$$

In other words, by definition of T, $v \leqslant v'$ implies $Tv \leqslant Tv'$.

To show the convexity of T, fix $v, v' \in \mathscr{C}$ and $\lambda \in [0, 1]$. For any given $(x, a) \in \mathbb{G}$, we have, by part (a) of assumption 2.2,

$$\begin{aligned} Q(x, a, \lambda v + (1-\lambda)v') &\leqslant \lambda Q(x, a, v) + (1-\lambda) Q(x, a, v') \\ &\leqslant \lambda \max_{a \in \Gamma(x)} Q(x, a, v) + (1-\lambda) \max_{a \in \Gamma(x)} Q(x, a, v') \\ &= \lambda T v(x) + (1-\lambda) T v'(x). \end{aligned}$$

Since $(x, a) \in \mathbb{G}$ was arbitrary, the above inequality implies

$$\max_{a \in \Gamma(x)} Q(x, a, \lambda v + (1 - \lambda) v') \leqslant \lambda T v(x) + (1 - \lambda) T v'(x)$$

for each $x \in \mathsf{X}$, which in turn means that $T[\lambda v + (1 - \lambda) v'] \leqslant \lambda T v + (1 - \lambda) T v'$.

The first part of condition (ii) follows directly from (2), since, for each $x \in \mathsf{X}$,

$$T w_1(x) = \max_{a \in \Gamma(x)} Q(x, a, w_1) \geqslant Q(x, a, w_1) \geqslant w_1(x).$$

To see that the second part of condition (ii) is satisfied, it follows from part (b) of assumption 2.2 that

$$T w_2(x) = \max_{a \in \Gamma(x)} Q(x, a, w_2) \leqslant w_2(x) - \varepsilon$$

for each $x \in \mathsf{X}$ and for some $\varepsilon > 0$. Hence, $T w_2 \ll w_2$, as was to be shown. □

Theorem 6.4. *If T_σ is asymptotically stable on $\mathscr{V}$ for all $\sigma \in \Sigma$ and T is asymptotically stable on $\mathscr{C}$, then the conclusions of theorem 2.3 hold.*

Proof. Let v^* be the maximum value function and let $\bar{v}$ be the unique fixed point of T in $\mathscr{C}$. To see that $\bar{v} = v^*$, first observe that $\bar{v} \in \mathscr{C}$ and hence $\bar{v}$ has at least one maximal-greedy policy σ. For this policy we have, by definition, $T_\sigma \bar{v}(x) = T\bar{v}(x)$ at each x, from which it follows that $\bar{v} = T\bar{v} = T_\sigma \bar{v}$. Since T_σ is asymptotically stable on $\mathscr{V}$, we know that its unique fixed point is v_σ, so $\bar{v} = v_\sigma$. But then $\bar{v} \leqslant v^*$, by the definition of v^*.

To see that the reverse inequality holds, pick any $\sigma \in \Sigma$. We have $T_\sigma \bar{v} \leqslant T\bar{v} = \bar{v}$. Iterating on this inequality and using the isotonicity of T_σ gives $T_\sigma^k \bar{v} \leqslant \bar{v}$ for all k. Taking the limit with respect to k and using the asymptotic stability of T_σ then gives $v_\sigma \leqslant \bar{v}$. Hence $v^* \leqslant \bar{v}$, and we can now conclude that $\bar{v} = v^*$.

Since $\bar{v} \in \mathscr{C}$, we have $v^* \in \mathscr{C}$. It follows that v^* is the unique solution to the Bellman maximization equation in $\mathscr{C}$, and that $T^n v \to v^*$ whenever $v \in \mathscr{C}$. Parts (a) and (b) of theorem 2.3 are now established.

Regarding part (c) and (d), by the definition of maximal-greedy policies, we know that σ is v^*-maximal-greedy iff $Q(x, \sigma(x), v^*) = \max_{a \in \Gamma(x)} Q(x, a, v^*)$ for all $x \in \mathsf{X}$. Since v^* satisfies the Bellman maximization equation, we then have

$$\sigma \text{ is } v^*\text{-maximal-greedy} \quad \iff \quad Q(x, \sigma(x), v^*) = v^*(x), \;\; \forall\, x \in \mathsf{X}.$$

But, by proposition 2.2, the right hand side is equivalent to the statement that $v^* = v_\sigma$. Hence, by this chain of logic and the definition of optimality,

$$\sigma \text{ is } v^*\text{-maximal-greedy} \iff v^* = v_\sigma \iff \sigma \text{ is optimal} \tag{43}$$

Moreover, the fact that v^* is in $\mathscr{C}$ combined with lemma 6.2 assures us that at least one v^*-maximal-greedy policy exists. Each such policy is optimal, so the set of optimal policies is nonempty. □

6.2. **Proofs for the Concave Case.**

Lemma 6.5. *If assumption 2.3 holds, then, for each $\sigma \in \Sigma$, the operator T_σ is asymptotically stable on $\mathscr{V}$.*

Proof of lemma 6.5. Fix $\sigma \in \Sigma$. We aim to apply theorem 3.1 of Du (1990). To this end, it is sufficient to show that

(i) T_σ is isotone and concave on $\mathscr{V}$, and
(ii) $T_\sigma w_1 \gg w_1$ and $T_\sigma w_2 \leqslant w_2$.

Clearly, T_σ is isotone, since, by the monotonicity restriction (1),

$$v \leqslant v' \implies Q(x, \sigma(x), v) \leqslant Q(x, \sigma(x), v') \quad \text{for all } x \in \mathsf{X}.$$

In other words, $v \leqslant v'$ implies $T_\sigma v \leqslant T_\sigma v'$.

Regarding the concavity of T_σ, fix $v, v' \in \mathscr{V}$ and $\lambda \in [0, 1]$. For any given $x \in \mathsf{X}$, by virtue of part (a) of assumption 2.3, we obtain

$$\begin{aligned} T_\sigma(\lambda v + (1-\lambda)v')(x) &= Q(x, \sigma(x), \lambda v + (1-\lambda)v') \\ &\geqslant \lambda Q(x, \sigma(x), v) + (1-\lambda) Q(x, \sigma(x), v') \\ &= \lambda T_\sigma v(x) + (1-\lambda) T_\sigma v'(x). \end{aligned}$$

Since $x \in \mathsf{X}$ was arbitrary, the concavity of T_σ follows.

To see that the first part of condition (ii) is satisfied, it follows from part (b) of assumption 2.3 that

$$T_\sigma w_1(x) = Q(x, \sigma(x), w_1) \geqslant w_1(x) + \varepsilon$$

for each $x \in \mathsf{X}$ and for some $\varepsilon > 0$. Hence, $T_\sigma w_1 \gg w_1$, as was to be shown.

The second part of condition (ii) follows directly from (2), since, for each $x \in \mathsf{X}$,

$$T_\sigma w_2(x) = Q(x, \sigma(x), w_2) \leqslant w_2(x).$$

This completes the proof. □

Proof of proposition 2.4. This follows directly from lemma 6.5. □

Given $v \in \mathscr{V}$, a policy σ in Σ will be called *v-minimal-greedy* if

$$\sigma(x) \in \operatorname*{argmin}_{a \in \Gamma(x)} Q(x, a, v) \quad \text{for all } x \in \mathsf{X}. \tag{44}$$

Lemma 6.6. *If $v \in \mathscr{C}$, then there exists at least one v-minimal-greedy policy.*

Proof. The proof of lemma 6.6 is essentially identical to that of lemma 6.2, and hence is omitted here. □

Lemma 6.7. *If assumption 2.3 holds, then S is asymptotically stable on $\mathscr{C}$.*

Proof of lemma 6.7. It follows from Berge's theorem of the minimum that, when v is in $\mathscr{C}$, we have

$$Sv(x) = \min_{a \in \Gamma(x)} Q(x, a, v)$$

and Sv is an element of $\mathscr{C}$.

In order to apply theorem 3.1 of Du (1990), it suffices to show that

(i) S is isotone and concave on $\mathscr{C}$, and
(ii) $Sw_1 \gg w_1$ and $Sw_2 \leqslant w_2$.

The isotonicity of S on $\mathscr{C}$ is obvious, since, by the monotonicity restriction (1),

$$v \leqslant v' \implies \min_{a \in \Gamma(x)} Q(x, a, v) \leqslant \min_{a \in \Gamma(x)} Q(x, a, v) \quad \text{for all } x \in \mathsf{X}.$$

In other words, by definition of S, $v \leqslant v'$ implies $Sv \leqslant Sv'$.

To show the concavity of S, fix $v, v' \in \mathscr{C}$ and $\lambda \in [0,1]$. For any given $(x,a) \in \mathbb{G}$, by part (a) of assumption 2.3, we have

$$\begin{aligned} Q(x,a,\lambda v + (1-\lambda)v') &\geqslant \lambda Q(x,a,v) + (1-\lambda)Q(x,a,v') \\ &\geqslant \lambda \min_{a \in \Gamma(x)} Q(x,a,v) + (1-\lambda) \min_{a \in \Gamma(x)} Q(x,a,v') \\ &= \lambda S v(x) + (1-\lambda) S v'(x). \end{aligned}$$

Since $(x,a) \in \mathbb{G}$ was arbitrary, the above inequality implies

$$\min_{a \in \Gamma(x)} Q(x,a,\lambda v + (1-\lambda)v') \geqslant \lambda S v(x) + (1-\lambda) S v'(x)$$

for each $x \in \mathsf{X}$; namely, $S[\lambda v + (1-\lambda)v'] \geqslant \lambda S v + (1-\lambda) S v'$, as desired.

To see that the first part of condition (ii) is satisfied, it follows from part (b) of assumption 2.3 that

$$S w_1(x) = \min_{a \in \Gamma(x)} Q(x,a,w_1) \geqslant w_1(x) + \varepsilon$$

for each $x \in \mathsf{X}$ and some $\varepsilon > 0$. Hence, $S w_1 \gg w_1$.

The second part of condition (ii) directly follows from (2), since, for each $x \in \mathsf{X}$,

$$S w_2(x) = \min_{a \in \Gamma(x)} Q(x,a,w_2) \leqslant Q(x,a,w_2) \leqslant w_2(x).$$

This finishes the proof. $\square$

Theorem 6.8. *If T_σ is asymptotically stable on $\mathscr{V}$ for all $\sigma \in \Sigma$ and S is asymptotically stable on $\mathscr{C}$, then the conclusions of theorem 2.5 hold.*

Proof. Let v^* be the minimum cost function and let $\bar{v}$ be the unique fixed point of S in $\mathscr{C}$. To see that $\bar{v} = v^*$, first observe that $\bar{v} \in \mathscr{C}$ and hence $\bar{v}$ has at least one minimal-greedy policy σ. For this policy we have, by definition, $T_\sigma \bar{v}(x) = S\bar{v}(x)$ at each x, from which it follows that $\bar{v} = S\bar{v} = T_\sigma \bar{v}$. Since T_σ is asymptotically stable on $\mathscr{V}$, we know that its unique fixed point is v_σ, so $\bar{v} = v_\sigma$. But then $\bar{v} \geqslant v^*$, by the definition of v^* in (7).

To see that the reverse inequality holds, pick any $\sigma \in \Sigma$. We have $T_\sigma \bar{v} \geqslant S\bar{v} = \bar{v}$. Iterating on this inequality and using the isotonicity of T_σ gives $T_\sigma^k \bar{v} \geqslant \bar{v}$ for all k. Taking the limit with respect to k and using the asymptotic stability of T_σ then gives $v_\sigma \geqslant \bar{v}$. Hence $v^* \geqslant \bar{v}$, and we can now conclude that $\bar{v} = v^*$.

Since $\bar{v} \in \mathscr{C}$, we have $v^* \in \mathscr{C}$. Moreover, for v in $\mathscr{C}$ we can replace the inf in the definition of S with a min, and solutions to the Bellman equation in $\mathscr{C}$ exactly coincide with fixed points of S in that set. It follows that v^* is the unique solution to the Bellman equation in $\mathscr{C}$, and that $S^n v \to v^*$ whenever $v \in \mathscr{C}$. Parts (a) and (b) of theorem 2.5 are now established.

Regarding part (c) and (d), by the definition of minimal-greedy policies, we know that σ is v^*-minimal-greedy iff $Q(x, \sigma(x), v^*) = \min_{a \in \Gamma(x)} Q(x, a, v^*)$ for all $x \in \mathsf{X}$. Since v^* satisfies the Bellman equation, we then have

$$\sigma \text{ is } v^*\text{-minimal-greedy} \quad \iff \quad Q(x, \sigma(x), v^*) = v^*(x), \;\; \forall\, x \in \mathsf{X}.$$

But, by proposition 2.4, the right hand side is equivalent to the statement that $v^* = v_\sigma$. Hence, by this chain of logic and the definition of optimality,

$$\sigma \text{ is } v^*\text{-minimal-greedy} \iff v^* = v_\sigma \iff \sigma \text{ is optimal} \tag{45}$$

Moreover, the fact that v^* is in $\mathscr{C}$ combined with lemma 6.6 assures us that at least one v^*-minimal-greedy policy exists. Each such policy is optimal, so the set of optimal policies is nonempty. □

6.3. **Proofs for Section 3.4.** In this section, we prove some properties of the state-action aggregator Q defined in section 3.4.

For the sake of exposition, fix $\theta \in \Theta$, we first define an operator R_θ on $bm(\mathsf{S} \times \mathsf{Z})_+$ by

$$(R_\theta w)(y, z) := \left[\int w(y, z')^{\xi_1} \pi_\theta(z, \mathrm{d}z') \right]^{1/\xi_1} \quad \text{for all } (y, z) \in \mathsf{S} \times \mathsf{Z}. \tag{46}$$

From this foundation, we then define an operator R that is a map sending w in $bm(\mathsf{S} \times \mathsf{Z})_+$ into

$$Rw(y, z, \theta) := (R_\theta w)(y, z) \quad \text{for all } (y, z, \theta) \in \mathsf{S} \times \mathsf{Z} \times \Theta. \tag{47}$$

The following lemma shows some useful properties of the operator R_θ.

Lemma 6.9. *For fixed $\theta \in \Theta$, if ξ_1 lies in $(0, 1)$, then the operator R_θ defined in (46) is isotone and concave on $bm(\mathsf{S} \times \mathsf{Z})_+$.*

Moreover, the function $R_\theta w$ is nonnegative, bounded, and Borel measurable on $\mathsf{S} \times \mathsf{Z}$ whenever $w \in bm(\mathsf{S} \times \mathsf{Z})_+$ and continuous on $\mathsf{S} \times \mathsf{Z}$ whenever $w \in bc(\mathsf{S} \times \mathsf{Z})_+$.

Proof of lemma 6.9. Fix $\theta \in \Theta$. The isotonicity of R_θ is obvious, since the scalar function $\mathbb{R}_+ \ni t \mapsto t^{\xi_1} \in \mathbb{R}_+$ and its inverse are both strictly increasing on $\mathbb{R}_+$.

Since $\xi_1 \in (0,1)$, by virtue of Theorem 198 of Hardy et al. (1934), we know that R_θ is super-additive in the sense that for any $w, w' \in m(\mathsf{S} \times \mathsf{Z})_+$, $R_\theta(w + w') \geqslant R_\theta(w) + R_\theta(w')$.[16] As a result, the super-additivity and the positive homogeneity of R_θ together yield the concavity of R_θ.[17] Indeed, pick any $\lambda \in [0,1]$ and $w, w' \in m(\mathsf{S} \times \mathsf{Z})_+$, by the convexity of $m(\mathsf{S} \times \mathsf{Z})_+$, we have

$$\begin{aligned} R_\theta[\lambda w + (1-\lambda)w'] &\geqslant R_\theta(\lambda w) + R_\theta((1-\lambda)w') \quad \text{(by super-additivity)} \\ &= \lambda R_\theta(w) + (1-\lambda) R_\theta(w') \quad \text{(by positive homogeneity)}, \end{aligned}$$

as was to be shown.

Regarding the second claim of lemma 6.9, non-negativity and boundedness of $R_\theta w$ is immediate and it is easy to see that $R_\theta w$ is Borel measurable on $\mathsf{S} \times \mathsf{Z}$ whenever $w \in bm(\mathsf{S} \times \mathsf{Z})_+$. Now fix $w \in bc(\mathsf{S} \times \mathsf{Z})_+$. We note that the function w^{ξ_1} also lies in $bc(\mathsf{S} \times \mathsf{Z})_+$. Then, by virtue of Feller property of π_θ, the mapping $\mathsf{S} \times \mathsf{Z} \ni (y,z) \mapsto \int w(y,z')^{\xi_1} \pi_\theta(z, \mathrm{d}z') \in \mathbb{R}_+$ is bounded and continuous on $\mathsf{S} \times \mathsf{Z}$. Furthermore, it follows that the mapping $\mathsf{S} \times \mathsf{Z} \ni (y,z) \mapsto [\int w(y,z')^{\xi_1} \pi_\theta(z, \mathrm{d}z')]^{1/\xi_1} \in \mathbb{R}_+$ is continuous on $\mathsf{S} \times \mathsf{Z}$, since the inverse of the map $t \mapsto t^{\xi_1}$ is also continuous on $\mathbb{R}_+$. Therefore, our claim follows. $\square$

As an application of lemma 6.9, we now present the next result.

Lemma 6.10. *The operator R defined in (47) is a well-defined map from $bm(\mathsf{S} \times \mathsf{Z})_+$ into $bm(\mathsf{S} \times \mathsf{Z} \times \Theta)_+$.*

Proof of lemma 6.10. Fix w in $bm(\mathsf{S} \times \mathsf{Z})_+$. Since boundedness and non-negativity of the function Rw are obvious, it remains to show that Rw is measurable on $\mathsf{S} \times \mathsf{Z} \times \Theta$.

On one hand, for each $\theta \in \Theta$, it follows from lemma 6.9 that the function $Rw(\cdot,\cdot,\theta) = R_\theta w \colon \mathsf{S} \times \mathsf{Z} \to \mathbb{R}_+$ is Borel measurable. On the other hand, for

[16] This result can also be reviewed as the reverse Minkowski inequality, see, for example, Proposition 3.2 in page 225 of DiBenedetto (2002).

[17] An operator A defined on the positive cone $bm\mathsf{X}_+$ of $bm\mathsf{X}$ is called *positively homogeneous* (of the first degree) if for any v in $bm\mathsf{X}_+$ and any real number $t \geqslant 0$, we have $A(tv) = tAv$.

each $(y,z) \in \mathsf{S} \times \mathsf{Z}$, the function $Rw(y,z,\cdot)\colon \Theta \to \mathbb{R}_+$ is continuous, since Θ is a finite set (endowed with the discrete topology).

In this connection, we conclude that the function $Rw\colon \mathsf{S} \times \mathsf{Z} \times \Theta \to \mathbb{R}$ is a Carathéodory function, in the sense that

(1) for each $\theta \in \Theta$, the function $Rw(\cdot,\cdot,\theta)\colon \mathsf{S} \times \mathsf{Z} \to \mathbb{R}$ is Borel measurable; and
(2) for each $(y,z) \in \mathsf{S} \times \mathsf{Z}$, the function $Rw(y,z,\cdot)\colon \Theta \to \mathbb{R}$ is continuous.

By virtue of lemma 4.51 in Aliprantis and Border (2006), it follows that the Carathéodory function Rw is jointly measurable on $\mathsf{S} \times \mathsf{Z} \times \Theta$, as desired. $\square$

In this connection, the state-action aggregator Q defined in (29) can be simply expressed as a composition of two operators R and $\tilde{Q}$ as follows

$$Q((s,z),y,\hat{v}) := \tilde{Q}((s,z),y,R\hat{v}), \tag{48}$$

with

$$\tilde{Q}((s,z),y,h) := \left\{ r(s,y,z) + \beta \left\{ \int h(y,z,\theta)\mu(z,\mathrm{d}\theta) \right\}^{1/\xi_2} \right\}^{\xi_2} \tag{49}$$

for all $((s,z),y) \in \mathbb{G}$ and $h \in bm(\mathsf{S} \times \mathsf{Z} \times \Theta)_{++}$.

It is worth noting that the formula of $\tilde{Q}$ defined in (49) is almost identical to that of Q defined in (13). Hence, recalling the results associated with Q in section 3.2.2, we have

Lemma 6.11. *If $\xi_2 < 0$, then $\tilde{Q}$ defined in (49) is isotone and concave in its third argument on $bm(\mathsf{S} \times \mathsf{Z} \times \Theta)_{++}$.*

In addition, the map $((s,z),y) \mapsto \tilde{Q}((s,z),y,h)$ is Borel measurable on $\mathbb{G}$ whenever $h \in bm(\mathsf{S} \times \mathsf{Z} \times \Theta)_{++}$, and continuous on $\mathbb{G}$ whenever the map $h(\cdot,\cdot,\theta)\colon \mathsf{S} \times \mathsf{Z} \to \mathbb{R}_{++}$ is continuous, for each $\theta \in \Theta$.

Proof of lemma 6.11. Analogous to the proofs in section 3.2.2, for any fixed $b > 0$, we consider the scalar map

$$\psi(t) := (b + \beta t^{1/\xi_2})^{\xi_2} \qquad (t > 0).$$

Since $\xi_2 < 0$, it is clear that the scalar function ψ is continuous, strictly increasing and strictly concave on $\mathbb{R}_{++}$.[18]

The first part of claim is immediate from the monotonicity and concavity of ψ, along with monotonicity and linearity of the integral.

Regarding the remaining part, fix a function h in $bm(\mathsf{S} \times \mathsf{Z} \times \Theta)_{++}$. Borel measurability of the map $((s,z),y) \mapsto \tilde{Q}((s,z),y,h)$ is obvious. Now fix a function h satisfying that the map $h(\cdot,\cdot,\theta)\colon \mathsf{S} \times \mathsf{Z} \to \mathbb{R}_{++}$ is continuous, for every $\theta \in \Theta$. By virtue of the continuity property imposed on the distribution $\mu(\cdot,\cdot)$, the map $\mathsf{S} \times \mathsf{Z} \ni (y,z) \mapsto \int h(y,z,\theta)\mu(z,\mathrm{d}\theta) \in \mathbb{R}_{++}$ is continuous on $\mathsf{S} \times \mathsf{Z}$.[19] It then follows from the continuity imposed on r and the continuity of ψ that the map $((s,z),y) \mapsto \tilde{Q}((s,z),y,h)$ is continuous on $\mathbb{G}$. $\square$

Lemma 6.12. *If $\xi_1 \in (0,1)$ and $\xi_2 < 0$, then the state-action aggregator Q defined in (29) is isotone and concave in its third argument on $bm(\mathsf{S} \times \mathsf{Z})_{++}$.*

In addition, the map $((s,z),y) \mapsto Q((s,z),y,v)$ is Borel measurable on $\mathbb{G}$ whenever $v \in bm(\mathsf{S} \times \mathsf{Z})_{++}$ and continuous on $\mathbb{G}$ whenever $v \in bc(\mathsf{S} \times \mathsf{Z})_{++}$.

Proof of lemma 6.12. Since the aggregator Q is a composition of $\tilde{Q}$ and R, by lemmas 6.9 to 6.11, the isotonicity, Borel measurability and continuity of Q immediately follow from those of $\tilde{Q}$ and R.

It only remains to show the concavity of Q. To see this, fix $((s,z),y) \in \mathbb{G}$, $\lambda \in [0,1]$ and w, w' in $bm(\mathsf{S} \times \mathsf{Z})_{++}$. For any given $\theta \in \Theta$, by concavity of R_θ and convexity of $bm(\mathsf{S} \times \mathsf{Z})_{++}$, we have

$$R_\theta[\lambda w + (1-\lambda)w'](y,z) \geqslant \lambda R_\theta w(y,z) + (1-\lambda)R_\theta w'(y,z);$$

that is, for each $(y,z,\theta) \in \mathsf{S} \times \mathsf{Z} \times \Theta$,

$$R[\lambda w + (1-\lambda)w'](y,z,\theta) \geqslant \lambda R w(y,z,\theta) + (1-\lambda)R w'(y,z,\theta).$$

In operator notation, this translates to $R[\lambda w + (1-\lambda)w'] \geqslant \lambda R w + (1-\lambda) R w'$.

[18] Please refer to the relevant results of such ψ in section 3.2.2.

[19] Since Θ is finite, this map becomes to a sum of a finite number of functions that are continuous in (y,z), and thus it is continuous in (y,z) as well.

Observe that due to isotonicity and concavity of $\tilde{Q}$, we now obtain

$$\begin{aligned}
Q((s,z),y,\lambda w + (1-\lambda)w') &= \tilde{Q}((s,z),y,R[\lambda w + (1-\lambda)w']) \\
&\geqslant \tilde{Q}((s,z),y,\lambda Rw + (1-\lambda)Rw') \\
&\geqslant \lambda \tilde{Q}((s,z),y,Rw) + (1-\lambda)\tilde{Q}((s,z),y,Rw') \\
&= \lambda Q((s,z),y,w) + (1-\lambda)Q((s,z),y,w'),
\end{aligned}$$

where the first and last equalities follow immediately from the definition of Q in (48), while the first and second inequalities follow from isotonicity and concavity of $\tilde{Q}$, respectively. This completes the proof. □

Analogously, the state-action aggregator Q defined in (35) can be expressed as

$$Q((s,z),y,\hat{v}) = \tilde{Q}((s,z),y,R\hat{v}),$$

with the operator R defined as above, but

$$\tilde{Q}((s,z),y,h) := \exp\left((1-\eta)\left\{r(s,y,z) + \frac{\beta}{1-\eta}\ln\left[\int h(y,z,\theta)\mu(z,\mathrm{d}\theta)\right]\right\}\right) \tag{50}$$

for all $((s,z),y) \in \mathbb{G}$ and $h \in bm(\mathsf{S} \times \mathsf{Z} \times \Theta)_{++}$.

Observe that the formula of $\tilde{Q}$ defined above is almost identical to that of Q defined in (22). In this connection, recalling the results associated with Q in section 3.3, it is easy to see that

Lemma 6.13. *If $\eta > 1$, then $\tilde{Q}$ defined in (50) is isotone and concave in its third argument on $bm(\mathsf{S} \times \mathsf{Z} \times \Theta)_{++}$.*

In addition, the map $((s,z),y) \mapsto \tilde{Q}((s,z),y,h)$ is Borel measurable on $\mathbb{G}$ whenever $h \in bm(\mathsf{S} \times \mathsf{Z} \times \Theta)_{++}$, and continuous on $\mathbb{G}$ whenever the map $h(\cdot,\cdot,\theta)\colon \mathsf{S} \times \mathsf{Z} \to \mathbb{R}_{++}$ is continuous, for each $\theta \in \Theta$.

Proof of lemma 6.13. Analogous to the proof of lemma 6.11, for fixed $b \in \mathbb{R}$, we consider the scalar map

$$\psi(t) := \exp\left[(1-\eta)\left(b + \frac{\beta}{1-\eta}\ln t\right)\right] \qquad (t > 0).$$

It is clear that this scalar function ψ is continuous, strictly increasing and strictly concave on $\mathbb{R}_{++}$.[20] As a consequence, the remaining proof of lemma 6.13 is identical to that of lemma 6.11, and thus omitted here. □

Lemma 6.14. *If $\xi_1 \in (0,1)$ and $\eta > 1$, then the state-action aggregator Q defined in (35) is isotone and concave in its third argument on $bm(\mathsf{S} \times \mathsf{Z})_{++}$.*

In addition, the map $((s,z),y) \mapsto Q((s,z),y,v)$ is Borel measurable on $\mathbb{G}$ whenever $v \in bm(\mathsf{S} \times \mathsf{Z})_{++}$ and continuous on $\mathbb{G}$ whenever $v \in bc(\mathsf{S} \times \mathsf{Z})_{++}$.

Proof of lemma 6.14. Invoking lemmas 6.9, 6.10 and 6.13, the proof is identical to that of lemma 6.12, and hence is omitted. □

6.4. **Proofs in section 4.** Recalling the definition of the weight function ℓ in section 4, the finite ℓ-norm turns the real normed vector space $b_\ell m\mathsf{X} := \{f \in m\mathsf{X} : f \text{ is } \ell\text{-bounded}\}$ into a real Banach space.[21]

Recalling the definition of Q and the construction of bracketing functions w_1 and w_2 in section 4, we obtain the following results.

Lemma 6.15. *If assumption 4.1 holds, the state-action aggregator Q is isotone and concave in its third argument on $\mathscr{V}$.*

Proof of lemma 6.15. The isotonicity and concavity of such aggregator Q have been already proved in section 3.2.3. □

Lemma 6.16. *If assumption 4.1 holds, then the state-action aggregator Q defined in* (13) *possesses a strict lower solution w_1 and an upper solution w_2 in the sense that*

- *(SL) there exists an $\varepsilon > 0$ such that $Q((s,z),y,w_1) \geqslant w_1(s,z) + \varepsilon\kappa(s,z)^\theta$ for all $((s,z),y) \in \mathbb{G}$.*
- *(U) $Q((s,z),y,w_2) \leqslant w_2(s,z)$ for all $((s,z),y) \in \mathbb{G}$.*

---

[20] For more details of the relevant results of such ψ, please refer to section 3.3.

[21] The *ℓ-norm* of f is defined by $\|f\|_\ell := \sup_{(s,z)\in\mathsf{S}\times\mathsf{Z}}\{|f(s,z)|/\ell(s,z)\}$. It is worth noting that when the weight function ℓ is bounded, the ℓ-norm $\|\cdot\|_\ell$ and the supremum norm $\|\cdot\|$ are equivalent. Therefore, the weighted supremum norms become relevant when ℓ is unbounded.

Proof of lemma 6.16. Observe that, for fixed $((s,z),y) \in \mathbb{G}$, we have

$$\begin{aligned} Q((s,z),y,w_1) &= \left\{ r(s,y,z) + \beta \left[\int (L-\delta)^\theta \cdot \kappa(y,z') P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^\theta \\ &\geqslant \left\{ L\kappa(s,z) + \beta \left[\int (L-\delta)^\theta \cdot \kappa(y,z') P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^\theta \\ &\geqslant \left\{ L\kappa(s,z) + \beta(L-\delta) \left[d\kappa(s,z) \right]^{1/\theta} \right\}^\theta \qquad (51) \\ &\geqslant \left\{ \left[L + \beta(L-\delta) d^{1/\theta} \right] \cdot \kappa(s,z)^{1/\theta} \right\}^\theta, \end{aligned}$$

where the first and second inequalities immediately follow from (38) and (40) in assumption 4.1, respectively, while the last one from the fact that $\kappa^{1/\theta} \leqslant \kappa$. Further, with some rearranging, we obtain

$$Q((s,z),y,w_1) \geqslant \left[L - \delta + \beta L d^{1/\theta} + \delta(1 - \beta d^{1/\theta}) \right]^\theta \kappa(s,z) > (L-\delta)^\theta \kappa(s,z),$$

and the last term is equal to $w_1(s,z)$. Similarly, we have

$$\begin{aligned} Q((s,z),y,w_2) &= \left\{ r(s,y,z) + \beta \left(\frac{M}{1-\beta c^{1/\theta}} \right) \left[\int \kappa(y,z')^\theta P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^\theta \\ &\leqslant \left\{ M\kappa(s,z) + \beta \left(\frac{M}{1-\beta c^{1/\theta}} \right) \left[c\kappa(s,z)^\theta \right]^{1/\theta} \right\}^\theta \\ &= \left[\frac{M}{1-\beta c^{1/\theta}} \right]^\theta \kappa(s,z)^\theta = w_2(s,z) \end{aligned}$$

where the inequality follows from (37) and (39) in assumption 4.1. Hence condition (U) of lemma 6.16 is satisfied.

So far we only show w_1 is a lower solution of Q. It remains to prove that it is a strict lower solution. Observe from (51) that to show condition (SL), it is sufficient to show that there exists an $\varepsilon > 0$ such that

$$\left\{ L\kappa(s,z) + \beta(L-\delta) \left[d\kappa(s,z) \right]^{1/\theta} \right\}^\theta \geqslant w_1(s,z) + \varepsilon\kappa(s,z)^\theta \qquad (52)$$

for all $(s,z) \in \mathsf{S} \times \mathsf{Z}$. To this end, for fixed $(s,z) \in \mathsf{S} \times \mathsf{Z}$, consider

$$
\begin{aligned}
&\frac{\left\{L\kappa(s,z) + \beta(L-\delta)\left[d\kappa(s,z)\right]^{1/\theta}\right\}^{\theta} - w_1(s,z)}{\kappa(s,z)^{\theta}} \\
&= \left\{L + \beta(L-\delta)d^{1/\theta} \cdot \kappa(s,z)^{1/\theta - 1}\right\}^{\theta} - (L-\delta)^{\theta}\kappa(s,z)^{1-\theta} \\
&\geqslant L^{\theta} - (L-\delta)^{\theta}\kappa(s,z)^{1-\theta} \geqslant L^{\theta} - (L-\delta)^{\theta} > 0
\end{aligned}
$$

where the first and second inequalities follow from the facts that $\kappa^{1/\theta-1} \geqslant 0$ and that $\kappa^{1-\theta} \leqslant 1$, respectively. Hence, condition (52) holds when we take $\varepsilon := L^{\theta} - (L-\delta)^{\theta}$, which is what we needed to show for condition (SL). $\square$

Lemma 6.17. *If assumption 4.1 holds, the map $((s,z),y) \mapsto Q((s,z),y,\hat{v})$ is continuous on $\mathbb{G}$ whenever $\hat{v} \in \mathscr{C}$.*

Proof of lemma 6.17. To see that this is so, pick any $\hat{v} \in \mathscr{C}$. By assumption 4.1, and by virtue of lemma 12.2.20 in Stachurski (2009), we know that $(y,z) \mapsto \int \hat{v}(y,z')P(z,\mathrm{d}z')$ is continuous on $\mathsf{S} \times \mathsf{Z}$. It then follows from the continuity of r that the map $((s,z),y) \mapsto Q((s,z),y,\hat{v})$ is continuous on $\mathbb{G}$, as was to be shown. $\square$

Recall that the σ-value operator T_σ on $\mathscr{V}$ is defined by

$$
T_\sigma \hat{v}(s,z) = Q((s,z),\sigma(s,z),\hat{v}) = \left\{r_\sigma(s,z) + \beta\left[\int \hat{v}(\sigma(s,z),z')P(z,\mathrm{d}z')\right]^{1/\theta}\right\}^{\theta}
$$

for all $(s,z) \in \mathsf{S} \times \mathsf{Z}$ and $\hat{v} \in \mathscr{V}$, where $r_\sigma(s,z) := r(s,\sigma(s,z),z)$.

Lemma 6.18. *If assumption 4.1 holds, then, for each $\sigma \in \Sigma$, the operator T_σ is asymptotically stable on $\mathscr{V}$.*

Proof of lemma 6.18. First, we show that T_σ is a self-map on $\mathscr{V}$. Fix $\hat{v}$ in $\mathscr{V}$. Together with continuity of r, measurabilities of $\hat{v}$ and σ imply that $T_\sigma \hat{v}$ is Borel measurable on $\mathsf{S} \times \mathsf{Z}$. In addition, since $w_1 \leqslant \hat{v}$, making use of isotonicity of Q (as shown in lemma 6.15), we have $w_1(s,z) \leqslant Q((s,z),\sigma(s,z),w_1) \leqslant Q((s,z),\sigma(s,z),\hat{v})$ for all $(s,z) \in \mathsf{S} \times \mathsf{Z}$, which in turn implies that $w_1 \leqslant T_\sigma \hat{v}$. A similar argument gives $T_\sigma \hat{v} \leqslant w_2$. Therefore, $T_\sigma \hat{v} \in \mathscr{V}$, as was to be shown.

Now, invoking lemmas 6.15 and 6.16, theorem 3.1 of Du (1990) applies and implies the stated result. □

Proof of proposition 4.1. This follows immediately from lemma 6.18. □

Given $\hat{v} \in \mathscr{V}$, a policy σ in Σ will be called *$\hat{v}$-greedy* if

$$\sigma(s,z) \in \operatorname*{argmin}_{y \in \Gamma(s,z)} Q((s,z), y, \hat{v}) \text{ for all } (s,z) \in \mathsf{S} \times \mathsf{Z}.$$

Lemma 6.19. *If $\hat{v} \in \mathscr{C}$, then there exists at least one $\hat{v}$-greedy policy.*

Proof. The proof is identical to that of lemma 6.6. □

Lemma 6.20. *If assumption 4.1 holds, then the operator S is asymptotically stable on $\mathscr{C}$.*

Proof of lemma 6.20. By virtue of lemma 6.17, it follows from Berge's theorem of the minimum that, when $\hat{v}$ is in $\mathscr{C}$, we have

$$S\hat{v}(s,z) = \min_{y \in \Gamma(s,z)} Q((s,z), y, \hat{v}) = \min_{y \in \Gamma(s,z)} \left\{ r(s,y,z) + \beta \left[\int \hat{v}(y,z') P(z, \mathrm{d}z') \right]^{1/\theta} \right\}^{\theta}$$

and $S\hat{v}$ is an element of $\mathscr{C}$.

In order to apply Du's theorem to the Bellman operator S, it suffices to show that

(i) S is isotone and concave on $\mathscr{C}$, and
(ii) $Sw_1 \gg w_1$ and $Sw_2 \leqslant w_2$.[22]

Regarding part (i), making use of the result of lemma 6.15, the proof is essentially identical to that of lemma 6.7. In addition, making use of the result of lemma 6.16, the proof of part (ii) is also essentially identical to that of lemma 6.7. □

Proof of theorem 4.2. By virtue of lemmas 6.18 to 6.20, applying the proof of theorem 6.8 yields the stated results in theorem 4.2. □

[22] The symbol $\gg$ denotes the strong partial order in the Banach space $b_{\kappa^\theta}c(\mathsf{S} \times \mathsf{Z})$ of all κ^θ-bounded continuous functions on $\mathsf{S} \times \mathsf{Z}$, in the sense that $w \gg v$ means $w - v$ lies in the interior of $b_{\kappa^\theta}c(\mathsf{S} \times \mathsf{Z})_+$. For more details, please refer to Guo et al. (2004) or Zhang (2012).

References

ALIPRANTIS, C. D. AND K. C. BORDER (2006): *Infinite dimensional analysis*, Springer, Berlin, 3rd ed.

BACKUS, D. K., B. R. ROUTLEDGE, AND S. E. ZIN (2004): "Exotic preferences for macroeconomists," *NBER Macroeconomics Annual*, 19, 319–390.

BALBUS, L. (2016): "On non-negative recursive utilities in dynamic programming with non-linear aggregator and CES," Working paper, Available at SSRN: https://ssrn.com/abstract=2703975.

BANSAL, R. AND A. YARON (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 59, 1481–1509.

BASU, S. AND B. BUNDICK (2017): "Uncertainty Shocks in a Model of Effective Demand," *Econometrica*, 85, 937–958.

BÄUERLE, N. AND A. JAŚKIEWICZ (2018): "Stochastic optimal growth model with risk sensitive preferences," *Journal of Economic Theory*, 173, 181–200.

BECKER, R. A. AND J. P. RINCÓN-ZAPATERO (2017): "Recursive Utility and Thompson Aggregators," Tech. rep., CAEPR (Center for Applied Economics and Policy Research), working paper, Available at SSRN: https://ssrn.com/abstract=3007788.

BELLMAN, R. (1957): *Dynamic programming*, Academic Press.

BERTSEKAS, D. P. (2013): *Abstract dynamic programming*, Athena Scientific.

BHANDARI, A., D. EVANS, AND M. GOLOSOV (2017): "Risk and Monetary Policy in a New Keynesian Model," Tech. rep.

BICH, P., J.-P. DRUGEON, AND L. MORHAIM (2018): "On temporal aggregators and dynamic programming," *Economic Theory*, 66, 787–817.

BLACKWELL, D. (1965): "Discounted dynamic programming," *The Annals of Mathematical Statistics*, 36, 226–235.

BLOISE, G. AND Y. VAILAKIS (2018): "Convex dynamic programming with (bounded) recursive utility," *Journal of Economic Theory*, 173, 118–141.

BOYD, J. H. (1990): "Recursive utility and the Ramsey problem," *Journal of Economic Theory*, 50, 326–345.

DATTA, M., L. J. MIRMAN, AND K. L. REFFETT (2002): "Existence and uniqueness of equilibrium in distorted dynamic economies with capital and labor," *Journal of Economic Theory*, 103, 377–410.

DIBENEDETTO, E. (2002): *Real Analysis*, Birkhäuser Boston.

DU, Y. (1989): "A fixed point theorem for a class of non-compact operators with application," *Acta Mathematica Sinica, Chinese Series*, 32, 618–627.

——— (1990): "Fixed points of increasing operators in ordered Banach spaces and allications," *Applicable Analysis*, 38, 1–20.

DURÁN, J. (2003): "Discounting long run average growth in stochastic dynamic programs," *Economic Theory*, 22, 395–413.

EPSTEIN, L. G. AND S. E. ZIN (1989): "Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

FARHI, E. AND I. WERNING (2008): "Optimal savings distortions with recursive preferences," *Journal of Monetary Economics*, 55, 21–42.

GOTTARDI, P., A. KAJII, AND T. NAKAJIMA (2015): "Optimal Taxation and Debt with Uninsurable Risks to Human Capital Accumulation," *American Economic Review*, 105, 3443–3470.

GUO, D., Y. J. CHO, AND J. ZHU (2004): *Partial ordering methods in non-linear problems*, Nova Science Publishers.

GUO, J. AND X. D. HE (2017): "Recursive Utility with Investment Gains and Losses: Existence, Uniqueness, and Convergence," Working paper, Available at SSRN: https://ssrn.com/abstract=2790768.

HALL, R. E. (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96, 339–357.

HANSEN, L. P. AND T. J. SARGENT (2008): *Robustness*, Princeton university press.

HANSEN, L. P. AND J. A. SCHEINKMAN (2012): "Recursive utility in a Markov environment with stochastic growth," *Proceedings of the National Academy of Sciences*, 109, 11967–11972.

HARDY, G. H., J. E. LITTLEWOOD, AND G. PÓLYA (1934): *Inequalities*, Cambridge at The University Press.

HAYASHI, T. AND J. MIAO (2011): "Intertemporal substitution and recursive smooth ambiguity preferences," *Theoretical Economics*, 6, 423–472.

JU, N. AND J. MIAO (2012): "Ambiguity, learning, and asset returns," *Econometrica*, 80, 559–591.

KAPLAN, G. AND G. L. VIOLANTE (2014): "A model of the consumption response to fiscal stimulus payments," *Econometrica*, 82, 1199–1239.

KARANTOUNIAS, A. G. (2018): "Optimal Fiscal Policy with Recursive Preferences," *The Review of Economic Studies*, 85, 2283–2317.

KLIBANOFF, P., M. MARINACCI, AND S. MUKERJI (2009): "Recursive smooth ambiguity preferences," *Journal of Economic Theory*, 144, 930–976.

KRASNOSELSKII, M. (1964): *Positive solutions of operator equations*, Noordhoff.

KREPS, D. M. AND E. L. PORTEUS (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200.

LE VAN, C., L. MORHAIM, AND Y. VAILAKIS (2008): "Monotone concave operators: An application to the existence and uniqueness of solutions to the Bellman equation," Working paper, Available at HAL: https://hal.archives-ouvertes.fr/hal-00294828.

LE VAN, C. AND Y. VAILAKIS (2005): "Recursive utility and optimal growth with bounded or unbounded returns," *Journal of Economic Theory*, 123, 187–209.

LUCAS, R. E. AND N. L. STOKEY (1984): "Optimal growth with many consumers," *Journal of Economic Theory*, 32, 139 – 171.

MARINACCI, M. AND L. MONTRUCCHIO (2010): "Unique solutions for stochastic recursive utilities," *Journal of Economic Theory*, 145, 1776–1804.

——— (2017): "Unique Tarski Fixed Points," Tech. rep., working paper, Available at IDEAS: ftp://ftp.igier.unibocconi.it/wp/2017/604.

MARTINS-DA ROCHA, V. F. AND Y. VAILAKIS (2010): "Existence and uniqueness of a fixed point for local contractions," *Econometrica*, 78, 1127–1141.

——— (2013): "Fixed point for local contractions: Applications to recursive utility," *International Journal of Economic Theory*, 9, 23–33.

MORAND, O. F. AND K. L. REFFETT (2003): "Existence and uniqueness of equilibrium in nonoptimal unbounded infinite horizon economies," *Journal of Monetary Economics*, 50, 1351–1373.

OZAKI, H. AND P. A. STREUFERT (1996): "Dynamic programming for non-additive stochastic objectives," *Journal of Mathematical Economics*, 25, 391 – 442.

PETROSKY-NADEAU, N., L. ZHANG, AND L.-A. KUEHN (2018): "Endogenous Disasters," *American Economic Review*, 108, 2212–2245.

RINCÓN-ZAPATERO, J. P. AND C. RODRÍGUEZ-PALMERO (2007): "Recursive utility with unbounded aggregators," *Economic Theory*, 33, 381–391.

RINCÓN-ZAPATERO, J. P. AND C. RODRÍGUEZ-PALMERO (2007): "Recursive utility with unbounded aggregators," *Economic Theory*, 33, 381–391.

SCHORFHEIDE, F., D. SONG, AND A. YARON (2018): "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach," *Econometrica*, 86, 617–654.

STACHURSKI, J. (2009): *Economic Dynamics: Theory and Computation*, MIT Press.

STOKEY, N., R. LUCAS, AND E. PRESCOTT (1989): *Recursive Methods in Economic Dynamics*, Harvard University Press.

ZHANG, Z. (2012): *Variational, topological, and partial order methods with their applications*, vol. 29, Springer Science & Business Media.