

Dynamic Programming with State-Dependent Discounting¹

John Stachurski^a and Junnan Zhang^b

^{a, b} Research School of Economics, Australian National University

September 2, 2019

ABSTRACT. This paper extends the core results of discrete time infinite horizon dynamic programming theory to the case of state-dependent discounting. The traditional constant-discount condition requires that the discount factor of the controller is strictly less than one. Here we replace the constant factor with a discount factor process and require, in essence, that the process is strictly less than one on average in the long run. We prove that, under this condition, the standard optimality results can be recovered, including Bellman's principle of optimality, convergence of value function iteration and convergence of policy function iteration. We also show that the condition cannot be weakened in many standard settings. The dynamic programming framework considered in the paper is general enough to contain features such as recursive preferences. Several applications are discussed.

Keywords: Dynamic programming; optimality; state-dependent discounting

JEL Classification: C61, C62

Driven by the need to better match a variety of empirical phenomena, researchers in economics and finance now routinely adopt settings where the subjective discount rate used by agents in their models varies with the state. For example, [Albuquerque et al. \(2016\)](#) propose an asset pricing model in which the one-period discount factor follows an AR(1) process. They show that the resulting demand shocks help explain the equity premium puzzle. [Mehra and Sah \(2002\)](#) find that small fluctuations in agents' discount factors can have large effect on equity price volatility. Other studies of asset pricing with state-dependent discount factors embedded in preferences include [Campbell \(1986\)](#), [Albuquerque et al. \(2015\)](#) and [Schorfheide et al. \(2018\)](#).

State-dependent and time-varying discount rates are also common in studies of savings, income and wealth. An early example is [Krusell and Smith \(1998\)](#). More recently, [Krusell et al. \(2009\)](#) model the discount factor process as a three state Markov chain and show how discount factor dispersion helps their heterogeneous agent model match

¹We thank Andrzej Nowak for his helpful comments. The first author gratefully acknowledges financial support from ARC grant FT160100423.

Email: john.stachurski@anu.edu.au, junnan.zhang@anu.edu.au

the wealth distribution. [Hubmer et al. \(2018\)](#) model discount factor dynamics using a discretized AR(1) process. [Fagereng et al. \(2019\)](#) use time-varying discount rates and portfolio adjustment frictions to explain the positive correlation between savings rates and wealth observed in Norwegian panel data.

State-dependent discounting is also found in analysis of fiscal and monetary policy. For example, [Eggertsson and Woodford \(2003\)](#) study monetary policy in the presence of zero lower bound restrictions with dynamic time preference shocks. [Woodford \(2011\)](#) considers the government expenditure multiplier in a similar environment. [Eggertsson \(2011\)](#) and [Christiano, Eichenbaum, and Rebelo \(2011\)](#) study the effect of fiscal policies at the zero lower bound on interest rates, while [Nakata and Tanaka \(2016\)](#) analyze the term structure of interest rates at the zero lower bound when agents have recursive preferences. In all of these models, state-dependent variation in discount rates plays a significant role.²

Among other fields, state-dependent discounting is also used regularly in studies of the business cycle and macroeconomic volatility. For example, [Primiceri et al. \(2006\)](#) argue that shocks to agents' rates of intertemporal substitution are a key source of macroeconomic fluctuations. [Justiniano and Primiceri \(2008\)](#) study the shifts in the volatility of macroeconomic variables in the US and find that a large portion of consumption volatility can be attributable to the variance in discount factors.³

In labor economics, state-dependent discounting has been adopted to help explain the excess unemployment volatility puzzle discussed in [Shimer \(2005\)](#). For example, [Mukoyama \(2009\)](#) enhances the Diamond–Mortensen–Pissarides model with state dependent discount factors for entrepreneurs and workers, which is then shown to increase unemployment volatility. Related analysis and extensions can be found in [Beraja et al. \(2016\)](#), [Hall \(2017\)](#) and [Kehoe et al. \(2018\)](#).

The standard theory of infinite horizon dynamic programming (also called the theory of Markov decision processes; see, e.g., [Bellman \(1957\)](#), [Blackwell \(1965\)](#), [Stokey et al. \(1989\)](#), [Puterman \(2014\)](#) or [Bertsekas \(2017\)](#)) does not accommodate state-dependent discounting. Instead, it assumes either zero discounting (and considers long-run average optimality) or a constant and positive discount rate, which corresponds to a discount factor strictly less than one. This implies that, in the canonical setting, the Bellman operator satisfies the conditions of Banach's contraction mapping theorem.

²Additional work from the same field with state-dependent discounting can be found in [Correia et al. \(2013\)](#), [Hills et al. \(2016\)](#), [Hills and Nakata \(2018\)](#) and [Williamson \(2019\)](#).

³Additional research in a similar vein can be found in [Justiniano et al. \(2010\)](#), [Justiniano et al. \(2011\)](#), [Christiano et al. \(2014\)](#), [Saijo \(2017\)](#), and [Bhandari et al. \(2013\)](#).

This contractive property in turn provides the foundations for a powerful optimality theory.

In this paper, we reconsider the theory of infinite horizon discrete time dynamic programming when the usual constant discount factor β is replaced by a discount factor process $\{\beta_t\}$. We then replace the traditional condition $\beta < 1$ for the constant discount case with the more general condition $r(L_\beta) < 1$, where $r(L_\beta)$ is the spectral radius of an operator generated by $\{\beta_t\}$. We show that, when this condition holds, the value function satisfies the Bellman equation, an optimal policy exists, and Bellman’s principle of optimality is valid. Moreover, value function iteration converges to the value function, as it does in the standard case, and Howard’s policy iteration algorithm is also convergent. We consider several applications of the theory, ranging from simple household problems to recursive preference models and optimal stopping problems. We finish with some extensions related to unbounded rewards.

The condition $r(L_\beta) < 1$ is, in several ways, the most natural generalization of the standard condition $\beta < 1$ from the constant discount case. For example, if $\{\beta_t\}$ happens to be constant at some value β , then, as we show below, $r(L_\beta) = \beta$, so that, in particular $r(L_\beta) < 1$ if and only if $\beta < 1$. More generally, if $\{\beta_t\}$ is IID with common mean $\bar{\beta}$, then $r(L_\beta) = \bar{\beta}$. At the same time, if $\{\beta_t\}$ is positively correlated, then, as we show below, $r(L_\beta)$ tends to increase not just with the unconditional mean of the process $\{\beta_t\}$, but also, with the strength of correlation and the amount of volatility. This provides insight not just into the determination of $r(L_\beta)$, but also into the strength of discounting as perceived by the controller across different specifications of the discount factor process.

We show that, unless one focuses on special cases, the condition $r(L_\beta) < 1$ cannot be significantly weakened. This matters for two reasons. One is that the theoretical results contained here have some permanent relevance. The other is that the weakness of the condition allows us to treat topical applications in the quantitative research literature. For example, the condition $r(L_\beta) < 1$ does not rule out the possibility that $\beta_t \geq 1$ with positive probability. This matters because, in some of the new Keynesian literature, the discount factor is allowed to temporarily attain or exceed one so that the zero lower bound on the nominal interest rates binds (see, e.g., [Christiano et al. \(2011\)](#), [Eggertsson \(2011\)](#) or [Correia et al. \(2013\)](#)). Other studies use an AR(1) specification for the log of the discount factor process (or the process itself), which allows the discount factor to become arbitrarily large.⁴

⁴See, for example, [Justiniano and Primiceri \(2008\)](#), [Justiniano et al. \(2010\)](#), [Justiniano et al. \(2011\)](#), [Christiano et al. \(2014\)](#), [Saijo \(2017\)](#) or [Schorfheide et al. \(2018\)](#). Although quantitative

In terms of computation, when $\{\beta_t\}$ takes only finitely many values, $r(L_\beta)$ reduces to the spectral radius of a matrix, and hence can be obtained by standard algorithms from numerical linear algebra. If $\{\beta_t\}$ is a continuous process, then the same ideas can be applied after discretization. One such application is considered in the paper.

This paper is not the first to reconsider dynamic programming problems when the discount factor is allowed to vary over time. For example, [Karni and Zilcha \(2000\)](#) study the saving behavior of agents with random discount factors in a steady-state competitive equilibrium. [Cao \(2018\)](#) proves the existence of sequential and recursive competitive equilibria in incomplete markets with aggregate shocks in which agents also have state-dependent discount factors. In the mathematics literature, [Wei and Guo \(2011\)](#), [Carmon and Shwartz \(2009\)](#), [Minjárez-Sosa \(2015\)](#), [Ilhuicatzí-Roldán et al. \(2017\)](#) and [González-Sánchez et al. \(2019\)](#) all address various issues in dynamic programming with state-dependent discounting. However, these papers assume that the discount factor process in the dynamic program is bounded above by some constant b such that $b < 1$. This is too strict for a range of valuable applications, as discussed above.

One paper from the mathematics literature that admits state-dependent discounting under weak conditions is [Schäl \(1975\)](#). He obtains some results on optimality, including existence of optimal policies and Bellman’s principle of optimality, under a condition that bounds the tail of expected discounted rewards. Although this condition is general, it directly assumes that expected discounted rewards are finite under any Markov policy, and hence restricts all primitives in the dynamic program simultaneously. This makes the condition difficult to test in applications. (No applications are provided in the paper.) In contrast, our condition is designed to be weak enough to include all of the economic applications listed above, while at the same time providing a simple, practical condition that can be tested in applications. Another advantage of our approach is that it allows for recursive preferences and unbounded rewards.

Our work is also related to several recent contributions in the literature on optimal savings that allow for relatively general discount factor processes. For example, an innovative study by [Toda \(2018\)](#) investigates an income fluctuation problem in which the agent has CRRA utility and obtains a necessary and sufficient condition for the existence of a solution to the optimal saving problem with state-dependent discount

implementations imply a finite upper bound, cases where $\beta_t \geq 1$ with positive probability at the benchmark parameterizations still occur. For example, [Hills et al. \(2016\)](#) assume that $\beta_t = \rho\beta_{t-1} + (1 - \rho)\mu + \mu\epsilon_t$, where $\mu = 0.996$, $\rho = 0.8$, and $\sigma_\epsilon = 0.0024$. In their discretization, the largest value for β is 1.006 (see p. 41).

factors. The CRRA restriction is relaxed in [Ma et al. \(2019\)](#). However, their techniques are based around methods for household problems that have no natural analog in the general theory of dynamic programming considered here. Neither of these papers attempts to provide a general theory of dynamic optimality when the discount rate can vary.⁵

As one extension of our results, we consider the case where rewards can be unbounded—a situation arises frequently in economics but introduces technical problems even in the standard case. One line of argument draws on [Alvarez and Stokey \(1998\)](#), which treats homogeneous dynamic programs when the discount factor is constant. A second line of argument extends the local contraction method of [Rincón-Zapatero and Rodríguez-Palmero \(2003\)](#), [Martins-da Rocha and Vailakis \(2010\)](#), [Matkowski and Nowak \(2011\)](#), and [Jaśkiewicz and Nowak \(2011\)](#) to the case of state-dependent discounting.

The rest of this paper is structured as follows. Section 1 sets out the model. Section 2 provides our main results. Section 3 reviews our key assumption. Section 4 gives applications, Section 5 discusses two extensions and Section 6 concludes.

1. A DYNAMIC PROGRAM

We begin with an extension of the abstract dynamic framework of [Bertsekas \(2013\)](#), which can handle non-separable features such as recursive preferences. Throughout, if Y is any metric space, then bmY and bcY denote the bounded (Borel) measurable and bounded continuous functions from Y to \mathbb{R} respectively.

1.1. Framework. The state of the world consists of a pair (x, z) , where x represents a set of endogenous variables, affected by the actions of an agent (the controller), and z is an exogenous state. These variables take values in metric spaces X and Z respectively. For convenience, we set $S := X \times Z$. The agent responds to the current state by taking action a from an action space A , which is a separable metric space. The choice a when facing state (x, z) , is restricted to $\Gamma(x, z) \subset A$, where Γ is the *feasible correspondence*. Let $\text{gr } \Gamma$ be all $((x, z), a)$ in $S \times A$ such that a is in $\Gamma(x, z)$. In other words, $\text{gr } \Gamma$ is all *feasible state-action pairs*.

The exogenous state evolves according to a Markov transition kernel Q , so that $Q(z, B)$ represents the probability of transitioning to Borel set $B \subset Z$ next period, given current exogenous state z . Throughout, \mathbb{E}_z represents expectation conditional on $Z_0 = z$.

⁵Also relevant are [Higashi et al. \(2009\)](#), which provides an axiomatic foundation for preferences with random discounting. Other work along this line includes [Krishna and Sadowski \(2014\)](#) and [Higashi et al. \(2017\)](#).

We combine the remaining elements of the dynamic programming problem into a single *continuation aggregator* H , with the understanding that $H(x, z, a, v)$ is the maximal value that can be obtained from the present time under the continuation value function v , given current state (x, z) and current action a . Thus, Bellman's equation takes the form

$$v(x, z) = \sup_{a \in \Gamma(x, z)} H(x, z, a, v). \quad (1)$$

We take \mathcal{V} to be a nonempty set of *candidate value functions*, each one of which is required to be real-valued, measurable and defined on \mathbf{S} . The continuation aggregator H maps each (x, z, a, v) in $\text{gr } \Gamma \times \mathcal{V}$ into a real number.

A *dynamic program* \mathcal{D} is a tuple $(\mathbf{X}, \mathbf{Z}, \mathbf{A}, \Gamma, H, \mathcal{V})$ with the structure imposed above. The next example illustrates in the context of a one-sector growth model.

Example 1.1. Consider the one-sector stochastic optimal growth model as found in, say, [Stokey et al. \(1989\)](#), except that the discount rate is state-dependent. An agent solves

$$\max_{\{C_t, K_t\}_{t=0}^{\infty}} \mathbb{E} \left\{ \sum_{t=0}^{\infty} \prod_{i=0}^{t-1} \beta_i u(C_t) \right\} \quad (2)$$

subject to $C_t = f(K_t, Z_t) - K_{t+1} \geq 0$. Here $\beta_t = \beta(Z_t)$ where β is a nonnegative function, $\{Z_t\}$ is a discrete time Markov process with transition kernel Q , u is a one-period return function and f is a production function. When interpreting (2), we adopt the convention that $\prod_{i=0}^{-1} \beta_i = 1$, and same convention is used for products over all other sequences.

The Bellman equation is

$$v(k, z) = \sup_{0 \leq a \leq f(k, z)} \left\{ u(f(k, z) - a) + \beta(z) \int v(a, z') Q(z, dz') \right\}. \quad (3)$$

This problem can be mapped to our framework by taking capital as the endogenous state, z as the exogenous state and consumption as the action. The endogenous state space \mathbf{X} and the action space \mathbf{A} can be set to \mathbb{R}_+ , while \mathbf{Z} to be some arbitrary metric space. The feasible correspondence $\Gamma(x, z)$ is all a such that $0 \leq a \leq f(x, z)$, and

$$H(x, z, a, v) = u(f(x, z) - a) + \beta(z) \int v(a, z') Q(z, dz'). \quad (4)$$

While the same shock process $\{Z_t\}$ affects both production and the discount rate under this formulation, there is no loss of generality because the shock space \mathbf{Z} is arbitrary and hence can support multiple independent processes.

1.2. Feasibility and Optimality. Given a dynamic program \mathcal{D} , let Σ be the set of *feasible policies*, defined as all Borel measurable maps σ from \mathbf{S} to \mathbf{A} such that $\sigma(x, z) \in \Gamma(x, z)$ for each (x, z) in \mathbf{S} . Given each such σ , let T_σ be the operator on \mathcal{V} given by

$$(T_\sigma v)(x, z) = H(x, z, \sigma(x, z), v). \quad (5)$$

Define the *Bellman operator* T on \mathcal{V} by

$$(Tv)(x, z) = \sup_{a \in \Gamma(x, z)} H(x, z, a, v). \quad (6)$$

Given v_0 in \mathcal{V} and σ in Σ , we can interpret $v_{n, \sigma}(x, z) := (T_\sigma^n v_0)(x, z)$ as the lifetime payoff of an agent who starts at state (x, z) , follows policy σ for n periods and uses v_0 to evaluate the terminal state. The σ -*value function* for an infinite-horizon problem is defined here as

$$v_\sigma(x, z) := \lim_{n \rightarrow \infty} v_{n, \sigma}(x, z).$$

The definition requires that this limit exists and is independent of v_0 . Below we impose conditions such that this is always the case.

We define the *value function* corresponding to our dynamic program by

$$v^*(x, z) = \sup_{\sigma \in \Sigma} v_\sigma(x, z) \quad (7)$$

at each (x, z) in \mathbf{S} . Existence of the value function requires that v_σ is well defined for each $\sigma \in \Sigma$, as well as some upper bound on the value that can be obtained from each policy. When the value function does exist, as it will in our setting, a policy $\sigma^* \in \Sigma$ is called *optimal* whenever it attains the supremum in (7), which is to say that

$$v^*(x, z) = v_{\sigma^*}(x, z) \quad (8)$$

at each (x, z) in \mathbf{S} .

1.3. Regularity. To present sharp optimality results, it is useful to add some degree of regularity via continuity and compactness. To this end, we define a dynamic program \mathcal{D} to be *regular* if

- (a) the set of candidate value functions \mathcal{V} equals $bm\mathbf{S}$,
- (b) the correspondence Γ is continuous, nonempty, and compact valued, and
- (c) the map $(x, z, a) \mapsto H(x, z, a, v)$ is measurable for all $v \in bm\mathbf{S}$, continuous for all $v \in bc\mathbf{S}$, and bounded for at least one $v \in bc\mathbf{S}$.

Many standard cases from the literature are regular.

Example 1.2. The optimal growth model in Example 1.1 is regular if u and f satisfy standard conditions, such as those in Section 5.1 of [Stokey et al. \(1989\)](#), Z is compact, β is a continuous function and Q is Feller.⁶

Example 1.3. All dynamic programs with finite state and action spaces are regular, since conditions (a)–(c) are automatically satisfied when we endow both spaces with the discrete topology.

(Although we obtain some fundamental dynamic programming results without regularity, this discussion is left to extensions and the appendix.)

1.4. Key Assumptions. Consider the following restrictions for the dynamic program \mathcal{D} , the first of which is standard.

Assumption 1.1 (Monotonicity). If $v, w \in \mathcal{V}$ and $v \leq w$, then, for all $(x, z) \in \mathbf{S}$ and $a \in \Gamma(x, z)$, we have $H(x, z, a, v) \leq H(x, z, a, w)$.

The importance of Assumption 1.1 is well known (see, e.g., [Bertsekas, 2013](#), Chapter 2). It clearly holds in (4) when the discount factor function β is nonnegative. In general, Assumption 1.1 is a mild condition that holds for all problems we consider.

The next assumption is new and central to what follows.

Assumption 1.2 (Long-Run Contractivity). There exists a positive function β in bmZ such that, for all $v, w \in \mathcal{V}$, we have

$$|H(x, z, a, v) - H(x, z, a, w)| \leq \beta(z) \int \sup_{x' \in X} |v(x', z') - w(x', z')| Q(z, dz') \quad (9)$$

for all $(x, z) \in \mathbf{S}$ and $a \in \Gamma(x, z)$. Moreover, β is such that the spectral radius $r(L_\beta)$ of the linear operator $L_\beta: bmZ \rightarrow bmZ$ defined by

$$(L_\beta h)(z) = \beta(z) \int_Z h(z') Q(z, dz') \quad (10)$$

satisfies $r(L_\beta) < 1$.

Here bmZ is the set of all bounded measurable functions from Z to \mathbb{R} , paired with the supremum norm. As usual, the spectral radius of a linear operator L from bmZ to itself

⁶ Q is called *Feller* if $v \in bcZ$ implies $z \mapsto \int v(z') Q(z, dz')$ is in bcZ .

is defined by $r(L) = \lim_{n \rightarrow \infty} \|L^n\|^{1/n}$, where $\|\cdot\|$ is the operator norm.⁷ The spectral radius condition from Assumption 1.2 is shown in the appendix to be equivalent to

$$\exists n \in \mathbb{N} \text{ such that } \sup_{z \in \mathcal{Z}} \mathbb{E}_z \prod_{t=0}^{n-1} \beta(Z_t) < 1, \quad (11)$$

where $\{Z_t\}$ is a Markov process generated by Q . In our applications, $\beta_t = \beta(Z_t)$ will have the interpretation of the time t value of the discount factor process.

Assumption 1.2 is the key condition of the paper. It is a generalization of the usual (see, e.g., Bertsekas, 2013, Chapter 2) discounting condition

$$|H(x, z, a, v) - H(x, z, a, w)| \leq b \|v - w\| \quad (12)$$

for some constant $b < 1$. This is easy to see because, if a b satisfying (12) exists, then we can set $\beta \equiv b$ as the function in Assumption 1.2. Evidently (9) then holds. In addition (11) is immediate, so the spectral radius condition $r(L_\beta) < 1$ also holds.

An extended discussion of the spectral radius condition $r(L_\beta) < 1$, including how to test it and how to understand its implications, is provided in Section 3.

2. RESULTS

We can now state our main results. We provide a fundamental optimality result, as well as a discussion of two standard algorithms: value function iteration and policy function iteration.

2.1. Key Optimality Results. In the statement of the next theorem and all of what follows, a map M from a metric space into itself is called *eventually contracting* if there exists an n in \mathbb{N} such that the n -th iterate M^n is a contraction mapping.⁸

Theorem 2.1. *Let \mathcal{D} be a regular dynamic program. If Assumptions 1.1 and 1.2 both hold, then the following statements are true:*

- a. T_σ is eventually contracting on $bm\mathcal{S}$ and T is eventually contracting on $bc\mathcal{S}$.
- b. For each feasible policy σ , the lifetime value v_σ is a well defined element of $bm\mathcal{S}$.
- c. The value function v^* is well defined, bounded and continuous. Moreover, the unique fixed point of T in $bc\mathcal{S}$ is v^* .

⁷This is the Gelfand representation of the spectral radius. See, for example, Krasnosel'skii et al. (1972), Ch. 1.

⁸More precisely, a self-map M on metric space (Y, ρ) is called eventually contracting if there exists an n in \mathbb{N} and a $\lambda < 1$ such that $\rho(M^n y, M^n y') \leq \lambda \rho(y, y')$ for all y, y' in Y .

d. *At least one optimal policy exists.*

e. *A policy $\sigma \in \Sigma$ is optimal if and only if it satisfies*

$$\sigma(x, z) \in \arg \max_{a \in \Gamma(x, z)} H(x, z, a, v^*) \quad \text{for each } (x, z) \text{ in } \mathbf{S}. \quad (13)$$

This theorem extends the core results of dynamic programming theory to the case of state-dependent discounting. In particular, the value function satisfies the Bellman equation, an optimal policy exists, and Bellman’s principle of optimality is valid (i.e., part (e) holds). Iteration with the Bellman operator leads to the value function, so that we have both existence of an optimal policy and a means to compute it.

Relative to the results that can be obtained under the standard one-step contraction condition in (12), as found for instance in Chapter 2 of Bertsekas (2013), the only significant weakening of the main findings is that T and T_σ are eventually contracting, rather than always contracting in one step. This is due to the fact that values of the discount factor greater than one are admitted, for reasons discussed in the introduction. Only in the long run are we guaranteed of contraction.

The proof of Theorem 2.1 is deferred to the appendix. In the proof, we only adopt regularity when necessary, so the results contained in the appendix also provide information on what can be said if the regularity condition is dropped. Throughout the paper we restrict our attention to optimality over stationary policies—that is, the agent chooses the same policy in every period, because stationary policies dominate nonstationary ones. The proof is similar to the standard case and hence omitted. See, for example, Chapter 2 of Bertsekas (2013).

2.2. Policy Iteration. Howard’s policy iteration algorithm generates a sequence of feasible policies in the following way. Choose any $\sigma_0 \in \Sigma$ and define σ_k for $k \in \mathbb{N}$ by

$$T_{\sigma_k} v_{\sigma_{k-1}} = T v_{\sigma_{k-1}}. \quad (14)$$

The next proposition shows that, under the conditions adopted in Theorem 2.1, the values of the policy sequence generated by Howard’s policy iteration algorithm converge to the maximum possible.

Theorem 2.2. *If the conditions of Theorem 2.1 hold and the policy sequence $\{\sigma_k\}$ satisfies (14), then $v_{\sigma_k} \rightarrow v^*$ uniformly as $k \rightarrow \infty$.*

3. UNDERSTANDING THE LONG-RUN CONTRACTIVITY CONDITION

Assumption 1.2 is specifically designed to handle models with state-dependent discount factors. Before moving on to such applications, we examine Assumption 1.2 in greater depth. In particular, we wish to know the implications of and restrictions imposed by the spectral radius condition $r(L_\beta) < 1$ on a given discount factor process $\{\beta_t\}$.

3.1. The Finite State Case. One common sub-case in terms of applications is the setting where the exogenous state space \mathbf{Z} is finite. In this case, the transition probability kernel Q can be represented as a Markov matrix of values Q_{ij} , giving the one-step probability of transitioning from z_i to z_j . As can be seen from its definition in (10), the linear operator L_β can be represented as the matrix

$$L_\beta := (\beta_i Q_{ij})_{1 \leq i, j \leq N}. \quad (15)$$

Here $\beta_i := \beta(z_i)$ and N is the number of elements in \mathbf{Z} . The spectral radius of L_β can be computed using standard routines, first by computing the eigenvalues of the matrix and next by taking the maximum in modulus.

For continuous state exogenous processes, the same procedure can be implemented numerically after discretizing the process. The next section discusses such a case.

3.2. A Representative Discount Factor Process. It is often assumed that the discount factor process obeys an AR(1) specification such as

$$\beta_{t+1} = \rho\beta_t + (1 - \rho)\mu + \sigma_\epsilon\epsilon_{t+1}, \quad (16)$$

where $-1 < \rho < 1$, σ_ϵ is positive and $\{\epsilon_t\}$ is IID and standard normal (see, e.g., Hills et al. (2016), Hubmer et al. (2018) or Schorfheide et al. (2018)). While this approach is problematic in its original form, since the discount process can be negative and, at the same time, is unbounded above, most quantitative analyses circumvent these issues by discretization. In this section we do the same and then investigate the factors that determine the size of $r(L_\beta)$. Discretization is carried out using the Rouwenhorst method, which is well suited to highly correlated Gaussian AR(1) processes.⁹

Our benchmark parameterization is $\mu = 0.985$, $\rho = 0.99$, and $\sigma_\beta = 0.01$, where σ_β is the unconditional standard deviation of β_t computed by $\sigma_\beta = \sigma_\epsilon / (\sqrt{1 - \rho^2})$. At this parameterization, the method for finite exogenous state problems described above yields

⁹See, Kopecky and Suen (2010) for an exposition of the method and some of its properties. We set the number of states $N = 50$, which as Kopecky and Suen (2010) show, is accurate enough for most applications.

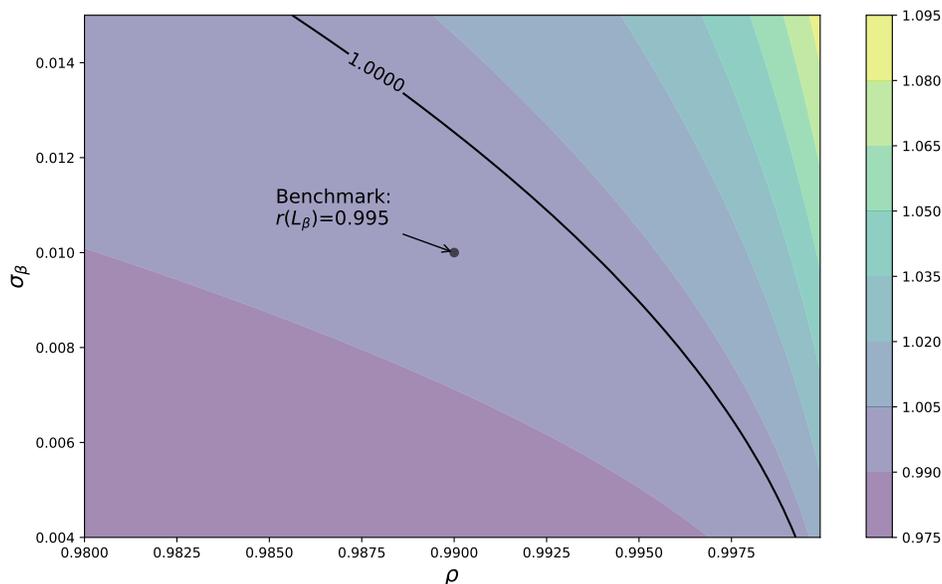


FIGURE 1. $r(L_\beta)$ as a function of ρ and σ_β ; $\mu = 0.985$

$r(L_\beta) = 0.995$. Figure 1 plots values of $r(L_\beta)$ as a function of ρ and σ_β , represented by contour lines, while μ is held constant at its benchmark value. The main message of the figure is that larger volatility and larger persistence both lead to a higher spectral radius for L_β , and the effect of increasing ρ becomes large when σ_β is large.

What is the intuition behind this result? To gain some understanding, consider a generic stationary discount factor process $\{\beta_t\}$ and let $\mu := \mathbb{E}\beta_0$ be the common mean. Suppose that the process exhibits positive autocorrelation. Then, for any t , the variates β_t and β_{t+1} are positively correlated, which is equivalent to the statement that $(\mathbb{E}\beta_t\beta_{t+1})^{1/2} \geq \mu$. For this bivariate case, then, positive correlation pushes the expected geometric time series average above the stationary mean. Moreover, the inequality is strict precisely when the joint distribution is not degenerate.

The situation is even clearer when we focus on the AR(1) model (16). In this case, we have $(\mathbb{E}\beta_t\beta_{t+1})^{1/2} = (\mu^2 + \rho\sigma_\beta^2)^{1/2} \geq \mu$. The key message is the same: positive correlation and positive volatility push the expected geometric time series average above the stationary mean. The impact increases with both ρ and σ_β , and the two effects reinforce each other.

The connection between this and $r(L_\beta)$ is that, as we show in (26) of the appendix, the spectral radius of L_β is in fact equal to an asymptotic long-run version of this geometric average. Putting the pieces together, we can see why increased positive correlation and

increased volatility in the discount factor process tend to drive up $r(L_\beta)$ relative to the stationary mean of β_t , as seen in Figure 1.

3.3. Necessity of the Spectral Radius Condition. Consider a standard Markov decision process (as in, say, Blackwell (1965)), where a controller seeks to maximize an expression of the form $\mathbb{E} \sum_{t \geq 0} \beta^t r_t$. Here $\beta > 0$ is a constant discount factor and $\{r_t\}$ is a bounded sequence of rewards. In this setting, the condition $\beta < 1$ cannot be relaxed without imposing specific conditions on rewards. For example, if there are constants $0 < a \leq b$ such that $a \leq r_t \leq b$ for all t , then we clearly have¹⁰

$$\mathbb{E} \sum_{t \geq 0} \beta^t r_t < \infty \text{ if and only if } \beta < 1. \quad (17)$$

The condition $r(L_\beta) < 1$ has this same distinction once we replace the constant discount factor β with a discount factor process $\{\beta_t\}$, as long as we are prepared to impose some regularity conditions. The next proposition gives one such result. In stating it, the process $\{\beta_t\}$ and the operator L_β are as defined in Assumption 1.2, while $\{r_t\}$ is as just described.

Proposition 3.1. *If the exogenous state space Z is compact and β is continuous, then*

$$\mathbb{E} \sum_{t \geq 0} \prod_{i=0}^{t-1} \beta_i r_t < \infty \text{ if and only if } r(L_\beta) < 1. \quad (18)$$

4. APPLICATIONS

In this section we turn to applications of Theorems 2.1–2.2.

4.1. One-Sector Stochastic Optimal Growth. Recall the stochastic optimal growth model with state-dependent discounting from Example 1.1. If u and f satisfy standard conditions, as in, say, Section 5.1 of Stokey et al. (1989), then the dynamic program is regular. The monotonicity condition in Assumption 1.1 certainly holds. Regarding Assumption 1.2, since H is given by (4), we have

$$\begin{aligned} |H(x, z, a, v) - H(x, z, a, w)| &= \beta(z) \left| \int [v(a, z') - w(a, z')] Q(z, dz') \right| \\ &\leq \beta(z) \int |v(a, z') - w(a, z')| Q(z, dz') \\ &\leq \beta(z) \int \sup_{x' \in X} |v(x', z') - w(x', z')| Q(z, dz'). \end{aligned}$$

¹⁰The equivalence in (17) is easy to see because, by the Monotone Convergence Theorem, we have $\mathbb{E} \sum_{t \geq 0} \beta^t r_t = \sum_{t \geq 0} \beta^t \mathbb{E} r_t$ and, moreover, $0 < a \leq \mathbb{E} r_t \leq b$.

Here the first inequality is by the triangle inequality for integrals and the second is obvious. Hence, if the process β_t satisfies $r(L_\beta) < 1$, then Assumption 1.2 also holds, and so do the conclusions of Theorems 2.1–2.2. Conditions under which $r(L_\beta) < 1$ holds were discussed in Section 3.

4.2. Job Search. Our framework is also able to deal with optimal stopping problems with proper definitions of the primitives. In this section, we demonstrate this using an elementary job search problem in McCall (1970) except that the agent has state-dependent discount factors. The basic structure considered here can be modified to deal with more complicated optimal stopping problems, such as Lucas and Prescott (1974) and Robin (2011).

An unemployed worker searching for a job is given a wage offer every period. He can accept the offer and receive this wage every period forever, or he can choose to receive an unemployment compensation c and wait for another offer next period. Uncertainty is driven by a Markov process on a metric space \mathbf{Z} with stochastic kernel Q . The wage offer is given by a function $w : \mathbf{Z} \rightarrow \mathbb{R}_+$. Since the discount factors are state dependent, the lifetime utility of accepting an offer at state z is $K(z)w(z)$ where

$$K(z) := \sum_{t=0}^{\infty} \left(\mathbb{E}_z \prod_{i=0}^{t-1} \beta(Z_i) \right), \quad \forall z \in \mathbf{Z}.$$

The Bellman equation is thus

$$v(z) = \max \left\{ w(z)K(z), c + \beta(z) \int_{\mathbf{Z}} v(z')Q(z, dz') \right\}.$$

We can translate this problem to our framework by letting $\mathbf{S} = \mathbf{Z}$, $\mathbf{A} = \{0, 1\}$, $\Gamma \equiv \mathbf{A}$, and

$$H(x, z, a, v) = aw(z)K(z) + (1 - a) \left[c + \beta(z) \int_{\mathbf{Z}} v(z')Q(z, dz') \right].$$

Then $\mathcal{D} = (\mathbf{X}, \mathbf{Z}, \mathbf{A}, \Gamma, H, \mathcal{V})$ is the associated dynamic program with $\mathcal{V} = bm\mathbf{S}$. Note that in this setting, the state space \mathbf{X} is redundant. In particular, the Bellman operator defined from \mathcal{D} is

$$\begin{aligned} (Tv)(z) &= \max_{a \in \{0,1\}} \left\{ aw(z)K(z) + (1 - a) \left[c + \beta(z) \int_{\mathbf{Z}} v(z')Q(z, dz') \right] \right\} \\ &= \max \left\{ w(z)K(z), c + \beta(z) \int_{\mathbf{Z}} v(z')Q(z, dz') \right\}. \end{aligned} \quad (19)$$

We have the following proposition.

Proposition 4.1. *If (i) w and β are bounded and continuous, (ii) $r(L_\beta) < 1$ where L_β is defined in (10), and (iii) Q has the Feller property, then \mathcal{D} is regular and Assumptions 1.1–1.2 hold. In particular, the conclusions of Theorems 2.1–2.2 are valid.*

4.3. A Household Problem with Taxation. Consider the household problem in Hills and Nakata (2018), where $u: \mathbb{R}_+ \rightarrow \mathbb{R}$ is one-period utility and the discount factor at time t is β_t . The agent chooses consumption $\{C_t\}$ and risk-free asset $\{B_t\}$ to solve

$$\max_{\{C_t, B_t\}_{t=0}^{\infty}} \mathbb{E} \left\{ \sum_{t=0}^{\infty} \prod_{i=0}^{t-1} \beta_i u(C_t) \right\}$$

subject to the constraint

$$C_t + \frac{B_{t+1}}{R_t P_t} \leq \frac{B_t}{P_t} + T_t.$$

The interest rate R_t , price level P_t , and lump-sum tax T_t are taken as given by the agent. We suppose that $R_t = R(Z_t)$, $P_t = P(Z_t)$, $T_t = T(Z_t)$ and $\beta_t = \beta(Z_t)$, where $\{Z_t\}$ is a Markov process on metric space Z with transition kernel Q and R, P, T and β are Borel measurable functions. Let $\mathbf{X} = \mathbf{A}$ be a compact subset of \mathbb{R}_+ and assume that R and P are bounded below away from zero. Define

$$H(x, z, a, v) = u(F(x, z, a)) + \beta(z) \int_Z v(a, z') Q(z, dz'),$$

where

$$F(x, z, a) = \frac{x}{P(z)} + T(z) - \frac{a}{R(z)P(z)}$$

corresponds to consumption given current-period asset x , next-period asset a , and shock z . Define the feasible correspondence by

$$\Gamma(x, z) = [0, xR(z) + T(z)R(z)P(z)].$$

Let $\mathbf{S} = \mathbf{X} \times Z$ and let $\mathcal{D} = (\mathbf{X}, Z, \mathbf{A}, \Gamma, H, \mathcal{V})$ be the associated dynamic program, where the class of candidate value functions \mathcal{V} is equal to $bm\mathbf{S}$.

Proposition 4.2. *If (i) R, P, T , and β are bounded and continuous, (ii) u is continuous, (iii) $r(L_\beta) < 1$ where L_β is defined in (10), and (iv) Q has the Feller property, then \mathcal{D} is regular and Assumptions 1.1–1.2 hold. In particular, the conclusions of Theorems 2.1–2.2 are valid.*

4.4. Recursive Preferences. Consider the Epstein-Zin utility with discount factor shocks studied in [Albuquerque et al. \(2016\)](#). The lifetime utility¹¹ of the agent at time t has a recursive formulation:

$$U_t = \max_{C_t} \left\{ C_t^{1-\rho} + \beta_t [\mathbb{E}_t U_{t+1}^{1-\gamma}]^{\frac{1-\rho}{1-\gamma}} \right\}^{\frac{1}{1-\rho}}.$$

Uncertainty in the economy is driven by a sequence of Markovian shocks $\{Z_t\}$ on Z with stochastic kernel Q . There is a single risky asset in unit supply with price $p(z)$ and dividend $d(z)$ given shock z . In each period, the agent maximizes lifetime utility subject to budget constraint

$$C_t + p(Z_t)X_{t+1} \leq d(Z_t)X_t + p(Z_t)X_t$$

where X_t is asset holding of the agent at time t . Given d and p , we seek a solution to the Bellman equation

$$v(x, z) = \max_{a \in \Gamma(x, z)} \left\{ F(x, z, a)^{1-\rho} + \beta(z) \left[\int v(a, z')^{1-\gamma} Q(z, dz') \right]^{\frac{1-\rho}{1-\gamma}} \right\}^{\frac{1}{1-\rho}}, \quad (20)$$

where $F(x, z, a) = d(z)x + p(z)(x - a)$ corresponds to consumption given current-period asset x , next-period asset a , and shock z , and

$$\Gamma(x, z) = \left\{ a : 0 \leq a \leq \min \left[1, \frac{d(z)x + p(z)x}{p(z)} \right] \right\}$$

is the constraint¹² on asset holding. We consider the case where $\rho < \gamma < 1$ so that utilities are positive and work on $\tilde{v} = v^{1-\gamma}$ instead. Since $\tilde{v} = v^{1-\gamma}$ is a strictly increasing transformation, for positive value functions the two formulations are equivalent.

To study this problem in our theoretical framework, let $\mathbf{X} = \mathbf{A} = \mathbb{R}_+$ and define

$$H(x, z, a, \tilde{v}) = \left\{ F(x, z, a)^{1-\rho} + \beta(z) \left[\int_{\mathbf{Z}} \tilde{v}(a, z') Q(z, dz') \right]^{\frac{1}{\theta}} \right\}^{\theta}$$

where $\theta = (1 - \gamma)/(1 - \rho) \in (0, 1)$. Then (20) can be written as $\tilde{v} = T\tilde{v}$ where T is defined in (6). Let $\mathbf{S} = \mathbf{X} \times \mathbf{Z}$ and $\mathcal{V} = bm\mathbf{S}$. This problem can be formulated as a dynamic program $\mathcal{D} = (\mathbf{X}, \mathbf{Z}, \mathbf{A}, \Gamma, H, \mathcal{V})$. Define operator L_θ by

$$L_\theta h(z) := \beta(z)^\theta \int_{\mathbf{Z}} h(z') Q(z, dz'). \quad (21)$$

We have the following proposition.

¹¹We use an alternative formulation of the recursive utility in their paper; see footnote 9 in [Albuquerque et al. \(2016\)](#).

¹²Asset holding cannot exceed one due to its fixed supply.

Proposition 4.3. *If (i) d , p , and β are bounded and continuous, (ii) $r(L_\theta) < 1$ where L_θ is defined in (21), and (iii) Q has the Feller property, then \mathcal{D} is regular and Assumptions 1.1–1.2 hold. In particular, the conclusions of Theorems 2.1–2.2 are valid.*

5. EXTENSIONS

In this section, we let $\mathsf{X} = \mathsf{A}$, and consider aggregators of the form

$$H(x, z, a, v) = u(x, z, a) + \beta(z) \int_{\mathsf{Z}} v(a, z') Q(z, dz') \quad (22)$$

where u is continuous but not necessarily bounded. All the other primitives are as defined in Section 1.1. Since value functions are not in $bm\mathsf{S}$ in this case, the associated dynamic program \mathcal{D} is no longer regular and most results in Section 2 do not apply.

We shall discuss two approaches that are common in dealing with dynamic programming with unbounded rewards. The first one is in the spirit of [Stokey et al. \(1989, Section 9.3\)](#) and [Alvarez and Stokey \(1998\)](#) that treat homogeneous programs. The second one uses a local contraction method. Each approach considers a different space for candidate value functions and establishes contraction of the Bellman operator on the new space when β is constant. This section aims to show that the Bellman operator is eventually contracting in the face of state-dependent discount factors as long as proper spectral conditions similar to Assumption 1.2 are satisfied.

5.1. Homogeneous Functions. Consider dynamic programming problems with return functions that are homogeneous of degree $\theta \in (0, 1]$ and feasible correspondences that are homogeneous of degree one. In particular, we have the following standard assumption.

Assumption 5.1. The feasible correspondence Γ is continuous, nonempty, and compact valued, and for any $(x, z) \in \mathsf{S}$,

$$a \in \Gamma(x, z) \implies \gamma a \in \Gamma(\gamma x, z), \quad \forall \gamma \geq 0.$$

The return function u is continuous, $u(\cdot, z, \cdot)$ is homogeneous of degree θ , and there exists $B > 0$ such that for any $(x, z) \in \mathsf{S}$

$$|u(x, z, a)| \leq B(\|x\| + \|a\|), \quad \forall a \in \Gamma(x, z).$$

The next assumption is a generalization of the standard growth restriction to problems with state-dependent discount factors. Note that if both β and α are constant, $r(L_\alpha) < 1$ regresses to the familiar condition $\alpha^\theta \beta < 1$.

Assumption 5.2. There exists a bounded measurable function $\alpha : \mathbf{Z} \rightarrow \mathbb{R}_+$ such that

$$\|a\| \leq \alpha(z)\|x\|, \quad \forall a \in \Gamma(x, z), \forall (x, z) \in \mathbf{S}.$$

The operator $L_\alpha : bm\mathbf{Z} \rightarrow bm\mathbf{Z}$ defined by

$$L_\alpha(z) := \beta(z)\alpha^\theta(z) \int_{\mathbf{Z}} h(z')Q(z, dz') \quad (23)$$

satisfies that $r(L_\alpha) < 1$.

Let $(H(\mathbf{S}; \theta), \|\cdot\|)$ be the space of bounded continuous functions that are homogeneous of degree θ with norm defined by

$$\|f\| := \sup_{z \in \mathbf{Z}} \sup_{x \in \mathbf{X}, \|x\|=1} |f(x, z)|. \quad (24)$$

Then $H(\mathbf{S}; \theta)$ is a Banach space (Stokey et al., 1989). We have the following result.

Proposition 5.1. *Let $\mathcal{V} = H(\mathbf{S}; \theta)$ and $\mathcal{D} = (\mathbf{X}, \mathbf{Z}, \mathbf{A}, \Gamma, H, \mathcal{V})$ be the associated dynamic program. Under Assumptions 5.1 and 5.2, the lifetime value $v_\sigma(x_0, z_0)$ is well defined and finite for any initial state (x_0, z_0) and feasible policy σ , the value function v^* is a unique fixed point of T on \mathcal{V} , $T^n v \rightarrow v^*$ for all $v \in \mathcal{V}$, there exists an optimal stationary policy, and the principle of optimality holds.*

Example 5.1. Consider the household saving problem in Toda (2018) without restricting the shock space to be finite. Instead, we assume that the shocks are Markovian with stochastic kernel Q on an arbitrary metric space \mathbf{Z} . The asset return $R(z)$ and discount function $\beta(z)$ are bounded continuous functions of the shocks. The utility function is $u(c) = c^{1-\gamma}/(1-\gamma)$ and we assume that $\gamma \in (0, 1)$. The budget constraint is $X_{t+1} = R(Z_t)(X_t - C_t)$ where X_t is the beginning-of-period wealth, C_t is consumption, and Z_t is shock at time t . The Bellman equation is thus

$$v(x, z) = \sup_{x' \in [0, R(z)x]} u\left(x - \frac{x'}{R(z)}\right) + \beta(z) \int_{\mathbf{Z}} v(x', z')Q(z, dz').$$

Then Assumption 5.1 is satisfied and so is Assumption 5.2 if $r(L_R) < 1$ with L_R defined by

$$L_R(z) := \beta(z)R^{1-\gamma}(z) \int_{\mathbf{Z}} h(z')Q(z, dz').$$

This is equivalent to the spectral radius condition¹³ in Proposition 1 of Toda (2018) if \mathbf{Z} is finite.

¹³The condition in Toda (2018) is $r(DP) < 1$ where $D = \text{diag}(\beta_1 R_1^{1-\gamma}, \dots, \beta_S R_S^{1-\gamma})$ and P is the transition matrix. Also see discussion in Section 3.2.

5.2. Local Contractions. Next we take a local contraction approach to dynamic programs with state dependent discounting and unbounded rewards. We follow the methods pioneered in [Rincón-Zapatero and Rodríguez-Palmero \(2003\)](#), [Martins-da Rocha and Vailakis \(2010\)](#) and [Matkowski and Nowak \(2011\)](#) for the constant discount case. Let $C(\mathbf{S})$ be the space of continuous functions on \mathbf{S} . Assume that \mathbf{Z} is compact and $\mathbf{X} = \bigcup_j \text{Int}(K_j)$ where $\{K_j\}$ is a sequence of strictly increasing and compact subsets of \mathbf{X} . Define $\|\cdot\|_j$ by

$$\|f\|_j = \sup_{x \in K_j, z \in \mathbf{Z}} |f(x, z)|, \quad \forall f \in C(\mathbf{S}).$$

Let $c > 1$ and $\{m_j\}$ be an unbounded sequence of increasing positive real numbers. Let $C_m(\mathbf{S})$ be the space of all $f \in C(\mathbf{S})$ such that

$$\|f\| := \sum_{j=1}^{\infty} \frac{\|f\|_j}{m_j c^j} < \infty.$$

Then $(C_m(\mathbf{S}), \|\cdot\|)$ is a Banach space ([Matkowski and Nowak, 2011](#)).

Assumption 5.3. Γ is continuous, nonempty, and compact valued, β is bounded, u is continuous, and Q is Feller.

Assumption 5.4. $\Gamma(x, z) \subset K_j$ for all $x \in K_j$, all $z \in \mathbf{Z}$, and all $j \in \mathbb{N}$, and $r(L_\beta) < 1$ where L_β is defined in (10).

Proposition 5.2. *Let $\mathcal{V} = C_m(\mathbf{S})$ and $\mathcal{D} = (\mathbf{X}, \mathbf{Z}, \mathbf{A}, \Gamma, H, \mathcal{V})$ be the associated dynamic program. Under Assumptions 5.3 and 5.4, the lifetime value $v_\sigma(x_0, z_0)$ is well defined and finite for any initial state (x_0, z_0) and feasible policy σ , there exists an increasing unbounded sequence $\{m_j\}$ such that the value function v^* is the unique fixed point of T on \mathcal{V} , $T^n v \rightarrow v^*$ for all $v \in \mathcal{V}$, there exists an optimal stationary policy, and the principle of optimality holds.*

Example 5.2. Consider Example 1.1¹⁴ with production function $f(k, z) = zf(k)$, discount function $\beta(z)$, and utility u any (unbounded) continuous function. Let $\mathbf{X} = \mathbf{A} = \mathbb{R}_+$ and \mathbf{Z} be a compact subset of \mathbb{R}_+ . Suppose f is positive, strictly concave in k , and $\lim_{k \rightarrow \infty} f'(k) = 0$. Then we can find a sequence¹⁵ $\{K_j\}$ of strictly increasing and compact sets covering \mathbf{X} such that $\Gamma(k, z) \subset K_j$ for all $x \in K_j$. Furthermore, if β satisfies $r(L_\beta) < 1$, then Proposition 5.2 can be applied.

¹⁴Also see [Matkowski and Nowak \(2011\)](#) for a similar example with constant β .

¹⁵For example, $K_j = [0, M + j]$ for some large M .

6. CONCLUSION

We provide a simple spectral radius condition under which standard infinite horizon dynamic programs with state-dependent discount rates are well defined and well behaved. In particular, under the stated condition, the value function satisfies the Bellman equation, an optimal policy exists and Bellman's principle of optimality is valid. Thus, the spectral radius condition we state is the natural analog of the condition $\beta < 1$ in the traditional setting, where the discount factor β is not state-dependent. Several applications were provided, including those with both bounded and unbounded rewards.

APPENDIX A. REMAINING PROOFS

In what follows, we consider the dynamic program described in Section 1.1. We will at times make use of the following lemma.

Lemma A.1. *The function β yields $r(L_\beta) < 1$ if and only if (11) holds.*

Proof. Let $\mathbb{1} \equiv 1$ on \mathbf{Z} . For each $z \in \mathbf{Z}$ and $n \in \mathbb{N}$, an inductive argument gives

$$\mathbb{E}_z \prod_{t=0}^{n-1} \beta(Z_t) = L_\beta^n \mathbb{1}(z). \quad (25)$$

Thus, condition (11) can be written as $\|L_\beta^n \mathbb{1}\| < 1$ for some $n \in \mathbb{N}$. Applying Theorem 9.1 of [Krasnosel'skii et al. \(1972\)](#), since (i) L_β is a positive linear operator on $bm\mathbf{Z}$, (ii) the positive cone in this set is solid and normal under the pointwise partial order¹⁶, and (iii) $\mathbb{1}$ lies interior to the positive cone in $bm\mathbf{Z}$, we have

$$r(L_\beta) = \lim_{n \rightarrow \infty} \|L_\beta^n \mathbb{1}\|^{1/n} = \lim_{n \rightarrow \infty} \left\{ \sup_{z \in \mathbf{Z}} \mathbb{E}_z \prod_{t=0}^{n-1} \beta(Z_t) \right\}^{1/n}, \quad (26)$$

where the second equality is due to (25), nonnegativity of β and the definition of the supremum norm. It follows immediately that $r(L_\beta) < 1$ implies the condition in (11).

¹⁶A cone is solid if it has an interior point; it is normal if $0 \leq x \leq y$ implies that $\|x\| \leq M\|y\|$. The cone of nonnegative functions in $bm\mathbf{Z}$ is both solid and normal.

To see that the converse is true, suppose there exists $n \in \mathbb{N}$ such that (11) holds. Then it follows from the Markov property that

$$\begin{aligned}
r(L_\beta) &= \lim_{m \rightarrow \infty} \left\{ \sup_{z \in \mathcal{Z}} \mathbb{E}_z \prod_{t=0}^{m-1} \beta(Z_t) \right\}^{1/m} \\
&= \lim_{m \rightarrow \infty} \left\{ \sup_{z \in \mathcal{Z}} \mathbb{E}_z \prod_{t=0}^{n-1} \beta(Z_t) \left[\mathbb{E}_{Z_{n-1}} \prod_{t=n}^{m-1} \beta(Z_t) \right] \right\}^{1/m} \\
&\leq \lim_{m \rightarrow \infty} \left\{ \sup_{z \in \mathcal{Z}} \mathbb{E}_z \prod_{t=0}^{n-1} \beta(Z_t) \sup_{z \in \mathcal{Z}} \mathbb{E}_z \prod_{t=0}^{m-n-1} \beta(Z_t) \right\}^{1/m} \\
&= \left\{ \sup_{z \in \mathcal{Z}} \mathbb{E}_z \prod_{t=0}^{n-1} \beta(Z_t) \right\}^{1/n} < 1. \quad \square
\end{aligned}$$

A.1. Proofs for Section 2.

A.1.1. *General Results.* We begin with a set of dynamic programming results in an abstract setting that does not impose regularity. Recall from Section 1.1 that \mathcal{V} is a subset of $bm\mathcal{S}$.

Assumption A.1. \mathcal{V} is a complete metric space under the sup metric and the operator T_σ is a self map on \mathcal{V} for all $\sigma \in \Sigma$. There exists a closed subset $\hat{\mathcal{V}} \subset \mathcal{V}$ equipped with the sup metric on which T is a self map.

Theorem A.2. *Under Assumptions 1.2 and A.1, there exists $n \in \mathbb{N}$ such that T_σ and T are n -step contraction mappings on \mathcal{V} and $\hat{\mathcal{V}}$, respectively.*

Proof. In view of T_σ defined in (5), Assumption 1.2 implies that for any $v, w \in \mathcal{V}$ and any $n \in \mathbb{Z}$

$$\begin{aligned}
& |(T_\sigma^n v)(x_0, z_0) - (T_\sigma^n w)(x_0, z_0)| \\
&= |H(x_0, z_0, \sigma(x_0, z_0), T_\sigma^{n-1} v) - H(x_0, z_0, \sigma(x_0, z_0), T_\sigma^{n-1} w)| \\
&\leq \beta(z_0) \int_{\mathcal{Z}} \sup_{x_1 \in \mathcal{X}} |(T_\sigma^{n-1} v)(x_1, z_1) - (T_\sigma^{n-1} w)(x_1, z_1)| Q(z_0, dz_1).
\end{aligned}$$

Iterating on the above inequality gives

$$\begin{aligned}
& |(T_\sigma^n v)(x_0, z_0) - (T_\sigma^n w)(x_0, z_0)| \\
&\leq \beta(z_0) \int_{\mathcal{Z}} \left[\beta(z_1) \int_{\mathcal{Z}} \sup_{x_2 \in \mathcal{X}} |(T_\sigma^{n-2} v)(x_2, z_2) - (T_\sigma^{n-2} w)(x_2, z_2)| Q(z_1, dz_2) \right] Q(z_0, dz_1).
\end{aligned}$$

It follows from an inductive argument that

$$\begin{aligned} |(T_\sigma^n v)(x_0, z_0) - (T_\sigma^n w)(x_0, z_0)| &\leq \mathbb{E}_{z_0} \prod_{t=0}^{n-1} \beta(z_t) \sup_{x_n \in \mathbf{X}} |v(x_n, z_n) - w(x_n, z_n)| \\ &\leq \mathbb{E}_{z_0} \prod_{t=0}^{n-1} \beta(z_t) \|v - w\|. \end{aligned}$$

Taking the supremum gives

$$\|T_\sigma^n v - T_\sigma^n w\| \leq \sup_{z_0 \in \mathbf{Z}} \mathbb{E}_{z_0} \prod_{t=0}^{n-1} \beta(z_t) \|v - w\|.$$

Since Lemma A.1 implies that (11) holds for some n , T_σ is an n -step contraction on \mathcal{V} .

As to the Bellman operator T defined in (6), we have

$$\begin{aligned} |(Tv)(x_0, z_0) - (Tw)(x_0, z_0)| &= \left| \sup_{a \in \Gamma(x_0, z_0)} H(x_0, z_0, a, v) - \sup_{a \in \Gamma(x_0, z_0)} H(x_0, z_0, a, w) \right| \\ &\leq \sup_{a \in \Gamma(x_0, z_0)} |H(x_0, z_0, a, v) - H(x_0, z_0, a, w)|, \end{aligned}$$

for all $v, w \in \hat{\mathcal{V}}$. It then follows from Assumption 1.2 that

$$\begin{aligned} |(T^n v)(x_0, z_0) - (T^n w)(x_0, z_0)| &\leq \sup_{a \in \Gamma(x_0, z_0)} |H(x_0, z_0, a, T^{n-1} v) - H(x_0, z_0, a, T^{n-1} w)| \\ &\leq \beta(z_0) \int_{\mathbf{Z}} \sup_{x_1 \in \mathbf{X}} |(T^{n-1} v)(x_1, z_1) - (T^{n-1} w)(x_1, z_1)| Q(z_0, dz_1). \end{aligned}$$

Iterating on the above inequality gives

$$\begin{aligned} |(T^n v)(x_0, z_0) - (T^n w)(x_0, z_0)| &\leq \beta(z_0) \int_{\mathbf{Z}} \left[\beta(z_1) \int_{\mathbf{Z}} \sup_{x_2 \in \mathbf{X}} |(T^{n-2} v)(x_2, z_2) - (T^{n-2} w)(x_2, z_2)| Q(z_1, dz_2) \right] Q(z_0, dz_1). \end{aligned}$$

A similar inductive argument shows that

$$\|T^n v - T^n w\| \leq \sup_{z_0 \in \mathbf{Z}} \mathbb{E}_{z_0} \prod_{t=0}^{n-1} \beta(z_t) \|v - w\|.$$

In light of Lemma A.1, T is an n -step contraction on $\hat{\mathcal{V}}$. □

We have the following immediate corollary.

Corollary A.3. *If $v_0 \in \mathcal{V}$, the σ -value function v_σ is the unique fixed point of T_σ in \mathcal{V} . The Bellman operator T has a unique fixed point \bar{v} in $\hat{\mathcal{V}}$. Moreover, $T_\sigma^n v \rightarrow v_\sigma$ for all $v \in \mathcal{V}$ and $T^n w \rightarrow \bar{v}$ for all $w \in \hat{\mathcal{V}}$.*

Proof. By Assumption A.1 and a generalized Contraction Mapping Theorem (see, e.g., Cheney, 2013, Section 4.2), T_σ and T are globally stable on \mathcal{V} and $\hat{\mathcal{V}}$ respectively. Hence, if $v_0 \in \mathcal{V}$, v_σ is the unique fixed point of T_σ in \mathcal{V} by definition and $T_\sigma^n v \rightarrow v_\sigma$ for all $v \in \mathcal{V}$. Similarly, T has a unique fixed point $\bar{v} \in \hat{\mathcal{V}}$ and $T^n v \rightarrow \bar{v}$ for all $v \in \hat{\mathcal{V}}$. \square

Corollary A.3 establishes the global stability of T and T_σ . Moreover, it shows that the unique fixed point of T_σ is the σ -value function, the lifetime utility to be optimized. They are both crucial in deriving our next optimality result.

Theorem A.4. *Under Assumptions 1.1, 1.2, and A.1, if $v_0 \in \mathcal{V}$ and there exists some $\sigma \in \Sigma$ such that $T_\sigma \bar{v} = T\bar{v}$, then $\bar{v} = v^*$ and σ is the optimal policy.*

Proof. First note that $\bar{v}(x, z) = (T\bar{v})(x, z) \geq (T_\sigma \bar{v})(x, z)$ for all $x \in \mathbf{X}$, $z \in \mathbf{Z}$, and $\sigma \in \Sigma$ by definition. Iterating T_σ on both sides and using Assumption 1.1, we have

$$\bar{v}(x, z) \geq T_\sigma^n \bar{v}(x, z), \quad \forall x \in \mathbf{X}, z \in \mathbf{Z}$$

for all $n \in \mathbb{Z}$. It follows from corollary A.3 that $\bar{v}(x, z) \geq v_\sigma(x, z)$ for all $x \in \mathbf{X}$, $z \in \mathbf{Z}$, and $\sigma \in \Sigma$. Taking the supremum over Σ gives $\bar{v} \geq v^*$.

For the other direction, since there exists $\sigma \in \Sigma$ such that $T_\sigma \bar{v} = T\bar{v}$, we have $T_\sigma \bar{v} = \bar{v}$. Because $\bar{v} \in \hat{\mathcal{V}} \subset \mathcal{V}$ and T_σ has a unique fixed point in \mathcal{V} , $\bar{v} = v_\sigma$. By the definition of v^* , we have $v^* \geq v_\sigma = \bar{v}$. Therefore, $v_\sigma = v^* = \bar{v}$. By definition, σ is the optimal policy. \square

A.1.2. *Regular Dynamic Programs.* Next we derive some nice properties for regular dynamic programs so that the assumptions of Theorem A.4 can be easily verified.

Proposition A.5. *If \mathcal{D} is a regular dynamic program (see Section 1.3) and Assumption 1.2 holds, then Assumption A.1 holds with $\hat{\mathcal{V}} = bc\mathbf{S}$ and for all $v \in \hat{\mathcal{V}}$ there exists $\sigma \in \Sigma$ such that $T_\sigma v = Tv$.*

Proof of Proposition A.5. Since \mathbf{S} is a metric space, $\mathcal{V} = bm\mathbf{S}$ and $\hat{\mathcal{V}} = bc\mathbf{S}$ are complete metric spaces under the sup metric.

By definition, for all $\sigma \in \Sigma$ and all $v \in \mathcal{V}$ we have

$$T_\sigma v(x, z) = H(x, z, \sigma(x, z), v).$$

The regularity of the dynamic program together with the measurability of all $\sigma \in \Sigma$ implies that $T_\sigma v$ is measurable. Since there exists $v' \in bc\mathbf{S}$ such that $(x, z, a) \mapsto$

$H(x, z, a, x)$ is bounded, we have

$$\begin{aligned} |H(x, z, \sigma(x, z), v)| &\leq |H(x, z, \sigma(x, z), v) - H(x, z, \sigma(x, z), v')| + |H(x, z, \sigma(x, z), v')| \\ &\leq \beta(z)\|v - v'\| + |H(x, z, \sigma(x, z), v')| < \infty \end{aligned}$$

where the last two inequalities follow from assumption 1.2. Hence, T_σ is a self map on \mathcal{V} for all $\sigma \in \Sigma$.

To show that T is a self map on $\hat{\mathcal{V}}$, we use the maximum theorem. Since Γ is continuous, nonempty, and compact valued, and $(x, z, a) \mapsto H(x, z, a, v)$ is continuous for all $v \in \hat{\mathcal{V}}$, Tv defined by

$$Tv(x, z) = \sup_{a \in \Gamma(x, z)} H(x, z, a, v)$$

is continuous. The boundedness of Tv can be derived similarly as above.

Finally, to show the last part, we need the maximizer correspondence to admit a measurable selection. The fact that $(x, z, a) \mapsto H(x, z, a, v)$ is continuous for all $v \in \hat{\mathcal{V}}$, \mathcal{A} is a separable metric space, and Γ is continuous, nonempty, and compact valued ensures that we can apply the measurable maximum theorem (see, for example, Theorem 18.19 in Aliprantis and Border (2006)). Hence, for all $v \in \hat{\mathcal{V}}$ there exists a measurable maximizer σ such that $T_\sigma v = Tv$. \square

Proof of Theorem 2.1. Since \mathcal{D} is a regular dynamic program and Assumption 1.2 holds, it follows from Proposition A.5 that Assumption A.1 holds with $\hat{\mathcal{V}} = bcS$ and for all $v \in \hat{\mathcal{V}}$ there exists $\sigma \in \Sigma$ such that $T_\sigma v = Tv$. Then Theorem 2.1 follows from applying Theorem A.2, Corollary A.3, and Theorem A.4. \square

A.1.3. Policy Iterations.

Proposition A.6. *Suppose there exists $\hat{\Sigma} \subset \Sigma$ such that for all $\sigma \in \hat{\Sigma}$, we have $v_\sigma \in \hat{\mathcal{V}}$. Suppose for all $v \in \hat{\mathcal{V}}$, there exists $\sigma \in \hat{\Sigma}$ such that $T_\sigma v = Tv$. Let $\{\sigma_k\}_{k=0}^\infty$ be generated by the policy iteration algorithm with $\sigma_0 \in \hat{\Sigma}$. Under Assumptions 1.1, 1.2, and A.1, $v_{\sigma_k} \rightarrow v^*$.*

Proof. The proof is adapted from Bertsekas (2013, Proposition 2.4.1). Since $\sigma_0 \in \hat{\Sigma}$, it follows from the assumptions that all $\sigma_k \in \hat{\Sigma}$ and all $v_{\sigma_k} \in \hat{\mathcal{V}}$. By definition, $T_{\sigma_k} v_{\sigma_{k-1}} = Tv_{\sigma_{k-1}} \geq T_{\sigma_{k-1}} v_{\sigma_{k-1}} = v_{\sigma_{k-1}}$. By Assumption 1.1, applying T_{σ_k} to both sides repeatedly gives $T_{\sigma_k}^n v_{\sigma_{k-1}} \geq Tv_{\sigma_{k-1}} \geq v_{\sigma_{k-1}}$. Taking n to infinity, it follows from Corollary A.3 that $v_{\sigma_k} \geq Tv_{\sigma_{k-1}} \geq v_{\sigma_{k-1}}$. An inductive argument implies that $v^* \geq v_{\sigma_k} \geq T^k v_{\sigma_0}$. Taking k to infinity, Corollary A.3 then implies that $v_{\sigma_k} \rightarrow v^*$. \square

The purpose of the assumptions for Proposition A.6 is to ensure that the policies generated by the algorithm are well defined. We can impose further convexity condition on Γ and concavity condition on H to make the assumptions hold. See Stokey et al. (1989, Chapter 9) for examples of such conditions in more specific models.

Proof of Theorem 2.2. Since \mathcal{D} is regular and satisfies Assumptions 1.1 and 1.2, it follows from Proposition A.5 that Assumption A.1 holds. Then the rest follows from applying Proposition A.6. \square

A.2. Proofs for Section 3.

Proof of Proposition 3.1. Suppose first that $r(L_\beta) < 1$. Then, since $r_t \leq b$, we have

$$\mathbb{E} \sum_{t \geq 0} \prod_{i=0}^{t-1} \beta_i r_t \leq b \sum_{t \geq 0} \mathbb{E} \prod_{i=0}^{t-1} \beta_i \leq b \sum_{t \geq 0} \sup_z \mathbb{E}_z \prod_{i=0}^{t-1} \beta_i.$$

Let $m_t := \sup_z \mathbb{E}_z \prod_{i=0}^{t-1} \beta_i$. By Cauchy's root convergence criterion, the sum $\sum_{t \geq 0} m_t$ will be finite whenever $\limsup_{n \rightarrow \infty} m_n^{1/n} < 1$. This does in fact hold when $r(L_\beta) < 1$, in view of (26).

Now suppose instead that $r(L_\beta) \geq 1$. By compactness of L_β , positivity of the function β from Assumption 1.2 and the Krein–Rutman Theorem (see, e.g., Theorem 1.2 in Du (2006)), there exists an everywhere positive function $e \in bcZ$ such that $L_\beta e = r(L_\beta)e$. Choosing $\gamma > 0$ such that $\gamma e \leq \mathbb{1}$ holds pointwise on Z , we have

$$\mathbb{E} \sum_{t \geq 0} \prod_{i=0}^{t-1} \beta_i r_t \geq a \sum_{t \geq 0} L_\beta^t \mathbb{1}(z_0) \geq a\gamma \sum_{t \geq 0} L_\beta^t e(z_0) \geq a\gamma \sum_{t \geq 0} r(L_\beta)^t e(z_0).$$

(Here z_0 is the initial condition of the state process $\{Z_t\}$.) Since e is everywhere positive and $r(L_\beta) \geq 1$, the sum diverges to infinity. \square

A.3. Proofs for Section 4.

We first prove a useful lemma.

Lemma A.7. *If $r(L_\beta) < 1$ where L_β is defined in (10), then we have*

$$\sum_{t=1}^{\infty} \sup_{z \in Z} \mathbb{E}_z \left(\prod_{i=0}^{t-1} \beta(Z_i) \right) < \infty.$$

Proof. Since $r(L_\beta) < 1$, Lemma A.1 yields an $N \in \mathbb{N}$ such that $\sup_{z \in Z} \mathbb{E}_z \prod_{t=0}^{n-1} \beta(Z_t) < 1$ for all $n \geq N$. Let

$$\alpha := \sup_{z \in Z} \mathbb{E}_z \prod_{t=0}^{N-1} \beta(Z_t) < 1 \quad \text{and} \quad M := \sup_{z \in Z, i \leq N} \mathbb{E}_z \prod_{t=0}^{i-1} \beta(Z_t) < \infty.$$

Then we have, for any n ,

$$\begin{aligned} \sup_{z \in \mathbf{Z}} \mathbb{E}_z \prod_{t=0}^{n-1} \beta(Z_t) &= \sup_{z \in \mathbf{Z}} \mathbb{E}_z \left(\prod_{t=0}^{N-1} \beta(Z_t) \mathbb{E}_{Z_{N-1}} \prod_{i=N}^{n-1} \beta(Z_i) \right) \\ &\leq \alpha \sup_{z \in \mathbf{Z}} \mathbb{E}_z \mathbb{E}_{Z_{N-1}} \prod_{i=N}^{n-1} \beta(Z_i) \\ &\leq \alpha \sup_{z \in \mathbf{Z}} \mathbb{E}_z \prod_{t=0}^{n-N-1} \beta(Z_t) \leq \alpha^m M \end{aligned}$$

where $n = mN + i$ for some $m \in \mathbf{Z}$ and $i < N$. Therefore,

$$\sum_{t=1}^{\infty} \sup_{z \in \mathbf{Z}} \mathbb{E}_z \left(\prod_{i=0}^{t-1} \beta(Z_i) \right) \leq NM + \alpha NM + \dots = \frac{NM}{1 - \alpha} < \infty. \quad \square$$

Proof of Proposition 4.1. For any $v_1, v_2 \in bc\mathbf{S}$, we have

$$\begin{aligned} |H(x, z, a, v_1) - H(x, z, a, v_2)| &= (1 - a)\beta(z) \left| \int_{\mathbf{Z}} (v_1(z') - v_2(z')) Q(z, dz') \right| \\ &\leq \beta(z) \int_{\mathbf{Z}} |v_1(z') - v_2(z')| Q(z, dz'). \end{aligned}$$

Then Assumption 1.2 is satisfied. It is apparent that Assumption 1.1 also holds and the feasible correspondence $\Gamma \equiv \{0, 1\}$ is continuous, nonempty, and compact valued. It follows from Lemma A.7 that $K(z)$ is well defined and finite for all $z \in \mathbf{Z}$. Since w and β are bounded and continuous and Q is Feller, \mathcal{D} is regular if we can show that K is continuous.

Since Q is Feller, $S_N(z) := \sum_{t=1}^N \mathbb{E}_z \prod_{i=0}^{t-1} \beta(Z_i)$ is bounded and continuous for all $N \in \mathbb{N}$. Since S_N is nonnegative, it follows from Tonelli's theorem that $\lim_{N \rightarrow \infty} S_N$ is continuous. Therefore, K is continuous and \mathcal{D} is a regular dynamic program. The proposition then follows from Theorem 2.1 and 2.2. \square

Proof of Proposition 4.2. It is obvious that Assumptions 1.1 and 1.2 hold. Next we check that \mathcal{D} is regular so that we can apply Theorem 2.1. Since R , T , and P are bounded and continuous, Γ is continuous, nonempty, and compact valued. Since β and $u(F(\cdot))$ is bounded and continuous, and Q has the Feller property, (c) of the regularity requirements is also satisfied (Stokey et al., 1989, Lemma 12.14). The rest follows from Theorem 2.1 and 2.2. \square

Proof of Proposition 4.3. Consider the function $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ defined by $f(x) = (a + x^{\frac{1}{\theta}})^{\theta}$. Since $\theta \in (0, 1)$, f is strictly increasing and strictly convex and we have

$$|f(x) - f(y)| \leq |x - y|, \quad \forall x, y \in \mathbb{R}_+.$$

Hence, H satisfies that

$$\begin{aligned} |H(x, z, a, v) - H(x, z, a, w)| &\leq \left| \beta(z)^{\theta} \int_{\mathbf{Z}} [v(a, z') - w(a, z')] Q(z, dz') \right| \\ &\leq \beta(z)^{\theta} \int_{\mathbf{Z}} |v(a, z') - w(a, z')| Q(z, dz') \\ &\leq \beta(z)^{\theta} \int_{\mathbf{Z}} \sup_{x' \in \mathbf{X}} |v(x', z') - w(x', z')| Q(z, dz'). \end{aligned}$$

Then Assumption 1.2 hold when $r(L_{\theta}) < 1$.

Since d and p are bounded and continuous, Γ is continuous, nonempty, and compact valued. The space \mathbf{X} and \mathbf{A} are compact due to Γ , so F is bounded and continuous. Since Q is Feller and β is bounded and continuous, the dynamic program is regular. Then the rest follows from Theorems 2.1 and 2.2. \square

A.4. Proofs for Section 5.

Proof of Proposition 5.1. We first show that T is eventually contracting on $\mathcal{V} = H(\mathbf{S}; \theta)$. By Assumption 5.1, T maps \mathcal{V} to itself. Note that for any $v \in \mathcal{V}$, we have $v(x, z) = \|x\|^{\theta} v(x/\|x\|, z)$. It follows from Assumption 5.2 that for any $v, w \in \mathcal{V}$,

$$\begin{aligned} & |(T^n v)(x_0, z_0) - (T^n w)(x_0, z_0)| \\ &\leq \sup_{x_1 \in \Gamma(x_0, z_0)} \beta(z_0) \int_{\mathbf{Z}} |(T^{n-1} v)(x_1, z_1) - (T^{n-1} w)(x_1, z_1)| Q(z_0, dz_1) \\ &\leq \sup_{x_1 \in \Gamma(x_0, z_0)} \beta(z_0) \int_{\mathbf{Z}} \|x_1\|^{\theta} \left| (T^{n-1} v) \left(\frac{x_1}{\|x_1\|}, z_1 \right) - (T^{n-1} w) \left(\frac{x_1}{\|x_1\|}, z_1 \right) \right| Q(z_0, dz_1) \\ &\leq \sup_{x_1 \in \Gamma(x_0, z_0)} \beta(z_0) \alpha^{\theta}(z_0) \|x_0\|^{\theta} \int_{\mathbf{Z}} \left| (T^{n-1} v) \left(\frac{x_1}{\|x_1\|}, z_1 \right) - (T^{n-1} w) \left(\frac{x_1}{\|x_1\|}, z_1 \right) \right| Q(z_0, dz_1). \end{aligned}$$

An inductive argument gives that

$$\begin{aligned} & |(T^n v)(x_0, z_0) - (T^n w)(x_0, z_0)| \\ &\leq \|x_0\|^{\theta} \sup_{x_1 \in \Gamma(x_0, z_0)} \mathbb{E}_{z_0} \prod_{t=0}^{n-1} \beta(z_t) \alpha^{\theta}(z_t) \left| v \left(\frac{x_n}{\|x_n\|}, z_n \right) - w \left(\frac{x_n}{\|x_n\|}, z_n \right) \right| \\ &\leq \|x_0\|^{\theta} \left(\mathbb{E}_{z_0} \prod_{t=0}^{n-1} \beta(z_t) \alpha^{\theta}(z_t) \right) \|v - w\| \end{aligned}$$

where the norm $\|\cdot\|$ is defined in (24). Therefore, we have

$$\|T^n v - T^n w\| \leq \sup_{z_0 \in \mathcal{Z}} \left(\mathbb{E}_{z_0} \prod_{t=0}^{n-1} \beta(z_t) \alpha^\theta(z_t) \right) \|v - w\|.$$

Since $r(L_\alpha) < 1$, T is eventually contracting on \mathcal{V} by Lemma A.1. Hence, T has a unique fixed point \bar{v} on \mathcal{V} and $T^n v \rightarrow \bar{v}$ for any $v \in \mathcal{V}$. Similarly, T_σ is also eventually contracting on \mathcal{V} for any $\sigma \in \Sigma$ and thus $v_\sigma(x_0, z_0)$ is well defined and finite. Since we can find $\sigma \in \Sigma$ such that $T_\sigma \bar{v} = T\bar{v}$ by the measurable maximum theorem, the optimality results follow from Theorem A.4. \square

Proof of Proposition 5.2. Define

$$u_j(x, z) := \max_{a \in \Gamma(x, z)} |u(x, z, a)| \text{ if } x \in K_j \text{ and } r_j := \sup_{x \in K_j, z \in \mathcal{Z}} u_j(x, z).$$

Since u is continuous and every K_j is compact, $r_j < \infty$ for all j . For any initial state (x_0, z_0) , we can find j such that $x_0 \in K_j$. It follows from Assumption 5.4 that

$$|u(\pi_{t-1}(z^{t-1}), z_t, \pi_t(z^t))| \leq r_j$$

for all $t \in \mathbb{N}$.

Choose any increasing and unbounded $\{m_j\}$ such that $m_j \geq r_j$. Since Q is Feller, Tv is continuous on every K_j for $v \in C_m(\mathbf{S})$. It follows from Remark 1(a) of Matkowski and Nowak (2011) that $T : C_m(\mathbf{S}) \rightarrow C(\mathbf{S})$.

Since $\Gamma(x, z) \subset K_j$ for all $x \in K_j$, we have on K_j

$$\begin{aligned} |(T^n v)(x, z) - (T^n w)(x, z)| &\leq \sup_{a \in \Gamma(x, z)} \beta(z) \int_{\mathcal{Z}} |T^{n-1} v(a, z') - T^{n-1} w(a, z')| Q(z, dz') \\ &\leq \sup_{a \in K_j} \beta(z) \int_{\mathcal{Z}} |T^{n-1} v(a, z') - T^{n-1} w(a, z')| Q(z, dz') \\ &\leq \beta(z) \|T^{n-1} v - T^{n-1} w\|_j. \end{aligned}$$

An inductive argument gives

$$|(T^n v)(x, z) - (T^n w)(x, z)| \leq \mathbb{E}_z \prod_{t=0}^{n-1} \beta(Z_t) \|v - w\|_j.$$

Taking the supremum, we have

$$\|T^n v - T^n w\|_j \leq \sup_{z \in \mathcal{Z}} \mathbb{E}_z \prod_{t=0}^{n-1} \beta(Z_t) \|v - w\|_j.$$

Since $r(L_\beta) < 1$, T is a 0-local contraction¹⁷ by Lemma A.1. Then it follows from Proposition 1 of Matkowski and Nowak (2011) that T has a unique fixed point \bar{v} in $C_m(\mathbf{S})$. It can be proved in the same way that T_σ is also a 0-local contraction and hence v_σ is well defined and finite for any initial state. Since we can find σ such that $T_\sigma \bar{v} = T\bar{v}$ by the measurable maximum theorem, the optimality results follow from a similar argument to the proofs of Theorem A.4. \square

REFERENCES

- ALBUQUERQUE, R., M. EICHENBAUM, V. X. LUO, AND S. REBELO (2016): “Valuation risk and asset pricing,” *The Journal of Finance*, 71, 2861–2904.
- ALBUQUERQUE, R., M. EICHENBAUM, D. PAPANIKOLAOU, AND S. REBELO (2015): “Long-run bulls and bears,” *Journal of Monetary Economics*, 76, S21–S36.
- ALIPRANTIS, C. D. AND K. C. BORDER (2006): *Infinite Dimensional Analysis: A Hitchhiker’s Guide*, Springer.
- ALVAREZ, F. AND N. L. STOKEY (1998): “Dynamic programming with homogeneous functions,” *Journal of Economic Theory*, 82, 167–189.
- BELLMAN, R. (1957): *Dynamic programming*, Academic Press.
- BERAJA, M., E. HURST, AND J. OSPINA (2016): “The aggregate implications of regional business cycles,” Tech. rep., National Bureau of Economic Research.
- BERTSEKAS, D. P. (2013): *Abstract dynamic programming*, Athena Scientific Belmont, MA.
- (2017): *Dynamic programming and optimal control*, vol. 4, Athena Scientific.
- BHANDARI, A., D. EVANS, M. GOLOSOV, AND T. J. SARGENT (2013): “Taxes, debts, and redistributions with aggregate shocks,” Tech. rep., National Bureau of Economic Research.
- BLACKWELL, D. (1965): “Discounted dynamic programming,” *The Annals of Mathematical Statistics*, 36, 226–235.
- CAMPBELL, J. Y. (1986): “Bond and stock returns in a simple exchange model,” *The Quarterly Journal of Economics*, 101, 785–803.
- CAO, D. (2018): “Recursive equilibrium in Krusell and Smith (1998),” *Available at SSRN 2863349*.
- CARMON, Y. AND A. SHWARTZ (2009): “Markov decision processes with exponentially representable discounting,” *Operations Research Letters*, 37, 51–55.
- CHENEY, W. (2013): *Analysis for applied mathematics*, vol. 208, Springer Science & Business Media.

¹⁷We say an operator $T : C_m(\mathbf{S}) \rightarrow C(\mathbf{S})$ is a 0-local contraction if there exists $\beta \in (0, 1)$ such that $\|Tf - Tg\|_j \leq \beta \|f - g\|_j$ for all $f, g \in C_m(\mathbf{S})$ and all $j \in \mathbb{N}$.

- CHRISTIANO, L., M. EICHENBAUM, AND S. REBELO (2011): “When is the government spending multiplier large?” *Journal of Political Economy*, 119, 78–121.
- CHRISTIANO, L. J., R. MOTTO, AND M. ROSTAGNO (2014): “Risk shocks,” *American Economic Review*, 104, 27–65.
- CORREIA, I., E. FARHI, J. P. NICOLINI, AND P. TELES (2013): “Unconventional fiscal policy at the zero bound,” *American Economic Review*, 103, 1172–1211.
- DU, Y. (2006): *Order structure and topological methods in nonlinear partial differential equations: Vol. 1: Maximum principles and applications*, vol. 2, World Scientific.
- EGGERTSSON, G. B. (2011): “What fiscal policy is effective at zero interest rates?” *NBER Macroeconomics Annual*, 25, 59–112.
- EGGERTSSON, G. B. AND M. WOODFORD (2003): “Zero bound on interest rates and optimal monetary policy,” *Brookings papers on economic activity*, 2003, 139–211.
- FAGERENG, A., M. B. HOLM, B. MOLL, AND G. NATVIK (2019): “Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains,” Tech. rep., Princeton.
- GONZÁLEZ-SÁNCHEZ, D., F. LUQUE-VÁSQUEZ, AND J. A. MINJÁREZ-SOSA (2019): “Zero-Sum Markov Games with Random State-Actions-Dependent Discount Factors: Existence of Optimal Strategies,” *Dynamic Games and Applications*, 9, 103–121.
- HALL, R. E. (2017): “High discounts and high unemployment,” *American Economic Review*, 107, 305–30.
- HIGASHI, Y., K. HYOGO, AND N. TAKEOKA (2009): “Subjective random discounting and intertemporal choice,” *Journal of Economic Theory*, 144, 1015–1053.
- HIGASHI, Y., K. HYOGO, N. TAKEOKA, AND H. TANAKA (2017): “Comparative impatience under random discounting,” *Economic Theory*, 63, 621–651.
- HILLS, T., T. NAKATA, AND S. SCHMIDT (2016): “The risky steady state and the interest rate lower bound,” Tech. rep., ECB Working Paper.
- HILLS, T. S. AND T. NAKATA (2018): “Fiscal multipliers at the zero lower bound: the role of policy inertia,” *Journal of Money, Credit and Banking*, 50, 155–172.
- HUBMER, J., P. KRUSELL, AND A. A. SMITH (2018): “A Comprehensive Quantitative Theory of the US Wealth Distribution,” Tech. rep., Yale.
- ILHUICATZI-ROLDÁN, R., H. CRUZ-SUÁREZ, AND S. CHÁVEZ-RODRÍGUEZ (2017): “Markov decision processes with time-varying discount factors and random horizon,” *Kybernetika*, 53, 82–98.
- JAŚKIEWICZ, A. AND A. S. NOWAK (2011): “Discounted dynamic programming with unbounded returns: application to economic models,” *Journal of Mathematical Analysis and Applications*, 378, 450–462.

- JUSTINIANO, A. AND G. E. PRIMICERI (2008): “The time-varying volatility of macroeconomic fluctuations,” *American Economic Review*, 98, 604–41.
- JUSTINIANO, A., G. E. PRIMICERI, AND A. TAMBALOTTI (2010): “Investment shocks and business cycles,” *Journal of Monetary Economics*, 57, 132–145.
- (2011): “Investment shocks and the relative price of investment,” *Review of Economic Dynamics*, 14, 102–121.
- KARNI, E. AND I. ZILCHA (2000): “Saving behavior in stationary equilibrium with random discounting,” *Economic Theory*, 15, 551–564.
- KEHOE, P. J., V. MIDRIGAN, AND E. PASTORINO (2018): “Evolution of modern business cycle models: Accounting for the great recession,” *Journal of Economic Perspectives*, 32, 141–66.
- KOPECKY, K. A. AND R. M. SUEN (2010): “Finite state Markov-chain approximations to highly persistent processes,” *Review of Economic Dynamics*, 13, 701–714.
- KRASNOSEL'SKII, M. A., G. M. VAINIKKO, P. P. ZABREIKO, Y. B. RUTITSKII, AND V. Y. STETSENKO (1972): *Approximate Solution of Operator Equations*, Springer Netherlands.
- KRISHNA, R. V. AND P. SADOWSKI (2014): “Dynamic preference for flexibility,” *Econometrica*, 82, 655–703.
- KRUSELL, P., T. MUKOYAMA, A. ŞAHIN, AND A. A. SMITH (2009): “Revisiting the welfare effects of eliminating business cycles,” *Review of Economic Dynamics*, 12, 393–404.
- KRUSELL, P. AND A. A. SMITH (1998): “Income and wealth heterogeneity in the macroeconomy,” *Journal of Political Economy*, 106, 867–896.
- LUCAS, R. E. AND E. C. PRESCOTT (1974): “Equilibrium search and unemployment,” *Journal of Economic Theory*, 7, 188–209.
- MA, Q., J. STACHURSKI, AND A. A. TODA (2019): “The Income Fluctuation Problem and the Evolution of Wealth,” Tech. rep., arXiv preprint arXiv:1905.13045.
- MARTINS-DA ROCHA, V. F. AND Y. VAILAKIS (2010): “Existence and uniqueness of a fixed point for local contractions,” *Econometrica*, 78, 1127–1141.
- MATKOWSKI, J. AND A. S. NOWAK (2011): “On discounted dynamic programming with unbounded returns,” *Economic Theory*, 46, 455–474.
- MCCALL, J. J. (1970): “Economics of information and job search,” *The Quarterly Journal of Economics*, 113–126.
- MEHRA, R. AND R. SAH (2002): “Mood fluctuations, projection bias, and volatility of equity prices,” *Journal of Economic Dynamics and Control*, 26, 869–887.
- MINJÁREZ-SOSA, J. A. (2015): “Markov control models with unknown random state-action-dependent discount factors,” *TOP*, 23, 743–772.

- MUKOYAMA, T. (2009): “A Note on Cyclical Discount Factors and Labor Market Volatility,” Tech. rep.
- NAKATA, T. AND H. TANAKA (2016): “Equilibrium Yield Curves and the Interest Rate Lower Bound,” Tech. rep., FEDS Working Paper.
- PRIMICERI, G. E., E. SCHAUMBURG, AND A. TAMBALOTTI (2006): “Intertemporal disturbances,” Tech. rep., National Bureau of Economic Research.
- PUTERMAN, M. L. (2014): *Markov Decision Processes.: Discrete Stochastic Dynamic Programming*, John Wiley & Sons.
- RINCÓN-ZAPATERO, J. P. AND C. RODRÍGUEZ-PALMERO (2003): “Existence and uniqueness of solutions to the Bellman equation in the unbounded case,” *Econometrica*, 71, 1519–1555.
- ROBIN, J.-M. (2011): “On the dynamics of unemployment and wage distributions,” *Econometrica*, 79, 1327–1355.
- SAIJO, H. (2017): “The uncertainty multiplier and business cycles,” *Journal of Economic Dynamics and Control*, 78, 1–25.
- SCHÄL, M. (1975): “Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal,” *Probability theory and related fields*, 32, 179–196.
- SCHORFHEIDE, F., D. SONG, AND A. YARON (2018): “Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach,” *Econometrica*, 86, 617–654.
- SHIMER, R. (2005): “The cyclical behavior of equilibrium unemployment and vacancies,” *American Economic Review*, 95, 25–49.
- STOKEY, N. L., R. E. LUCAS, AND E. C. PRESCOTT (1989): *Recursive methods in economic dynamics*, Harvard University Press.
- TODA, A. A. (2018): “Wealth distribution with random discount factors,” *Journal of Monetary Economics*.
- WEI, Q. AND X. GUO (2011): “Markov decision processes with state-dependent discount factors and unbounded rewards/costs,” *Operations Research Letters*, 39, 369–374.
- WILLIAMSON, S. D. (2019): “Low real interest rates and the zero lower bound,” *Review of Economic Dynamics*, 31, 36–62.
- WOODFORD, M. (2011): “Simple Analytics of the Government Expenditure Multiplier,” *American Economic Journal: Macroeconomics*, 3, 1–35.