Existence and uniqueness of recursive utilities without boundedness*

Timothy M. Christensen†

January 14, 2022.

Abstract

This paper derives primitive, easily verifiable sufficient conditions for existence and uniqueness of (stochastic) recursive utilities for several important classes of preferences. In order to accommodate models commonly used in practice, we allow both the state space and per-period utilities to be unbounded. For many of the models we study, existence and uniqueness is established under a single, primitive "thin tail" condition on the distribution of growth in per-period utilities. We present several applications to robust preferences, models of ambiguity aversion and learning about hidden states, and Epstein–Zin preferences.

Keywords: Stochastic recursive utility, ambiguity, model uncertainty, existence, uniqueness.

JEL codes: C62, C65, D81, E7, G10

*This paper is a revised version of the preprint arXiv:1812.11246 which was posted to arXiv in December 2018. I thank J. Borovička, L. Hansen, T. Sargent, and participants of the "Blue Collar Working Group" at the University of Chicago, and several seminar audiences for helpful comments. This material is based upon work supported by the National Science Foundation under Grant No. SES-1919034.

†Department of Economics, New York University, 19 W. 4th Street, 6th floor, New York, NY 10012, USA. E-mail address: `timothy.christensen@nyu.edu`

1 Introduction

Recursive utilities[1] play a central role in contemporary macroeconomics and finance. Under recursive preferences, the value of a stream of per-period utilities is defined as the solution to a nonlinear, stochastic, forward-looking difference equation (or "recursion"). Despite the importance of recursive utilities, existence and uniqueness remains an unresolved issue as the recursions are typically not contraction mappings when state variables and per-period utilities are unbounded. In this paper, we derive primitive, easily verifiable sufficient conditions for existence and uniqueness of recursive utilities in stationary, infinite-horizon Markovian environments, with an emphasis on robust preferences, models of ambiguity aversion and learning about hidden states, and Epstein–Zin preferences. To accommodate models used extensively in macroeconomics and finance, we allow both the support of the Markov state vector and per-period utilities to be unbounded.

There are a large number of existence and uniqueness results for recursive utilities in models with compact state space, and possibly also bounded per-period utilities.[2] However, many models used in macroeconomics and finance feature unbounded (i.e., non-compact) state spaces and unbounded utilities. For instance, the extensive long-run risks literature following Bansal and Yaron (2004) typically models state variables as vector autoregressive processes with unbounded shocks.[3] A seemingly reasonable approach for models with non-compact state space is to truncate (i.e., compactifty) the state space and apply existing results for compact state spaces. After all, this truncation occurs implicitly when computing solutions numerically. However, truncation can materially alter the existence and uniqueness properties of the recursions we study. Knowing when the original model without truncation has a unique solution remains important for reconciling numerical solutions with the original (un-truncated) model envisioned by the researcher.

To illustrate this point, in Section 2 we present two empirically relevant examples to show how non-existence and non-uniqueness can arise under unboundedness. For both examples, we focus on a recursion arising under preferences for "robustness" (Hansen and Sargent, 1995, 2001; Hansen, Sargent, Turmuhambetova, and Williams, 2006) and under Epstein–Zin

[1] Throughout the paper, by "recursive utility" we mean "stochastic recursive utility".

[2] See, e.g., Epstein and Zin (1989), Alvarez and Jermann (2005), Marinacci and Montrucchio (2010), Guo and He (2017), Becker and Rincon-Zapatero (2017), Bloise and Vailakis (2018), Balbus (2020), Borovička and Stachurski (2020), Ren and Stachurski (2020), and references therein.

[3] See, e.g., Hansen, Heaton, and Li (2008), Barillas, Hansen, and Sargent (2009), Wachter (2013), Bansal, Kiku, Shaliastovich, and Yaron (2014), Croce, Lettau, and Ludvigson (2015), Bidder and Smith (2018), Collard, Mukerji, Sheppard, and Tallon (2018), Schorfheide, Song, and Yaron (2018), and additional references listed in Sections 4–6.

preferences with unit intertemporal elasticity of substitution. The first example features a simplified version of the consumption growth process from Schorfheide et al. (2018), for which existence fails. The second example is from Bidder and Smith (2018) and Wachter (2013), for which uniqueness fails. When the state space is truncated, however, the recursion has a unique solution (irrespective of the truncation level) in both examples. This stark difference between the compact and unbounded case arises because the properties of the recursion depend delicately on the tail behavior of state variables and truncation, even at an arbitrarily high truncation level, materially alters tail behavior.

For many of the models we study, the single primitive sufficient condition we require for both existence *and* uniqueness is that the distribution of growth in per-period utilities has thin tails, in a sense we make precise. We verify this condition for robust preferences, models of ambiguity aversion and learning about hidden states, and Epstein–Zin preferences with unit intertemporal elasticity of substitution (IES). We consider both canonical linear-Gaussian environments which are pertinent to the long-run risks literature as well as environments featuring regime-switching and stochastic volatility.

As with much of the literature, we identify recursive utilities with fixed points of a nonlinear operator acting on a suitable function class. One strand of the literature on existence and uniqueness of (deterministic or stochastic) recursive utilities under unboundedness uses contraction mapping arguments for function classes defined via weighted sup-norms.[4] However, it is not always easy to find a suitable weighting function under which operators defining recursive utilities are a contraction.[5] Our arguments instead rely on monotonicity and concavity/convexity properties of the recursions we study, as with earlier work by Marinacci and Montrucchio (2010); see also Becker and Rincon-Zapatero (2017), Bloise and Vailakis (2018), and Ren and Stachurski (2020), primarily for the compact case.[6] While our approach has some similarities with these earlier works, it differs in terms of the function class and technical arguments used so as to accommodate a broad class of empirically relevant models with unbounded state space. In particular, our arguments do not rely on certain topological properties of the space of bounded functions, such as the such the "solidness" of the cone of non-negative functions.

[4]See, e.g., Boyd (1990) and Durán (2003) for deterministic and stochastic utilities, respectively. Le Van and Vailakis (2005) provide a related approach for deterministic utilities under Lipschitz conditions.

[5]See, e.g., Le Van and Vailakis (2005) for a discussion.

[6]Marinacci and Montrucchio (2010) and Becker and Rincon-Zapatero (2017) allow for processes that are bounded with probability one but growing over time using weighted ℓ^∞-norms. See also Ren and Stachurski (2020) for a particular parameterization of Epstein–Zin preferences with unbounded state space using a weighted sup-norm, where the weighting function is tightly related to per-period utilities and the law of motion of the Markov state.

Our point of departure is to embed a transformation of the value function, such as its logarithm, in a class of unbounded but thin-tailed functions. The class is an exponential-Orlicz class used in empirical process theory in statistics (van der Vaart and Wellner, 1996) and modern high-dimensional probability (Vershynin, 2018).[7] Exponential-Orlicz classes are naturally suited to the recursions we study, which involve the composition of exponential and logarithmic transforms and expected values.

The key high-level condition we use to establish uniqueness is that a subgradient (in the convex case) or supergradient (in the concave case) of the recursion is monotone and its spectral radius is strictly less than one. For many of the models we study, the recursion is convex and its subgradient is a discounted conditional expectation under a distorted law of motion. Verifying the spectral radius condition in these models amounts to checking a primitive thin-tail condition on the change-of-measure distorting the law of motion. We specialize this condition to particular models, deriving more primitive thin-tail conditions on the distribution of growth in per-period utility which are easy to verify: one simply has to know the tail behavior of the distribution.

To illustrate the usefulness of our results, we then present applications to three classes of models.

Section 4 studies a recursion arising under preferences for "robustness", namely risk-sensitive preferences (Hansen and Sargent, 1995), multiplier preferences (Hansen and Sargent, 2001), constraint preferences (Hansen et al., 2006), and also under Epstein and Zin (1989) preferences with unit IES. There are currently no uniqueness results in the literature for this recursion allowing non-compact state space and unbounded utilities (see the discussion in Section 4), both of which are important for models in macroeconomics and finance. We establish new existence and uniqueness results under a single primitive thin-tail condition on utility growth. We verify this condition in canonical linear-Gaussian environments and environments featuring regime-switching and stochastic volatility, thereby establishing new existence and uniqueness results for such settings.

Section 5 considers models with learning. We study extensions by Hansen and Sargent (2007, 2010) of multiplier preferences to accommodate both model uncertainty and uncertainty about hidden states, dynamic models of ambiguity aversion studied by Ju and Miao (2012) and Klibanoff, Marinacci, and Mukerji (2009), and Epstein–Zin preferences with unit IES and learning. There are currently no existence and uniqueness results in the literature

[7] Previously, Hindy and Huang (1992) and Hindy, Huang, and Kreps (1992) used Orlicz classes to define topologies for consumption paths in continuous time.

allowing non-compact state space and unbounded utilities (see the discussion in Section 5). We establish existence and uniqueness under a single primitive thin-tail condition on utility growth. We verify the condition, and therefore establish existence and uniqueness results, for regime-switching environments (Ju and Miao, 2012) and Gaussian state-space models (Hansen and Sargent, 2007, 2010; Croce et al., 2015; Collard et al., 2018).

Finally, in Section 6 we examine Epstein–Zin recursive utilities with IES not equal to one. There are no uniqueness results for models with unbounded state space when risk aversion and intertemporal substitution are in a range normally encountered in the long-run risks literature (see the discussion in Section 6). Here we establish existence under an eigenvalue condition from Hansen and Scheinkman (2012) and a thin-tail condition on its corresponding eigenfunction. We verify this condition for linear-Gaussian environments which are pertinent to the long-run risks literature. Appendix A gives definitions of mathematical terms used in the main text. All proofs are in Appendix B.

2 Non-existence and non-uniqueness without boundedness

In this section, we present two empirically relevant examples of non-existence and non-uniqueness in models with unbounded state spaces. The first features a simplified version of the consumption growth process from Schorfheide et al. (2018), for which existence fails. The second is the model from Bidder and Smith (2018) and Wachter (2013), for which uniqueness fails. In both examples, however, there is always a unique solution when the support of state variables are truncated (irrespective of the truncation level).

2.1 Non-existence

Consider the following simplified[8] model of consumption growth $\{g\}_{t\geq 0}$ from Schorfheide et al. (2018):

$$g_{t+1} = \nu_g + e^{h_t}\eta^g_{t+1}\,, \qquad h_{t+1} = \nu_h + \rho h_t + \sigma\eta^h_{t+1}\,, \tag{1}$$

where $|\rho| < 1$, and the η^g_t and η^h_t are all i.i.d. $N(0,1)$. Let $X_t = (g_t, h_t)$. Both g_t and h_t have support $\mathbb{R}$.

[8] We have removed a stochastic growth component from model (4) of Schorfheide et al. (2018) to simplify presentation. Non-uniqueness arises here because of the form of the stochastic volatility process, and not because of the absence of a stochastic growth component.

Suppose we seek a solution v to the recursion

$$v(x) = \beta \log \mathbb{E}^Q \left[e^{v(X_{t+1})+\alpha g_{t+1}} \middle| X_t = x \right], \tag{2}$$

where $\beta \in (0,1)$ and $\alpha \in \mathbb{R}$ are preference parameters and $\mathbb{E}^Q$ denotes expectation under the law of motion (1). This recursion is studied in Section 4 and arises under preferences for robustness as well as under Epstein–Zin preferences with unit IES. As the conditional distribution of X_{t+1} given $X_t = (g,h)$ depends only on h, the right-hand side conditional expectation, and therefore v, must depend only on h. Using (1), we see that recursion (2) simplifies to

$$v(h) = \mathsf{a} + \mathsf{b}e^{2h} + \beta \log \mathbb{E}^Q \left[e^{v(h_{t+1})} \middle| h_t = h \right] =: \mathbb{T}v(h), \tag{3}$$

where $\mathsf{a} = \alpha\beta\nu_g$ and $\mathsf{b} = \frac{1}{2}\alpha^2\beta$.

Let L^1 denote the space of (equivalence classes of) functions f for which $\mathbb{E}^\mu[|f(h_t)|] < \infty$, where $\mathbb{E}^\mu$ denotes expectation under the stationary distribution μ implied by (1) (see Appendix A).

Proposition 2.1. *Let $\alpha \neq 0$ and let consumption growth g evolve according to (1). Then: recursion (3) has no solution in L^1.*

Now suppose instead that the support of h is truncated to some compact interval $\mathcal{H} := [-H, H]$ for $H \in (0,\infty)$. Under this truncation, $\mathbb{T}$ satisfies Blackwell's sufficient conditions for a contraction mapping on the space $B(\mathcal{H})$ of bounded functions on $\mathcal{H}$. Therefore, $\mathbb{T}$ has a unique fixed point in $B(\mathcal{H})$, irrespective of the truncation level H.

To understand the difference between the bounded and unbounded cases, note from (2) that for $\mathbb{T}v$ to be well defined we need the tails of the (conditional) distribution of $v(X_{t+1})+\alpha g_{t+1}$ to decay sufficiently quickly (i.e., sub-exponentially). While this condition is always satisfied in the bounded case, it is violated in model (1) due to the specification of the stochastic volatility process. In Section 4 we present a different form of stochastic volatility with thinner tails for which existence and uniqueness can be guaranteed without truncation.

2.2 Non-uniqueness

Consider the model from Bidder and Smith (2018) (see also Wachter (2013)) in which consumption growth g evolves as

$$g_{t+1} = \nu_g + w_{z,t+1} + \sigma w_{g,t+1} \,, \tag{4}$$

where $w_{g,t+1} \sim N(0,1)$ and $w_{z,t+1}|j_{t+1} \sim N(\nu_j j_{t+1}, \sigma_j^2 j_{t+1})$ with $\nu_j < 0$, and where $j_{t+1}|h_t$ is Poisson-distributed with mean h_t, where $\{h\}_{t\ge 0}$ follows an autoregressive gamma process with parameters (φ, c, δ) (see appendix H of Backus, Chernov, and Zin (2014) and references therein for details). Here consumption growth is subject to occasional "disasters" which arrive at rate h_t. We again seek a solution to recursion (2) with $X_t = (g_t, h_t)$. The support of g_t is $\mathbb{R}$ and the support of h_t is $\mathbb{R}_+$. As with the previous example, here it suffices to consider solutions depending only on h. Using (4), we may rewrite recursion (2) as

$$v(h) = \mathsf{a} + \mathsf{b}h + \beta \log \mathbb{E}^Q[e^{v(h_{t+1})}|h_t = h] =: \mathbb{T}v(h) \,, \tag{5}$$

where $\mathsf{a} = \alpha\beta\nu_g + \frac{1}{2}\alpha^2\beta\sigma^2$ and $\mathsf{b} = \beta(e^{\alpha\nu_j + \frac{1}{2}\alpha^2\sigma_j^2} - 1)$. Let $\mathsf{q} = 1 + c\mathsf{b} - \beta\varphi$.

Proposition 2.2. *Let consumption growth g evolve according to (4) and let $\mathsf{q}^2 - 4c\mathsf{b} > 0$. Then: recursion (5) has two solutions of the form $v_i(h) = a_i + b_i h$, $i = 1, 2$, where*

$$b_1 = \frac{\mathsf{q} - \sqrt{\mathsf{q}^2 - 4c\mathsf{b}}}{2c} \,, \qquad b_2 = \frac{\mathsf{q} + \sqrt{\mathsf{q}^2 - 4c\mathsf{b}}}{2c} \,,$$

and $a_i = \frac{\mathsf{a} - \beta\delta\log(1 - b_i c)}{1-\beta}$, $i = 1, 2$.

Note that the condition $\mathsf{q}^2 - 4c\mathsf{b} > 0$ is satisfied for the parameterization in Bidder and Smith (2018), so uniqueness fails for that parameterization.

One may again verify that $\mathbb{T}$ is a contraction mapping on the space $B(\mathcal{H})$ of bounded functions on $\mathcal{H} := [0, H]$ when the support of h is truncated to $[0, H]$ with $H \in (0, \infty)$. Therefore, $\mathbb{T}$ has a unique fixed point in $B(\mathcal{H})$, irrespective of the truncation level H.

In this example, the stability properties of fixed points also differ under truncation and unboundedness. Under truncation, the recursion is a (global) contraction on $B(\mathcal{H})$ so the unique fixed point is globally attracting. In the unbounded case, suppose we restrict $\mathbb{T}$ to affine functions of the form $v(h) = a + bh$. Here the two solutions (a_i, b_i), $i = 1, 2$, solve the recursion $(a, b) = T(a, b)$ (see the proof of Proposition 2.2 for a derivation), where

$$T(a,b) = \left(\mathsf{a} + \beta a - \beta\delta \log(1-bc)\,,\; \mathsf{b} + \frac{\beta\varphi b}{1-bc}\right).$$

Fixed point iteration of T on an initial point (a_0, b_0) converges to (a_1, b_1) if $b_0 < b_2$, converges to (a_2, b_2) if $b_0 = b_2$, and diverges otherwise. In the latter case, iterations diverge because the tails of $a_0 + b_0 h_{t+1} + \alpha g_{t+1}$ become increasingly heavy under repeated application of $\mathbb{T}$, eventually becoming sufficiently heavy that $\mathbb{T}v$ is no longer finite.

3 Preliminaries

Section 3.1 presents a basic existence and uniqueness result which serves as a useful starting point for organizing the discussion that follows. The key condition for uniqueness is a spectral radius condition on a sub- or supergradient of the operator. In many models with forward-looking agents—including models we study in the later sections—the subgradient is a discounted conditional expectation under a distorted law of motion. We then show in Section 3.3 that the spectral radius condition holds in these models under a "thin tail" condition on the change-of-measure distorting the law of motion. We shall use this result to derive more primitive conditions for recursive utilities in Sections 4 and 5.

3.1 A basic fixed-point result

In this section, we present a basic existence and uniqueness result for an operator $\mathbb{T}$ acting on a Banach lattice $\mathcal{E}$ with norm $\|\cdot\|$ and partial order $\leq$ (see Appendix A). We also require $\mathcal{E}$ has a *monotone convergence property*: any increasing sequence $\{f_n\}_{n\geq 1} \subset \mathcal{E}$ bounded above by some $g \in \mathcal{E}$ converges to some $f \leq g$. Spaces with this property include L^p spaces for $1 \leq p < \infty$ and Orlicz spaces (see Section 3.2). We say that $\mathbb{T}$ is *monotone* if $\mathbb{T}f \leq \mathbb{T}g$ whenever $f \leq g$. A bounded linear operator $\mathbb{D}_f$ on $\mathcal{E}$ is a *subgradient* of $\mathbb{T}$ at f if

$$\mathbb{T}g - \mathbb{T}f \geq \mathbb{D}_f(g - f) \tag{6}$$

for each $g \in \mathcal{E}$, and a *supergradient* of $\mathbb{T}$ at f if inequality (6) is reversed:

$$\mathbb{T}g - \mathbb{T}f \leq \mathbb{D}_f(g - f) \tag{7}$$

for each $g \in \mathcal{E}$. Let $\|\mathbb{D}\| := \sup\{\|\mathbb{D}f\| : f \in \mathcal{E}, \|f\| = 1\}$ denote the norm of a linear operator $\mathbb{D}$ on $\mathcal{E}$, and $\rho(\mathbb{D};\mathcal{E}) := \lim_{n\to\infty} \|\mathbb{D}^n\|^{1/n}$ denote the spectral radius of $\mathbb{D}$, where $\mathbb{D}^n$ denotes $\mathbb{D}$ applied n times in succession.

Proposition 3.1. *(i) Existence: Let $\mathbb{T}$ be a continuous and monotone operator on $\mathcal{E}$ and let there exist $\underline{v}, \bar{v} \in \mathcal{E}$ such that either (a) $\mathbb{T}\bar{v} \leq \bar{v}$ and $\{\mathbb{T}^n\bar{v}\}_{n\geq 1}$ is bounded from below by $\underline{v}$, or (b) $\mathbb{T}\underline{v} \geq \underline{v}$ and $\{\mathbb{T}^n\underline{v}\}_{n\geq 1}$ is bounded from above by $\bar{v}$. Then: $\mathbb{T}^n\bar{v}$ (if (a) holds) or $\mathbb{T}^n\underline{v}$ (if (b) holds) converges to a fixed point $v \in \mathcal{E}$ as $n \to \infty$, where $\underline{v} \leq v \leq \bar{v}$.*
(ii) Uniqueness: Suppose that inequality (6) holds at each fixed point $v \in \mathcal{E}$, or inequality (7) holds at each fixed point $v \in \mathcal{E}$, and $\mathbb{D}_v$ is monotone with $\rho(\mathbb{D}_v;\mathcal{E}) < 1$ for each fixed point $v \in \mathcal{E}$. Then: $\mathbb{T}$ has at most one fixed point in $\mathcal{E}$.

When uniqueness cannot be guaranteed, ordering and stability criteria can be used to refine the set of fixed points. Let $\mathcal{V}$ denote the set of fixed points of $\mathbb{T}$. Say v is the *smallest* (respectively *largest*) fixed point of $\mathbb{T}$ if $v \leq v'$ (resp. $v \geq v'$) holds for each $v' \in \mathcal{V}$. Say v is *stable* if $\rho(\mathbb{D}_v;\mathcal{E}) < 1$ (see, e.g., Amann (1976)).

Corollary 3.1. *Let the conditions of Proposition 3.1(i) hold, let $\mathbb{T}$ satisfy (6) (resp. (7)) at each of its fixed points, and let $v \in \mathcal{E}$ be a fixed point of $\mathbb{T}$ with $\rho(\mathbb{D}_v;\mathcal{E}) < 1$. Then: v is both the smallest (resp. largest) fixed point and the unique stable fixed point of $\mathbb{T}$ in $\mathcal{E}$.*

Stability of v is a useful property. In the examples we consider in Sections 4 and 5, the subgradient is of the form $\mathbb{D}_v = \beta\tilde{\mathbb{E}}$ with $\beta \in (0,1)$, where $\tilde{\mathbb{E}}$ denotes conditional expectation under a distorted probability measure. Stability ensures that discounted expected utilities under $\tilde{\mathbb{E}}$ are finite. Stability of v also helps ensure that fixed-point iteration on a neighborhood of v will converge to v (see Lemma B.4).

While Proposition 3.1(i) establishes that $\mathbb{T}^n\bar{v}$ (if (a) holds) or $\mathbb{T}^n\underline{v}$ (if (b) holds) converges to a fixed point of $\mathbb{T}$ as $n \to \infty$, it is also possible to strengthen this to a (partial) global convergence result.

Corollary 3.2. *Let the conditions of Proposition 3.1 hold, with the additional restriction that $\mathbb{T}$ satisfies (6) if (a) holds, or (7) if (b) holds, at v. Then: for any $w \in \mathcal{E}$ for which $w \leq \bar{v}$ (if (a) holds) or $w \geq \underline{v}$ (if (b) holds), we have $\lim_{n\to\infty} \mathbb{T}^n w = v$.*

We conclude this subsection by noting results similar to Proposition 3.1 appear in the existing literature. Proposition 3.1(i) is based on Theorem 4.1(b) of Krasnosel'skii (1964), which assumes the order interval $[\underline{v}, \bar{v}]$ be invariant under $\mathbb{T}$. Invariance is not necessary: all that is required is that either $\mathbb{T}\bar{v} \leq \bar{v}$ or $\mathbb{T}\underline{v} \geq \mathbb{T}\underline{v}$ and the sequence $\{\mathbb{T}^n \bar{v}\}_{n \geq 1}$ or $\{\mathbb{T}^n \underline{v}\}_{n \geq 1}$ is bounded from below or above, respectively.[9] Proposition 3.1(ii) uses similar techniques to the literature on fixed points of order-convex maps (see, e.g., Chapter 5 of Amann (1976)). Unlike much of this literature, Proposition 3.1(ii) does not require additional properties such as compactness and differentiability of $\mathbb{T}$ or strict positivity of $\mathbb{D}_v$, which may be difficult to verify in practice, or that the cone of non-negative functions in $\mathcal{E}$ has non-empty interior, which is a property not shared by L^p spaces with $1 \leq p < \infty$ or Orlicz classes. We do not view Proposition 3.1 as a contribution of this paper: we use it simply as a starting point to derive more primitive existence and uniqueness conditions in the following sections.

3.2 Thin-tailed functions

In the applications that follow, we will take $\mathcal{E}$ to be a class of unbounded but "thin-tailed" functions. Let μ be a probability measure on $(\mathcal{X}, \mathscr{X})$. Most of the applications will feature a stationary Markov process $\{X_t\}_{t \geq 0}$ with statespace $\mathcal{X}$, in which case we shall take μ to be the stationary distribution of the Markov process (see Appendix A). Let L^0 denote the vector space of (equivalence classes of) all measurable functions on $\mathcal{X}$. Define

$$\begin{aligned} L^{\phi_r} &= \{f \in L^0 : \mathbb{E}^\mu[\exp(|f(X)/c|^r)] < \infty \text{ for some } c > 0\}, \\ E^{\phi_r} &= \{f \in L^0 : \mathbb{E}^\mu[\exp(|f(X)/c|^r)] < \infty \text{ for all } c > 0\}, \end{aligned} \tag{8}$$

for $r \geq 1$, where $\mathbb{E}^\mu[\,\cdot\,]$ denotes expectation with respect to the distribution μ of the random variable X. Both L^{ϕ_r} and E^{ϕ_r} are Banach lattices when equipped with the (Luxemburg) norm

$$\|f\|_{\phi_r} = \inf\{c > 0 : \mathbb{E}^\mu[\exp(|f(X)/c|^r)] \leq 2\} \tag{9}$$

and the partial order $f \geq g$ if and only if $f(x) \geq g(x)$ μ-almost everywhere. The space L^{ϕ_r} is an *(exponential) Orlicz space* and E^{ϕ_r} is its *Orlicz heart*.[10] We will be mainly concerned with E^{ϕ_r} in what follows. Lemma B.5 shows that E^{ϕ_r} has the monotone convergence property.

The spaces L^{ϕ_r} and E^{ϕ_r} are related to $L^p(\mu)$ spaces through the embeddings $L^\infty(\mu) \hookrightarrow$

[9] Our requirement that $\mathcal{E}$ has the monotone convergence property is equivalent to the requirement from Krasnosel'skii (1964) that the cone of non-negative functions is "regular".

[10] More generally, Orlicz classes can be defined by replacing $\exp(x^r)$ in display (8) by a convex, increasing function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ for which $\phi(0) = 0$ and $\lim_{x \to \infty} \phi(x)/x = +\infty$ (Krasnosel'skii and Rutickii, 1961). We can then equip this space with the Luxemburg norm $\|f\|_\phi := \inf\{c > 0 : \mathbb{E}^\mu[\phi(|f(X)/c|)] \leq 1\}$. The spaces L^{ϕ_r} and E^{ϕ_r} and norm (9) correspond to the special case in which $\phi(x) = \phi_r(x) := \exp(x^r) - 1$, $r \geq 1$. We use the ϕ_r superscripts and subscripts to avoid confusion with L^p spaces and L^p norms.

$E^{\phi_r} \hookrightarrow L^{\phi_r} \hookrightarrow E^{\phi_s} \hookrightarrow L^{\phi_s} \hookrightarrow L^p(\mu)$ for $1 \le s < r < \infty$, with $\|f\|_p \le p!(\log 2)^{1/r-1}\|f\|_{\phi_r}$ for each $1 \le p < \infty$ where $\|\cdot\|_p$ denotes the $L^p(\mu)$ norm, and $\|f\|_{\phi_s} \le (\log 2)^{1/r-1/s}\|f\|_{\phi_r}$ (van der Vaart and Wellner, 1996, p. 95).

3.3 Verifying the spectral radius condition

In many models featuring forward-looking agents such as those we study in Sections 4 and 5, the subgradient is a discounted conditional expectation operator under a distorted probability measure. That is, there is a wedge between the probability measure describing the evolution of state variables and the probability measure under which the expectation is taken. In this section we show how to verify the key spectral radius condition from Proposition 3.1 under a thin-tail condition on the change of measure.

When there is no such wedge (e.g., time-separable preferences and rational expectations), the spectral radius condition is easily seen to hold. Let $\{X_t\}_{t\ge 0}$ be a time-homogeneous Markov process with transition kernel Q and stationary distribution μ (see Appendix A). In what follows, we define E^{ϕ_r} relative to the stationary distribution μ. Suppose $\mathbb{D}_v = \beta\mathbb{E}^Q$, where $\mathbb{E}^Q$ denotes conditional expectation under Q. Then for any $c > 0$ and $f \in E^{\phi_r}$,

$$\begin{aligned}\mathbb{E}^\mu[\exp(|\mathbb{D}_v f(X_t)/(\beta c)|^r)] &= \mathbb{E}^\mu[\exp(|\mathbb{E}^Q[f(X_{t+1})|X_t]/c|^r)] \\ &\le \mathbb{E}^\mu[\mathbb{E}^Q[\exp(|f(X_{t+1})|^r/c)|X_t]] \\ &= \mathbb{E}^\mu[\exp(|f(X_t)|^r/c)]\,,\end{aligned}$$

by Jensen's inequality and the fact that μ is the stationary distribution of $\{X_t\}_{t\ge 0}$. Taking f to be almost-everywhere constant, we see that the operator $\mathbb{D}_v$ has norm $\|\mathbb{D}_v\|_{\phi_r} = \beta$ on E^{ϕ_r} and $\rho(\mathbb{D}_v; E^{\phi_r}) = \beta$. A similar argument applies for $L^p(\mu)$ spaces.

This argument breaks down in the settings we study, in which $\mathbb{D}_v = \beta\tilde{\mathbb{E}}$, where $\tilde{\mathbb{E}}$ denotes conditional expectation under a distribution different from Q. Suppose

$$\tilde{\mathbb{E}} f(x) = \mathbb{E}^Q[m(X_t, X_{t+1}) f(X_{t+1})|X_t = x]\,, \tag{10}$$

where m is the (conditional) change-of-measure transforming $\mathbb{E}^Q$ into $\tilde{\mathbb{E}}$.[11] We shall verify the spectral radius condition under a thin-tail condition on m. For the intuition behind the result, note that applying $\mathbb{D}_v$ involves multiplying by m, taking conditional expectations under Q, and discounting. Therefore, provided the higher moments of m don't diverge too

[11] That is, for each $x \in \mathcal{X}$, $m(x, \cdot)$ takes values in $\mathbb{R}_+$ and $\int m(x, y) Q(\mathrm{d}y, x) = 1$.

quickly, repeatedly applying $\mathbb{D}_v$ to thin-tailed functions ensures that the effect of discounting eventually dominates and the spectral radius condition holds.

To formalize this reasoning, let $\log m \vee 0$ denote the pointwise maximum of $\log m$ and 0 and let $\mu \otimes Q$ denote the joint (stationary) distribution of (X_t, X_{t+1}) (see Appendix A).

Lemma 3.1. *Let $\mathbb{D} = \beta\tilde{\mathbb{E}}$ where $\beta \in (0,1)$ and $\tilde{\mathbb{E}}$ is of the form (10) with*

$$\mathbb{E}^{\mu\otimes Q}\left[\exp(|\log m(X_t, X_{t+1}) \vee 0|^r/c)\right] < \infty \tag{11}$$

for some $c > 0$ and $r > 1$. Then: $\mathbb{D}$ is a bounded linear operator on E^{ϕ_s} with $\rho(\mathbb{D}; E^{\phi_s}) < 1$ for each $s \geq 1$.

Remark 3.1. *Lemma 3.1 does not require stationarity (or any other property) of $\{X_t\}_{t\geq 0}$ under the law of motion corresponding to $\tilde{\mathbb{E}}$.*

Remark 3.2. *Lemma 3.1 establishes the spectral radius condition for all $\beta \in (0,1)$. When the change of measure m defining $\tilde{\mathbb{E}}$ has thin tails, any amount of discounting is sufficient to overwhelm the effect of the change of measure under repeated application of $\mathbb{D} = \beta\tilde{\mathbb{E}}$.*

4 Application 1: Robust (and related) preferences

4.1 Setting

Consider an infinite-horizon environment in which the continuation value V_t of a stream of per-period utilities $\{U_t\}_{t\geq 0}$ from date t forwards is defined recursively by

$$V_t = U_t - \beta\theta \log \mathbb{E}\left[e^{-\theta^{-1}V_{t+1}} \,\middle|\, \mathcal{F}_t \right] , \tag{12}$$

where $\mathcal{F}_t$ is the date-t information set, $\beta \in (0,1)$ is a time preference parameter, and $\theta > 0$. Recursion (12) arises under preferences for "robustness", namely risk sensitive preferences (Hansen and Sargent, 1995), multiplier preferences (Hansen and Sargent, 2001), and constraint preferences (Hansen et al., 2006). Recursion (12) is also equivalent to the recursion under Epstein and Zin (1989) preferences with unit IES, in which case θ is a transformation of the risk aversion parameter.[12]

[12]Specifically, $\theta = 1/(\gamma - 1)$ where γ is the coefficient of relative risk aversion. See, e.g., Section III in Hansen et al. (2008) for a derivation of recursion (12) from the Epstein–Zin recursion with unit IES.

We follow much of the literature and consider environments characterized by a stationary Markov state process $\{X_t\}_{t\geq 0}$ supported on a state space $\mathcal{X} \subseteq \mathbb{R}^d$. Let $\mathcal{F}_t$ denote the information set generated by the realization of the Markov state up to date t. Let Q denote the Markov transition kernel, $\mathbb{E}^Q$ denote conditional expectation with respect to Q, and μ denote the stationary distribution of $\{X_t\}_{t\geq 0}$ (see Appendix A). In such environments it follows for certain commonly used specifications of U_t that there exists $v : \mathcal{X} \to \mathbb{R}$ and $u : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and such that

$$v(X_t) = -\frac{1}{\theta}\left(V_t - \frac{1}{1-\beta}U_t\right), \qquad u(X_t, X_{t+1}) = U_{t+1} - U_t\,.$$

For instance, this is true when $U_t = \log(C_t)$ and consumption growth $\log(C_{t+1}/C_t)$ is a function of (X_t, X_{t+1}).[13] Under these conditions, the recursion may be rewritten in terms of the scaled continuation value function v:

$$v(x) = \beta \log \mathbb{E}^Q\left[e^{v(X_{t+1})+\alpha u(X_t,X_{t+1})} \Big| X_t = x\right], \tag{13}$$

where $\alpha = -(\theta(1-\beta))^{-1}$. Recursion (13) may be expressed as $v = \mathbb{T}v$, where

$$\mathbb{T}f(x) = \beta \log \mathbb{E}^Q\left[e^{f(X_{t+1})+\alpha u(X_t,X_{t+1})} \Big| X_t = x\right].$$

4.2 Existing results

Hansen and Scheinkman (2012) and Christensen (2017) studied this recursion in the context of Epstein–Zin preferences with unit IES and unbounded $\mathcal{X}$. Hansen and Scheinkman (2012) derived sufficient conditions for existence of a fixed point but not uniqueness. Their conditions restrict moments of a Perron–Frobenius eigenfunction of an operator and require convergence of a sequence of iterates of a related recursion. Christensen (2017) established uniqueness on a neighborhood for the same recursion under a spectral radius condition but did not establish existence or global uniqueness.

[13] Our results trivially extend to allow $\log(C_{t+1}/C_t) = g(X_t, X_{t+1}, Y_{t+1})$ where the conditional distribution of (X_{t+1}, Y_{t+1}) given (X_t, Y_t) depends only on X_t by redefining the state as (X_t, Y_t).

4.3 New results

Here we establish existence and uniqueness under a primitive thin-tail condition on the growth in per-period utility. Formally, we require that for some $r \geq 1$,

$$\mathbb{E}^{\mu\otimes Q}\left[\exp(|u(X_t, X_{t+1})|^r/c)\right] < \infty \quad \text{for all } c > 0, \tag{14}$$

where $\mathbb{E}^{\mu\otimes Q}$ denotes expectation with respect to the stationary distribution of (X_t, X_{t+1}) (see Appendix A). We verify this condition below in several examples. Note, however, that both examples in Section 2 violate this condition.

We shall establish existence and uniqueness by applying Proposition 3.1. The operator $\mathbb{T}$ is continuous, monotone, and convex under condition (14); see Lemma B.7. The proof of existence constructs an upper value $\bar{v}$ and shows the sequence of iterates $\{\mathbb{T}^n \bar{v}\}_{n\geq 1}$ is bounded from below. For uniqueness, by Jensen's inequality the operator $\mathbb{T}$ satisfies inequality (6) with subgradient

$$\mathbb{D}_v f(x) = \beta \mathbb{E}_v f(x)\,,$$

where $\mathbb{E}_v$ is a distorted conditional expectation:

$$\mathbb{E}_v f(x) = \mathbb{E}^Q[m_v(X_t, X_{t+1}) f(X_{t+1}) | X_t = x]\,, \tag{15}$$

$$m_v(X_t, X_{t+1}) = \frac{e^{v(X_{t+1}) + \alpha u(X_t, X_{t+1})}}{\mathbb{E}^Q[e^{v(X_{t+1}) + \alpha u(X_t, X_{t+1})} | X_t]}\,. \tag{16}$$

For robust preferences, $\mathbb{E}_v$ may be interpreted as expectation under the agent's "worst-case" model. The spectral radius condition is verified by applying Lemma 3.1; see Lemma B.8.

Theorem 4.1. *Let condition (14) hold. Then: $\mathbb{T}$ has a fixed point $v \in E^{\phi_r}$. Moreover, if $r > 1$ then: (i) v is the unique fixed point of $\mathbb{T}$ in E^{ϕ_s} for each $s \in (1, r]$, and (ii) v is both the smallest fixed point and the unique stable fixed point of $\mathbb{T}$ in E^{ϕ_1}.*

Example 1: Linear-Gaussian environments. Condition (14) holds for all $r \in [1, 2)$ when $u(X_t, X_{t+1}) = \lambda_0' X_t + \lambda_1' X_{t+1}$ and its stationary distribution is Gaussian.

This specification arises, for instance, with $U_t = \log(C_t e^{\lambda' X_t})$ where $\log(C_{t+1}/C_t)$ is a function of (X_t, X_{t+1}) and the process $\{X_t\}_{t\geq 0}$ is a stationary Gaussian VAR(1):

$$X_{t+1} = \nu + A X_t + u_{t+1}\,, \quad u_{t+1} \sim N(0, \Sigma)\,,$$

with all eigenvalues of A inside the unit circle. This setting was considered in Hansen et al. (2008), Barillas et al. (2009), and several other works. It is known that $\mathbb{T}$ has a fixed point of the form $v(x) = a + b'x$ where $b = \alpha\beta(I - \beta A')^{-1}(\lambda_0 + A'\lambda_1)$ and

$$a = \frac{\beta}{1-\beta}\Big((\alpha\lambda_1 + b)'\nu + \frac{1}{2}(\alpha\lambda_1 + b)'\Sigma(\alpha\lambda_1 + b)\Big).$$

Theorem 4.1 shows that $v(x) = a + b'x$ is the unique fixed point in E^{ϕ_s} for all $s \in (1, 2)$, and the smallest fixed point and unique stable fixed point in E^{ϕ_1}. $\square$

Example 2: Fat tails and rare disasters. Consider the model from Section 2.2. Here with $X_t = (g_t, h_t)$ we have $u(X_t, X_{t+1}) = g_{t+1}$. By iterated expectations we may deduce

$$\mathbb{E}^{\mu\otimes Q}\left[e^{cu(X_t,X_{t+1})}\right] = e^{c\nu_g + \frac{c^2\sigma^2}{2}}\mathbb{E}^{\mu}\left[\exp\left(h_t\left(\exp\left\{c\nu_j + \frac{c^2\sigma_j^2}{2}\right\} - 1\right)\right)\right].$$

Condition (14) is violated for this model: the expectation on the right-hand side is only finite if c is in a neighborhood of zero because the stationary distribution of h_t is a Gamma distribution. Note that uniqueness can fail for this model, as illustrated in Section 2.2.

One could modify this specification so that $w_{z,t+1}|j_{t+1} \sim N(\nu_j j_{t+1}^{\varsigma}, \sigma_j^2)$ for some $\varsigma \in [\frac{1}{2}, 1)$. Given the low frequency of jumps, this modification is likely difficult to distinguish empirically from the original specification. Under this modification, condition (14) holds for each $r \in [1, 1/\varsigma)$. Therefore, there is a unique fixed point $v \in E^{\phi_s}$ for all $s \in (1, 1/\varsigma)$, and v is both the smallest fixed point and the unique stable fixed point in E^{ϕ_1}. $\square$

Example 3: Regime-switching. Consider the same setup from Example 1 but suppose now that the parameters of the VAR are state-dependent (see, e.g., Hamilton (1989), Cecchetti, Lam, and Mark (1990, 2000), Hansen and Sargent (2010), and Ang and Timmermann (2012)):

$$X_{t+1} = \nu_{s_t} + A_{s_t}X_t + u_{t+1}\,, \quad u_{t+1} \sim N(0, \Sigma_{s_t})\,,$$

where s_t is stationary, exogenous Markov state taking values in $\{1, \ldots, N\}$, and all eigenvalues of A_s are inside the unit circle for each $s = 1, \ldots, N$. The full state vector is now (X_t, s_t), which is Markovian and stationary. The stationary distribution of growth in per-period utilities $u(X_t, X_{t+1})$ is sub-Gaussian (see, e.g., Vershynin, 2018, Section 2.5), and so condition (14) holds for all $r \in [1, 2)$. It follows by Theorem 4.1 there is a unique fixed point $v \in E^{\phi_s}$ for all $s \in (1, 2)$ (with E^{ϕ_s} defined with respect to the stationary distribution of (X_t, s_t)), and v is both the smallest fixed point and the unique stable fixed point in E^{ϕ_1}. $\square$

Example 4: Stochastic volatility. Consider the environment from section I.B of Bansal and Yaron (2004) in which consumption growth $g_{t+1} := \log(C_{t+1}/C_t)$ is modeled as

$$
\begin{aligned}
g_{t+1} &= \bar{g} + h_t + \sigma_t \eta^g_{t+1} \,, \\
h_{t+1} &= \rho_h h_t + \varphi_h \sigma_t \eta^h_{t+1} \,, \\
\sigma^2_{t+1} &= \bar{\sigma}^2 + \rho_\sigma(\sigma^2_t - \bar{\sigma}^2) + \varphi_\sigma \eta^\sigma_{t+1} \,,
\end{aligned}
$$

where η^g_t, η^h_t, and η^σ_t are all i.i.d. $N(0,1)$. We alter this model slightly in two respects. First, to focus on the implications of stochastic volatility and simplify exposition we set $\rho_h = 0$ though this is not essential to our analysis. Second, to deal with the complications arising when $\sigma^2_{t+1} < 0$ we take absolute values. This leads to the consumption growth process

$$
\begin{aligned}
g_{t+1} &= \bar{g} + \sqrt{|s_t|}\eta^g_{t+1} \,, \\
s_{t+1} &= \bar{s} + \rho_s(s_t - \bar{s}) + \varphi_s \eta^s_{t+1} \,,
\end{aligned}
$$

where η^g_t and η^s_t are i.i.d. $N(0,1)$. Defining $X_t = (g_t, s_t)$, we see that $u(X_t, X_{t+1}) = g_{t+1}$ when per-period utility is logarithmic in consumption. To verify condition (14), first note that

$$
\mathbb{E}^{\mu\otimes Q}[\exp(|(g_{t+1} - \bar{g})/c|^r)] = \mathbb{E}^\mu[\mathbb{E}[\exp(|\sqrt{|s_t|}\eta^g_{t+1}/c|^r)|s_t]] \,, \tag{17}
$$

where the inner expectation is taken with respect to $\eta^g_{t+1} \sim N(0,1)$. The inner expectation is equivalent to $\mathbb{E}[\exp(Y^r/a^r)]$ where $Y = |Z|$ with $Z \sim N(0,1)$ and $a = c/\sqrt{|s_t|} > 0$. In Appendix B we derive a crude bound on this expectation (see Lemma B.9) from which we may deduce that for $r \in [1, 2)$,

$$
\begin{aligned}
&\mathbb{E}^Q\left[\exp\left(\left|\frac{\sqrt{|s_t|}\eta^g_{t+1}}{c}\right|^r\right)\middle| s_t\right] \\
&\qquad \leq \frac{\sqrt{2}}{\sqrt{\pi}}\left(\left(\frac{2\sqrt{|s_t|}^r}{c^r}\right)^{\frac{1}{2-r}} \exp\left(\frac{(2|s_t|)^{\frac{r}{2-r}}}{c^{\frac{2r}{2-r}}}\right) + \left(\frac{4\sqrt{|s_t|}^r}{c^r}\right)^{\frac{1}{2-r}} + \sqrt{\pi}\right).
\end{aligned}
$$

As the stationary distribution of s_t is Gaussian, the exponent $\frac{r}{2-r}$ of the $|s_t|$ term appearing in the right-hand side exponential must be less than 2 (equivalently, $r \in [1, 4/3)$) so that that the expectation (17) is finite for all $c > 0$. It follows that (14) holds for all $r \in [1, 4/3)$. Therefore, there is a unique fixed point in $v \in E^{\phi_s}$ for all $s \in (1, 4/3)$, and v is both the smallest fixed point and the unique stable fixed point in E^{ϕ_1}. □

4.4 Convergence of compact approximations

While there are many different ways to construct versions of $\mathbb{T}$ over truncated state spaces, a natural approach is to simply restrict $\mathcal{X}$ to a large but compact set $\mathcal{C}$ and rescale the transition density of $\{X_t\}_{t\geq 0}$ accordingly. We close this section by showing that this construction yields an operator $\mathbb{T}_{\mathcal{C}}$ whose fixed point $v_{\mathcal{C}}$ approaches the unique stable fixed point v of $\mathbb{T}$ from below as $\mathcal{C}$ becomes large. Therefore, in view of Theorem 4.1(ii), $v_{\mathcal{C}}$ will *not* converge to any unstable fixed point of $\mathbb{T}$ (if unstable fixed points of $\mathbb{T}$ exist).

The operator $\mathbb{T}_{\mathcal{C}}$ is defined by truncating the support of $\{X_t\}_{t\geq 0}$ to a compact set $\mathcal{C} \subset \mathcal{X}$ and rescaling the transition distribution. Define

$$\mathbb{T}_{\mathcal{C}} f(x) = \beta \log \mathbb{E}^Q \left[e^{f(X_{t+1}) + \alpha u(X_t, X_{t+1})} \frac{1\!\!1\{X_{t+1} \in \mathcal{C}\}}{Q(\mathcal{C}, x)} \right| X_t = x \Bigg] \,, \quad x \in \mathcal{C}\,,$$

where $Q(\mathcal{C}, x)$ is the conditional probability (under the un-truncated transition kernel Q) that $X_{t+1} \in \mathcal{C}$ given $X_t = x$ and $1\!\!1\{x \in \mathcal{C}\} = 1$ if $x \in \mathcal{C}$ and 0 otherwise. Let $B(\mathcal{C})$ denote the space of bounded functions on $\mathcal{C}$ under the sup-norm.

Proposition 4.1. *Let* $\sup_{x\in\mathcal{C}} |\log \mathbb{E}^Q[e^{\alpha u(X_t,X_{t+1})} 1\!\!1\{X_{t+1} \in \mathcal{C}\}/Q(\mathcal{C}, x)|X_t = x]| < \infty$*. Then:* $\mathbb{T}_{\mathcal{C}}$ *has a unique fixed point* $v_{\mathcal{C}} \in B(\mathcal{C})$*. Moreover, if* $\inf_{x\in\mathcal{C}} Q(\mathcal{C}, x) > 0$ *then for any fixed point* v *of* $\mathbb{T}$*,*

$$\inf_{x\in\mathcal{C}} (v(x) - v_{\mathcal{C}}(x)) \geq \frac{\beta}{1-\beta} \left(\inf_{x\in\mathcal{C}} \log Q(\mathcal{C}, x) \right).$$

As $\epsilon_{\mathcal{C}} := -\frac{\beta}{1-\beta} (\inf_{x\in\mathcal{C}} \log Q(\mathcal{C}, x)) > 0$, Proposition 4.1 implies $v_{\mathcal{C}}(x) \leq v(x) + \epsilon_{\mathcal{C}}$ holds for all $x \in \mathcal{C}$. If $\mathbb{T}$ has a second (unstable) fixed point $v' \geq v$, then for any subset of $\mathcal{C}$ upon which v' and v differ by more than $\epsilon_{\mathcal{C}}$, we have $v_{\mathcal{C}}(x) \leq v(x) + \epsilon_{\mathcal{C}} < v'(x)$. As such, $v_{\mathcal{C}}$ cannot converge to v' as $\mathcal{C}$ becomes large (i.e., as $\epsilon_{\mathcal{C}} \to 0$).

5 Application 2: Learning and ambiguity

We now extend the setting from Section 4 to models in which the agent learns about a hidden state, e.g. a regime, stochastic volatility, growth process, or time-varying parameter. This setting is relevant for several types of preferences, including: (i) the extension of multiplier preferences by Hansen and Sargent (2007, 2010) to include concerns about misspecification of beliefs about the hidden state, (ii) generalized recursive smooth ambiguity preferences of Ju and Miao (2012) with unit IES, (iii) special cases of recursive smooth ambiguity preferences studied by Klibanoff et al. (2009), and (iv) Epstein and Zin (1989) recursive preferences with unit IES and learning as used, for example, by Croce et al. (2015).

5.1 Setting

We again consider environments characterized by a Markov state process $\{X_t\}_{t\geq 0}$ with transition kernel Q. Partition the state as $X_t = (\varphi_t, \xi_t)$ where the agent observes φ_t but does not observe ξ_t. Let $\mathcal{O}_t = \sigma(\{\varphi_t, \varphi_{t-1}, \ldots, \varphi_0\})$ denote the history of the observed state to date t. Beliefs about ξ_t are summarized by a posterior distribution Π_t conditional on $\mathcal{O}_t$. We consider environments in which the continuation value V_t of a stream of per-period utilities $\{U_t\}_{t\geq 0}$ from date t forward is defined recursively as

$$V_t = U_t - \beta\theta \log \mathbb{E}^{\Pi_t}\left[\mathbb{E}^{Q}\left[e^{-\vartheta^{-1}V_{t+1}}\middle|\,\mathcal{O}_t, \xi_t\right]^{\frac{\vartheta}{\theta}}\middle|\,\mathcal{O}_t\right], \tag{18}$$

for $\beta \in (0,1)$. This recursion is from Hansen and Sargent (2007, 2010), who introduce an extension of multiplier preferences to accommodate concerns about misspecification of the model (Q) and beliefs about the hidden state (Π_t), where $\vartheta > 0$ and $\theta > 0$ encode concerns about misspecification of Q and Π_t, respectively. When $U_t = \log C_t$, recursion (18) also arises under generalized recursive smooth ambiguity preferences of Ju and Miao (2012) with unit IES, where θ and ϑ are one-to-one transformations of their ambiguity aversion and risk aversion parameters, respectively. When $\vartheta = \theta$, recursion (18) reduces to

$$V_t = U_t - \beta\vartheta \log \mathbb{E}^{\Pi_t}\left[\mathbb{E}^{Q}\left[e^{-\vartheta^{-1}V_{t+1}}\middle|\,\mathcal{O}_t, \xi_t\right]\middle|\,\mathcal{O}_t\right].$$

With $U_t = \log C_t$, this recursion corresponds to Epstein–Zin recursive preferences with unit IES and learning about the hidden state. In the limit as $\vartheta \to \infty$ (thus, the agent is confident in Q but has doubts about the hidden state) recursion (18) becomes

$$V_t = U_t - \beta\theta \log \mathbb{E}^{\Pi_t}\left[e^{-\theta^{-1}\mathbb{E}^{Q}[V_{t+1}|\mathcal{O}_t,\xi_t]}\middle|\,\mathcal{O}_t\right]. \tag{19}$$

This recursion is obtained under recursive smooth ambiguity preferences of Klibanoff et al. (2009), when their ϕ function is $\phi(x) = \exp(-\theta^{-1}x)$.

We impose several (standard) conditions to make the problem tractable. First, the state is assumed to have a conventional hidden Markov structure, in which the conditional distribution of X_{t+1} given X_t factors into the product of a conditional distribution Q_φ for φ_{t+1}

given ξ_t and a conditional distribution Q_ξ for ξ_{t+1} given ξ_t. This nests models with regime-switching studied by Ju and Miao (2012) as well as models with learning about a hidden growth term as in Hansen and Sargent (2007, 2010), Croce et al. (2015) and Collard et al. (2018). Our analysis extends to allow φ_t to influence φ_{t+1}, but we maintain this simpler presentation for convenience.

Second, we assume Π_t is summarized by a finite-dimensional sufficient statistic $\hat{\xi}_t$:

$$\Pi_t(\xi_t) = \Pi_\xi(\xi_t|\hat{\xi}_t)$$

for some conditional distribution Π_ξ, where $\hat{\xi}$ is updated according to a time-invariant rule:

$$\hat{\xi}_{t+1} = \Xi(\hat{\xi}_t, \varphi_{t+1})\,.$$

These conditions are satisfied under Bayesian updating when the state ξ_t takes finitely many values (e.g. a hidden regime) and when X_t evolves as a Gaussian state-space model; see below. The rule for $\hat{\xi}_t$ could also represent belief updating in a boundedly-rational way. Let $\hat{X}_t = (\varphi_t, \hat{\xi}_t)$ and let $\mathcal{X}_{\hat{X}}$, $\mathcal{X}_{\hat{\xi}}$, and $\mathcal{X}_\varphi$ denote the support of $\hat{X}_t$, $\hat{\xi}_t$, and φ_t.

We assume learning is in a "steady state", i.e., $\{(\xi_t, \hat{X}_t)\}_{t\geq 0}$ is stationary. In linear-Gaussian environments, learning corresponds to the Kalman filter. If the filter is not initialized in its steady-state then this process will typically be non-stationary. The stationary problem studied here is a boundary problem representing convergence of the filter to its steady state. Solutions can be obtained by backwards iteration from the steady-state boundary solution.[14] Uniqueness of the limiting steady state recursion is necessary for uniqueness of the sequence of backward iterates.

Finally, we require that there exists $v : \mathcal{X}_{\hat{\xi}} \to \mathbb{R}$ and $u : \mathcal{X}_\varphi \to \mathbb{R}$ such that

$$v(\hat{\xi}_t) = -\frac{1}{\theta}\left(V_t - \frac{1}{1-\beta}U_t\right)\,, \qquad\qquad u(\varphi_{t+1}) = U_{t+1} - U_t\,.$$

Before proceeding, we give two examples of environments in which the preceding conditions hold. In both examples, $U_t = \log(C_t)$ and $\log(C_{t+1}/C_t)$ is a function of φ_{t+1}.

Example 1: Regime switching. Suppose that $\xi_t \in \{1, \dots, N\}$ denotes a hidden Markov state with transition matrix Λ. Let the conditional distribution of φ_{t+1} given $\xi_t = \xi$ have density $q(\cdot|\xi)$. The posterior Π_t is identified with a vector $\hat{\xi}_t$ of regime probabilities given

[14] A similar approach is taken by Collin-Dufresne, Johannes, and Lochstoer (2016) in models featuring Epstein–Zin preferences and learning about parameters of the data-generating process.

$\mathcal{O}_t$. Beliefs $\hat{\xi}_t$ are updated as

$$\hat{\xi}_{t+1} = \Lambda \frac{q(\varphi_{t+1}) \odot \hat{\xi}_t}{1'(q(\varphi_{t+1}) \odot \hat{\xi}_t)} \,,$$

where $q(\varphi_{t+1})$ is the N-vector whose entries are $q(\varphi_{t+1}|\xi)$ for $\xi \in \{1, \ldots, N\}$, $\odot$ denotes element-wise product, and 1 is a N-vector of ones (see, e.g., Hamilton, 1994, Section 4.2).

For example, Ju and Miao (2012) study an economy in which consumption and dividend growth is jointly dependent on a hidden regime ξ_t:

$$\log(C_{t+1}/C_t) = \kappa_{\xi_t} + u^C_{t+1} \,, \quad \log(D_{t+1}/D_t) = \zeta \log(C_{t+1}/C_t) + g_d + u^D_{t+1} \,,$$

where u^C_t and u^D_t are i.i.d. $N(0, \sigma^2_C)$ and $N(0, \sigma^2_D)$. The observable state is $\varphi_t = \log(C_t/C_{t-1})$. The stationary distribution of $u(\varphi_{t+1})$ is a finite mixture of Gaussians. Our results also allow the *volatility* of consumption and dividend growth to be state-dependent. □

Example 2: Gaussian state-space models. Suppose $\{X_t\}_{t \geq 0}$ evolves under Q according to:

$$\varphi_{t+1} = A\xi_t + u^\varphi_{t+1} \,, \quad \xi_{t+1} = B\xi_t + u^\xi_{t+1} \,,$$

where u^φ_t and u^ξ_t are i.i.d. $N(0, \Sigma_u)$ and $N(0, \Sigma_w)$ and all eigenvalues of B are inside the unit circle. This is the setting studied in Hansen and Sargent (2007, 2010), Croce et al. (2015), Collard et al. (2018), and several other works. If $\xi_0 \sim N(\hat{\mu}_0, \hat{\Sigma}_0)$ under Π_0 then $\xi_t \sim N(\hat{\mu}_t, \hat{\Sigma}_t)$ under Π_t. The matrix $\hat{\Sigma}_t$ will converge to a fixed matrix $\bar{\Sigma}$ as $t \to \infty$. In this steady state, the sufficient statistic for Π_t is $\hat{\xi}_t = \hat{\mu}_t$ which is updated using

$$\hat{\xi}_{t+1} = B\hat{\xi}_t + B\bar{\Sigma}A'(A\bar{\Sigma}A' + \Sigma_u)^{-1}(\varphi_{t+1} - A\hat{\xi}_t) \,.$$

The stationary distribution of $u(\varphi_t)$ is Gaussian. □

5.2 Existing results

The only related existence and uniqueness result we are aware of in any of these setting is that of Klibanoff et al. (2009) for recursive smooth ambiguity preferences (recursion (19)). Their result applies to bounded functions and requires bounded per-period utilities.

5.3 New results

Recursion (18) may be reformulated as the fixed-point equation $v = \mathbb{T}v$ where

$$\mathbb{T}f(\hat{\xi}_t) = \beta \log \mathbb{E}^{\Pi_\xi}\left[\mathbb{E}^{Q_\varphi}\left[e^{\frac{\theta}{\vartheta} f(\Xi(\hat{\xi}_t, \varphi_{t+1})) + \alpha u(\varphi_{t+1})} \middle| \xi_t, \hat{\xi}_t \right]^{\frac{\vartheta}{\theta}} \middle| \hat{\xi}_t \right].$$

Recursion (19) in the limiting case with $\vartheta = +\infty$ may be reformulated as $v = \mathbb{T}v$ where

$$\mathbb{T}f(\hat{\xi}_t) = \beta \log \mathbb{E}^{\Pi_\xi}\left[e^{\mathbb{E}^{Q_\varphi}\left[f(\Xi(\hat{\xi}_t, \varphi_{t+1})) + \alpha u(\varphi_{t+1}) | \xi_t, \hat{\xi}_t \right]} \middle| \hat{\xi}_t \right].$$

The existence and uniqueness results presented below apply to either case, though the proofs are presented only for the more involved setting in which $\vartheta < \infty$.

Let $E^{\phi_r}_{\hat{X}}$ be defined relative to the stationary distribution μ of $\hat{X}_t = (\varphi_t', \hat{\xi}_t')'$. Similarly, let $E^{\phi_r}_{\varphi} \subset E^{\phi_r}_{\hat{X}}$ and $E^{\phi_r}_{\hat{\xi}} \subset E^{\phi_r}_{\hat{X}}$ denote functions in $E^{\phi_r}_{\hat{X}}$ depending only on φ or $\hat{\xi}$, respectively. The key regularity condition is again that the stationary distribution of utility growth has thin tails:

$$u \in E^{\phi_r}_{\varphi} \tag{20}$$

for some $r \geq 1$. Note that this condition depends only on the marginal distribution of the observed state and is therefore easy to verify.

We establish existence and uniqueness of fixed points of $\mathbb{T}$ by applying Proposition 3.1. Further details on the form of the subgradient and verification of Lemma 3.1 are deferred to Appendix B.5.

Theorem 5.1. *Let condition (20) hold. Then: $\mathbb{T}$ has a fixed point $v \in E^{\phi_r}_{\hat{\xi}}$. Moreover, if $r > 1$, then: (i) v is the unique fixed point of $\mathbb{T}$ in $E^{\phi_s}_{\hat{\xi}}$ for all $s \in (1, r]$, and (ii) v is both the smallest fixed point and the unique stable fixed point of $\mathbb{T}$ in $E^{\phi_1}_{\hat{\xi}}$.*

Example 1: Regime switching (continued). In the example of Ju and Miao (2012), the stationary distribution of $u(\varphi_{t+1})$ is a finite mixture of Gaussians, so (20) holds for all $r \in [1, 2)$, including when the volatility of consumption and dividend growth is state-dependent. Therefore, there is a unique fixed point $v \in E^{\phi_s}_{\hat{\xi}}$ for all $s \in (1, 2)$, and v is both the smallest fixed point and the unique stable fixed point in $E^{\phi_1}_{\hat{\xi}}$. □

Example 2: Gaussian state-space models (continued). Here the stationary distribution of $u(\varphi_{t+1})$ is Gaussian, so (20) holds for all $r \in [1, 2)$. Therefore, there is a unique fixed point in $v \in E^{\phi_s}_{\hat{\xi}}$ for all $s \in (1,2)$, and v is both the smallest fixed point and the unique stable fixed point in $E^{\phi_1}_{\hat{\xi}}$. $\square$

It is straightforward (albeit more cumbersome notationally) to extend the preceding analysis to allow for u to depend on $(\varphi_t, \varphi_{t+1})$ and to allow the law of motion to be of the more general form in which the conditional distribution of φ_{t+1} given X_t depends on both φ_t and ξ_t. In this case, however, the effective state vector will be $\hat{X}_t$ rather than $\hat{\xi}_t$.

6 Application 3: Epstein–Zin preferences

In this section we study Epstein and Zin (1989) recursive utility with IES $\neq 1$. Existence and uniqueness when state variables have non-compact support is of particular importance as many prominent models, such as those in the long-run risks literature, have non-compact state space. There are currently no uniqueness results for the recursion we study with non-compact state space. This is a complicated issue and it is beyond the scope of the paper to provide a comprehensive treatment. Rather, we show how our approach may be used to derive primitive existence conditions in empirically relevant settings.

6.1 Setting

The continuation value V_t of the agent's consumption plan from time t forward solves

$$V_t = \left\{ (1-\beta)(C_t)^{1-\rho} + \beta \mathbb{E}[(V_{t+1})^{1-\gamma} | \mathcal{F}_t]^{\frac{1-\rho}{1-\gamma}} \right\}^{\frac{1}{1-\rho}} ,$$

where C_t is date-t consumption, $\mathcal{F}_t$ is date-t information, $\gamma \in (0,1) \cup (1,\infty)$ is the coefficient of relative risk aversion, and $1/\rho > 0$ is the elasticity of intertemporal substitution.

We consider the $\rho \neq 1$ case in this section as the $\rho = 1$ case is studied in Section 4. We again consider environments characterized by a stationary Markov process $\{X_t\}_{t \geq 0}$ with state space $\mathcal{X} \subseteq \mathbb{R}^d$. Let Q denote the Markov transition kernel and $\mathbb{E}^Q$ denote conditional expectation under Q. Also let $\log(C_{t+1}/C_t) = g(X_t, X_{t+1})$ for some function g.[15] Then

[15]Our results trivially extend to allow $\log(C_{t+1}/C_t) = g(X_t, X_{t+1}, Y_{t+1})$ where the conditional distribution of (X_{t+1}, Y_{t+1}) given (X_t, Y_t) depends only on X_t by redefining the state as (X_t, Y_t).

$(1-\rho)\log(V_t/C_t) =: v(X_t)$, where v solves

$$v(x) = \log\left((1-\beta) + \beta\mathbb{E}^Q\left[e^{\kappa v(X_{t+1})+(1-\gamma)g(X_t,X_{t+1})}\middle| X_t = x\right]^{\frac{1}{\kappa}}\right) \tag{21}$$

with $\kappa = \frac{1-\gamma}{1-\rho}$ (see, e.g., Hansen et al. (2008)). The properties of this recursion are different for $\kappa < 0$, $\kappa \in (0,1)$, and $\kappa \in [1,\infty)$. We focus on the case $\kappa < 0$, as it is the pertinent case in the long-run risks literature where typically $\gamma > 1$ and $1/\rho > 1$.

6.2 Existing results

Epstein and Zin (1989) and Marinacci and Montrucchio (2010) derived sufficient conditions for existence and uniqueness when consumption growth is bounded. Alvarez and Jermann (2005) establish existence and uniqueness when consumption growth is i.i.d. with bounded innovations. Guo and He (2017) establish sufficient conditions for existence and uniqueness with finite state space. Borovička and Stachurski (2020; BS hereafter) present necessary and sufficient conditions for existence when $\mathcal{X}$ is compact (under additional side conditions on Q). Our results below and those of BS are non-nested if $\mathcal{X}$ is compact: we do not impose any side conditions on Q, but we also do not establish uniqueness in the compact case.

Hansen and Scheinkman (2012; HS hereafter) and BS establish existence with unbounded $\mathcal{X}$ when $\kappa < 0$.[16] We also only present sufficient conditions for existence because the operator does not have a subgradient of the form studied in Section 3.3. Connections between our conditions and those in HS and BS are discussed in more detail below.

6.3 New results

Under general conditions (see Hansen and Scheinkman (2009) and Christensen (2015, 2017)), there exists a strictly positive function ι and scalar $\lambda > 0$ solving[17] the equation

$$\lambda\iota(x) = \mathbb{E}^Q[\iota(X_{t+1})(C_{t+1}/C_t)^{1-\gamma}|X_t = x]\,. \tag{22}$$

[16] Hansen and Scheinkman (2012) and Ren and Stachurski (2020) establish uniqueness when $\kappa \geq 1$.

[17] Note the function ι is defined only up to scale normalization.

Hansen and Scheinkman (2009) use ι and λ to define a distorted conditional expectation operator

$$\tilde{\mathbb{E}} f(x) = \mathbb{E}^Q \left[\frac{\iota(X_{t+1})(C_{t+1}/C_t)^{1-\gamma}}{\lambda \iota(X_t)} f(X_{t+1}) \middle| X_t = x \right] .$$

HS show that solving (21) is equivalent to finding a fixed point of

$$\mathbb{T} f(x) = \log \left((1-\beta)\iota(x)^{-\frac{1}{\kappa}} + \beta \lambda^{\frac{1}{\kappa}} \tilde{\mathbb{E}}[e^{\kappa f(X_{t+1})} | X_t = x]^{\frac{1}{\kappa}} \right) , \tag{23}$$

with the solution to recursion (21) and the fixed point of $\mathbb{T}$ differing additively by $\frac{1}{\kappa} \log \iota$.[18]

We follow HS and assume $\{X_t\}_{t \geq 0}$ is stationary under the law of motion corresponding to the distorted conditional expectation $\tilde{\mathbb{E}}$. Let $\tilde{\mu}$ denote the stationary distribution induced by $\tilde{\mathbb{E}}$ and let $\tilde{E}^{\phi_r}$ denote the corresponding Orlicz heart defined using $\tilde{\mu}$. Our first regularity condition requires that $\log \iota$ has thin tails, in the sense that

$$\log \iota \in \tilde{E}^{\phi_r} \quad \text{for some } r \geq 1. \tag{24}$$

Under this condition, Lemma B.12 shows that $\mathbb{T}$ is a continuous, monotone operator on $\tilde{E}^{\phi_s}$ for each $1 \leq s \leq r$. It is clear that $\mathbb{T} v \geq \log((1-\beta)\iota(x)^{-\frac{1}{\kappa}})$. Therefore, should there exist a $\bar{v} \in \tilde{E}^{\phi_r}$ for which $\mathbb{T}\bar{v} \leq \bar{v}$, the sequence of iterates $\mathbb{T}^n \bar{v}$ must be bounded from below. The remainder of the proof shows that the inequality $\mathbb{T}\bar{v} \leq \bar{v}$ holds for the function

$$\bar{v}(x) = \log \left((1-\beta) \sum_{n=0}^{\infty} (\beta \lambda^{\frac{1}{\kappa}})^n \tilde{\mathbb{E}}^n (\iota^{-\frac{1}{\kappa}})(x) \right) .$$

The sum is convergent under the eigenvalue condition from Hansen and Scheinkman (2012), namely

$$\beta \lambda^{\frac{1}{\kappa}} < 1 . \tag{25}$$

Remark 6.1. *Although $\mathbb{T}$ is not contractive, it follows from Proposition 3.1(i) that the sequence of iterates $\bar{v}, \mathbb{T}\bar{v}, \mathbb{T}^2\bar{v}, \ldots$ will converge to a fixed point of $\mathbb{T}$ under the conditions of Theorem 6.1 and Corollary 6.1 below. The same is true for the sequence of iterates $\underline{v}, \mathbb{T}\underline{v}, \mathbb{T}^2\underline{v}, \ldots$ with $\underline{v}(x) = \log(1-\beta) - \kappa^{-1} \log \iota(x)$.*

Theorem 6.1. *Let $\{X_t\}_{t \geq 0}$ be stationary under the law of motion corresponding to the distorted conditional expectation $\tilde{\mathbb{E}}$, $\kappa < 0$, and conditions (24) and (25) hold. Then: $\mathbb{T}$ has a fixed point in $\tilde{E}^{\phi_s}$ and therefore the recursion (21) has a solution $v \in \tilde{E}^{\phi_s}$ for all $s \in [1, r]$.*

[18]The version of recursion (22) above appears on p. 11968 of HS. In our notation, their recursion is $\hat{\mathbb{U}} g(x) = (1-\beta)\iota(x)^{-\frac{1}{\kappa}} + \beta \lambda^{\frac{1}{\kappa}} \tilde{\mathbb{E}}[g(X_{t+1})^\kappa | X_t = x]^{\frac{1}{\kappa}}$. Recursion (22) is obtained by setting $\mathbb{T} f = \log(\hat{\mathbb{U}}(\exp(f)))$.

Condition (25) is the eigenvalue condition under which HS establish existence in $L^1(\tilde{\mu})$. BS showed this condition is necessary for existence (under some additional operator-theoretic side conditions). Condition (24) is stronger than the integrability conditions imposed on ι in Assumptions 4 and 5 of HS. However, this condition does not seem to bite for models commonly encountered (see the linear-Gaussian example below) and also ensures that the stochastic discount factor (SDF)

$$\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}\left[\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}^Q[V_{t+1}^{1-\gamma}|\mathcal{F}_t]}\right]^{\frac{\rho-\gamma}{1-\gamma}} \equiv \beta e^{-\rho g(X_t,X_{t+1})}\left[\frac{e^{\kappa v(X_{t+1})+(1-\gamma)g(X_t,X_{t+1})}}{\mathbb{E}^Q[e^{\kappa v(X_{t+1})+(1-\gamma)g(X_t,X_{t+1})}|X_t]}\right]^{\frac{\rho-\gamma}{1-\gamma}} \tag{26}$$

is well defined provided consumption growth g has sufficiently thin tails.

Theorem 6.1 has implications for existence in spaces defined relative to the stationary distribution μ of $\{X_t\}_{t\geq 0}$. Suppose that $\tilde{\mu}$ and μ are mutually absolutely continuous and let $\Delta = \frac{\mathrm{d}\tilde{\mu}}{\mathrm{d}\mu}$ denote the change of measure of $\tilde{\mu}$ with respect to μ. Consider the thin-tail condition

$$\mathbb{E}^\mu[\Delta(X_t)^{1+\varepsilon}] < \infty \quad \text{and} \quad \mathbb{E}^\mu[\Delta(X_t)^{-\varepsilon}] < \infty \quad \text{for some } \varepsilon > 0. \tag{27}$$

A sufficient condition for (27) is that $\log \Delta \in L^{\phi_1}$. The spaces $\tilde{E}^{\phi_r}$ (defined using $\tilde{\mu}$) and E^{ϕ_r} (defined using μ) are equivalent under condition (27); see Lemma B.3. We may therefore restate condition (24) as

$$\log \iota \in E^{\phi_r} \quad \text{for some } r \geq 1. \tag{28}$$

Corollary 6.1. *Let $\{X_t\}_{t\geq 0}$ be stationary under the law of motion corresponding to the distorted conditional expectation $\tilde{\mathbb{E}}$, $\kappa < 0$, and conditions (25), (27), and (28) hold. Then: $\mathbb{T}$ has a fixed point in E^{ϕ_s} and therefore the recursion (21) has a solution $v \in E^{\phi_s}$ for all $s \in [1, r]$.*

Example: Linear-Gaussian environments. Consider an environment studied in Section I.A of Bansal and Yaron (2004), Hansen et al. (2008), and Bansal et al. (2014), amongst others, in which

$$X_{t+1} = \nu + AX_t + u_{t+1}\,, \quad u_t \sim N(0, \Sigma)\,,$$

with all eigenvalues of A inside the unit circle, and $g(X_t, X_{t+1}) = \delta' X_{t+1}$ for some vector δ (this is trivially true if log consumption growth is itself a component of X_t). Solving (22),

$$\iota(x) = e^{(1-\gamma)\delta' A(I-A)^{-1}x}\,, \qquad \lambda = e^{\frac{(1-\gamma)^2}{2}\delta'(I-A)^{-1}\Sigma(I-A')^{-1}\delta+(1-\gamma)\delta'(I-A)^{-1}\nu}\,.$$

To apply Corollary 6.1 we must verify conditions (25), (27), and (28). To verify condition (27), first note

$$\frac{\iota(X_{t+1})(C_{t+1}/C_t)^{1-\gamma}}{\lambda\iota(X_t)} = e^{(1-\gamma)\delta'(I-A)^{-1}u_{t+1} - \frac{(1-\gamma)^2}{2}\delta'(I-A)^{-1}\Sigma(I-A')^{-1}\delta}$$

so the u_t are i.i.d. $N((1-\gamma)\delta'(I-A)^{-1}\Sigma, \Sigma)$ under $\tilde{\mathbb{E}}$. Equivalently, under $\tilde{\mathbb{E}}$ we have

$$X_{t+1} = \nu + (1-\gamma)\delta'(I-A)^{-1}\Sigma + AX_t + u_{t+1}\,, \quad u_t \sim N(0,\Sigma)\,.$$

This implies the stationary distributions μ and $\tilde{\mu}$ are both Gaussian, with different means but the same covariance. In consequence, $\log \Delta(x)$ is affine in x and so condition (27) holds for any $\varepsilon > 0$. As $\log \iota(x)$ is also affine in x, we have that $\log \iota \in E^{\phi_r}$ for all $r \in [1,2)$, which verifies condition (28). It follows that the single condition one needs to verify for existence of recursive utilities in linear-Gaussian environments is the eigenvalue condition (25), which reduces to

$$\beta e^{\frac{(1-\rho)(1-\gamma)}{2}\delta'(I-A)^{-1}\Sigma(I-A')^{-1}\delta + (1-\rho)\delta'(I-A)^{-1}\nu} < 1\,.$$

Note also that as $g(X_t, X_{t+1}) = \delta' X_{t+1}$, which belongs to E^{ϕ_r} for $r \in [1,2)$, the SDF (26) is therefore well defined and all of its moments exist. □

A Definitions of some mathematical terms

This appendix presents definitions of some mathematical terms used in the main text. We refer the reader to standard references (e.g. Aliprantis and Border (1999)) for further details.

A set $\mathcal{E}$ equipped with a partial order $\geq$ (i.e., a transitive, reflexive, and antisymmetric relation on $\mathcal{E}$) is a partially ordered set. Say that $\mathcal{E}$ is a *lattice* if each pair of elements $f, g \in \mathcal{E}$ has a supremum, denoted $f \vee g$, and infimum, denoted $f \wedge g$, in $\mathcal{E}$.[19] Say that $\mathcal{E}$ is a *Banach lattice* if it is a Banach space when equipped with a norm $\|\cdot\|$ and $\|\cdot\|$ has the property that for $f, g \in \mathcal{E}$, $|f| \leq |g|$ implies $\|f\| \leq \|g\|$, where $|f| = f \vee (-f)$. Orlicz spaces defined relative to a measure μ and $L^p(\mu)$ spaces are Banach lattices when equipped with their Orlicz (or Luxemburg) and $L^p(\mu)$ norms, respectively, and the partial order $f \geq g$ if and only if $f(x) \geq g(x)$ for μ-almost every x.

[19] That is, $f \leq (f \vee g)$, $g \leq (f \vee g)$, and $f \leq h$ and $g \leq h$ imply $(f \vee g) \leq h$. The infimum is defined analogously.

Let $(\mathcal{X}, \mathscr{X})$ be a measurable space. A stochastic process $\{X_t\}_{t\geq 0}$ is a Markov process[20] with statespace $\mathcal{X}$ if for each $t \geq 0$, X_t is a random variable taking values in $\mathcal{X}$, and for each $t \geq 0$ and $k \geq 1$, the conditional distribution of X_{t+k} given $\{X_s\}_{0\leq s\leq t}$ depends only on X_t. Say $\{X_t\}_{t\geq 0}$ is *time-homogeneous* if for each $t \geq 0$ and $x \in \mathcal{X}$, the conditional distribution of X_{t+1} given $X_t = x$ can be described by a *transition kernel* $Q(\cdot, x)$. That is, for each $x \in \mathcal{X}$, $Q(\cdot, x)$ is a probability measure on $(\mathcal{X}, \mathscr{X})$, and $Q(A, \cdot)$ is $\mathscr{X}$-measurable for each $A \in \mathscr{X}$. The conditional expectation of $f(X_t, X_{t+1})$ given $X_t = x$ is

$$\mathbb{E}^Q[f(X_t, X_{t+1})|X_t = x] := \int f(x, y)Q(\mathrm{d}y, x).$$

A probability measure μ on $(\mathcal{X}, \mathscr{X})$ is a *stationary distribution* if $\mu(A) = \int Q(A, x)\mu(\mathrm{d}x)$ for all $A \in \mathscr{X}$. In addition, say $\{X_t\}_{t\geq 0}$ is *stationary* if μ is unique and $\{X_t\}_{t\geq 0}$ is initialized by drawing X_0 from the stationary distribution μ. As such, the marginal distributions of $(X_t, \ldots, X_{t+k})$ do not depend on t. In particular, for any $t \geq 0$, the (unconditional) expected values of functions of X_t and (X_t, X_{t+1}) are given by

$$\mathbb{E}^\mu[f(X_t)] = \int f\,\mathrm{d}\mu\,, \qquad \mathbb{E}^{\mu\otimes Q}[h(X_t, X_{t+1})] = \int h(x, y)Q(\mathrm{d}y, x)\mu(\mathrm{d}x)\,,$$

for all bounded measurable $f : \mathcal{X} \to \mathbb{R}$ and $h : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$.

B Proofs

Remark B.1. *Several of the proofs below require showing that a function f is an element of E^{ϕ_s} with $s \geq 1$. That is, that $\mathbb{E}^\mu[\exp(|f(X_t)/c|^s)] < \infty$ holds for all $c > 0$. For any $0 < \bar{c} < c$ we have $(\bar{c}/c)^s < 1$ and therefore*

$$\exp(|f(X_t)/c|^s) = (\exp(|f(X_t)/\bar{c}|^s))^{(\bar{c}/c)^s} \leq \exp(|f(X_t)/\bar{c}|^s)$$

because $\exp(|f(X_t)/\bar{c}|^s) \geq 1$. In order to show that $f \in E^{\phi_s}$, one therefore only has to check that $\mathbb{E}^\mu[\exp(|f(X_t)/c|^s)] < \infty$ holds for all $c \in (0, \epsilon)$ for any fixed $\epsilon > 0$.

[20] In this paper we consider discrete-time processes, for which the time index t ranges over the non-negative integers.

B.1 Ancillary results

A version of this first Lemma appears in Chapter 2.3 of the manuscript Pollard (2015) and is used frequently to control the Orlicz norm $\|\cdot\|_{\phi_r}$. We include a proof for convenience.

Lemma B.1 (Pollard (2015))**.** *Let* $\mathbb{E}^\mu[\exp(|f(X)/C|^r)] - 1 \le C'$ *for finite constants* $C > 0$ *and* $C' \ge 1$*. Then:* $\|f\|_{\phi_r} \le CC'$*.*

Proof of Lemma B.1. Take $\tau \in [0,1]$. By convexity of $\phi_r(x) := \exp(|x|^r) - 1$, we have

$$\mathbb{E}^\mu[\phi_r(\tau|f(X)|/C)] \le \tau\mathbb{E}^\mu[\phi_r(|f(X)|/C)] + (1-\tau)\phi_r(0) = \tau\mathbb{E}^\mu[\phi_r(|f(X_t)|/C)]\,.$$

The result follows by setting $\tau = 1/C'$. □

For the next lemma, recall that $\|\cdot\|_p$ denotes the usual L^p norm for $1 \le p < \infty$.

Lemma B.2 (Karakostas (2008); Chen, Jia, and Jiao (2016))**.** *Let* $1 < p_i < \infty$ *for* $i \in \mathbb{N}$*, and* $\sum_{i=1}^\infty \frac{1}{p_i} = 1$*. If* $\prod_{i=1}^\infty \|f_i\|_{p_i} < \infty$ *then* $\prod_{i=1}^\infty f_i$ *is well defined and* $\|\prod_{i=1}^\infty f_i\|_1 \le \prod_{i=1}^\infty \|f_i\|_{p_i}$*.*

Let μ and ν be two probability measures on a measurable space $(\mathcal{X}, \mathscr{X})$. We make explicit the dependence of function classes and norms on the measures μ and ν. Let $\Delta = \frac{\mathrm{d}\mu}{\mathrm{d}\nu}$, and let $\|\Delta\|_{L^p(\nu)}$ denote its $L^p(\nu)$ norm.

Lemma B.3. *Let* $\mu \ll \nu$ *and* $\int \Delta^p \,\mathrm{d}\nu < \infty$ *for some* $p > 1$*. Then:* $E^{\phi_r}(\nu) \hookrightarrow E^{\phi_r}(\mu)$ *and* $L^{\phi_r}(\nu) \hookrightarrow L^{\phi_r}(\mu)$ *for each* $r \ge 1$*.*

Proof of Lemma B.3. To see that $E^{\phi_r}(\nu) \subseteq E^{\phi_r}(\mu)$, take any $f \in E^{\phi_r}(\nu)$ and $c > 0$. Then:

$$\mathbb{E}^\mu\left[e^{|f(X)/c|^r}\right] = \mathbb{E}^\nu\left[\Delta(X)e^{|f(X)/c|^r}\right] \le \|\Delta\|_{L^p(\nu)}\mathbb{E}^\nu\left[e^{\left|f(X)/(c/q^{1/r})\right|^r}\right]^{\frac{1}{q}} < \infty\,,$$

where $q > 1$ is the dual index of p. Therefore, $f \in E^{\phi_r}(\mu)$. Similarly, $L^{\phi_r}(\nu) \subseteq L^{\phi_r}(\mu)$.

For continuity of the embedding, take $f \in L^{\phi_r}(\nu)$ and $c = q^{\frac{1}{r}}\|f\|_{\phi_r(\nu)}$. Substituting into the above display yields

$$\mathbb{E}^\mu[e^{|f(X)/c|^r}] \le 2^{\frac{1}{q}}\|\Delta\|_{L^p(\nu)}\,.$$

Therefore, $\|f\|_{L^{\phi_r}(\mu)} \le ((2^{\frac{1}{q}}\|\Delta\|_{L^p(\nu)} - 1) \vee 1)q^{\frac{1}{r}}\|f\|_{L^{\phi_r}(\nu)}$ by Lemma B.1. □

B.2 Proofs for Section 2

Proof of Proposition 2.1. Suppose a solution $v \in L^1$ to (3) does indeed exist for some $\alpha \neq 0$. Then v is a fixed point the operator $\mathbb{T}$. Consider the related operator $\mathbb{S}$, given by

$$\mathbb{S}f(h) = \mathsf{a} + \mathsf{b}e^{2h} + \beta \mathbb{E}^Q \left[f(h_{t+1}) \middle| \, h_t = h \right].$$

As $\mathbb{S}$ is a contraction mapping on L^1, we may deduce it has a unique fixed point $w \in L^1$ given by

$$w(h) = \frac{\mathsf{a}}{1-\beta} + \mathsf{b} \sum_{i=0}^{\infty} \beta^i \mathbb{E}^Q [e^{2h_{t+i}} | h_t = h].$$

Note by Jensen's inequality that $\mathbb{T}f \geq \mathbb{S}f$ holds for any f, where $f \geq g$ means $f(h) \geq g(h)$ holds μ-almost everywhere. Note also that $w - v = \mathbb{S}w - \mathbb{T}v \leq \mathbb{S}w - \mathbb{S}v$, where $\mathbb{S}w(h) - \mathbb{S}v(h) = \beta \mathbb{E}^Q[w(h_{t+1}) - v(h_{t+1}) | h_t = h] =: \mathbb{D}(w-v)(h)$. Therefore, $(\mathbb{I} - \mathbb{D})(w - v) \leq 0$, where $\mathbb{I}$ denotes the identity operator. As $(\mathbb{I} - \mathbb{D})$ is invertible on L^1 (see the discussion in Section 3.3) and its inverse maps non-negative functions to non-negative functions, we have $w - v \leq 0$ and hence that $v \geq w$. Also note that $w \geq \underline{w}$, where

$$\underline{w}(h) = \frac{\mathsf{a}}{1-\beta} + \mathsf{b}e^{2h}.$$

By monotonicity and the fact that the fixed point v of $\mathbb{T}$ is bounded below by $\underline{w}$, we have

$$v = \mathbb{T}v \geq \mathbb{T}\underline{w}, \tag{29}$$

where

$$\mathbb{T}\underline{w}(h) = \mathsf{a} + \mathsf{b}e^{2h} + \beta \log \mathbb{E}^Q \left[\exp\left(\frac{\mathsf{a}}{1-\beta} + \mathsf{b}e^{2h_{t+1}} \right) \middle| \, h_t = h \right].$$

But note that the right-hand side expectation is $+\infty$ for every h because $\mathsf{b} > 0$. It follows by inequality (29) that $v(h) = +\infty$ almost everywhere, which contradicts $v \in L^1$. $\square$

Proof of Proposition 2.2. Substituting $v(h) = a + bh$ into (5) and using the conditional characteristic function for the autoregressive gamma process (Backus et al., 2014, Appendix H), we obtain

$$a + bh = \mathsf{a} + \mathsf{b}h + \beta a + \frac{\beta \varphi b}{1 - bc} h - \beta \delta \log(1 - bc).$$

Matching coefficients gives a quadratic equation in b. When $\mathsf{q} := 1 + c\mathsf{b} - \beta\varphi$ satisfies

$\mathsf{q}^2 - 4c\mathsf{b} > 0$, there are two solutions for b:

$$b_1 = \frac{\mathsf{q} - \sqrt{\mathsf{q}^2 - 4c\mathsf{b}}}{2c}, \qquad b_2 = \frac{\mathsf{q} + \sqrt{\mathsf{q}^2 - 4c\mathsf{b}}}{2c},$$

both of which satisfy $1 - bc > 0$. Therefore, there are two solutions of the form $v_i(h) = a_i + b_i h$, where $a_i = \frac{\mathsf{a} - \beta\delta \log(1-b_i c)}{1-\beta}$, $i = 1, 2$. □

B.3 Proofs for Section 3

Proof of Proposition 3.1. Existence: we prove this for case (a); similar arguments apply for (b). The sequence $\{\bar{v}_n\}_{n\geq 1}$ with $\bar{v}_n = \mathbb{T}^n \bar{v}$ is monotone and bounded below by $\underline{v}$. It follows by the monotone convergence property that $\{\bar{v}_n\}_{n\geq 1}$ converges to some $v \in \mathcal{E}$ with $v \geq \underline{v}$. Finally, $\|\mathbb{T}v - v\| \leq \|\mathbb{T}v - \mathbb{T}\bar{v}_n\| + \|\mathbb{T}\bar{v}_n - v\| = \|\mathbb{T}v - \mathbb{T}\bar{v}_n\| + \|\bar{v}_{n+1} - v\| \to 0$ by continuity of $\mathbb{T}$, hence $\mathbb{T}v = v$.

Uniqueness: Suppose $\mathbb{T}$ satisfies (6) at each fixed point. Let $v, v' \in \mathcal{E}$ be fixed points of $\mathbb{T}$. By (6), we have $v' - v = \mathbb{T}v' - \mathbb{T}v \geq \mathbb{D}_v(v' - v)$, which implies that

$$(\mathbb{I} - \mathbb{D}_v)(v' - v) \geq 0\,. \tag{30}$$

As $\rho(\mathbb{D}_v; \mathcal{E}) < 1$, we have $(\mathbb{I} - \mathbb{D}_v)^{-1} = \sum_{i=0}^{\infty} (\mathbb{D}_v)^i$ where the series converges in operator norm (Kress, 2014, Theorem 10.15). The operator $\mathbb{D}_v$ is monotone and so $(\mathbb{I} - \mathbb{D}_v)^{-1}$ is also monotone. Applying $(\mathbb{I} - \mathbb{D}_v)^{-1}$ to both sides of equation (30) yields $v' - v \geq 0$. A parallel argument yields $v - v' \geq 0$. Therefore, $v = v'$. The proof follows by similar arguments when $\mathbb{T}$ instead satisfies (7) at each of its fixed points. □

Proof of Corollary 3.1. Suppose $\mathbb{T}$ satisfies (6) at each fixed point. By (6), for $v, v' \in \mathcal{V}$:

$$v' - v = \mathbb{T}v' - \mathbb{T}v \geq \mathbb{D}_v(v' - v)$$

hence $(\mathbb{I} - \mathbb{D}_v)(v' - v) \geq 0$. When $\rho(\mathbb{D}_v; \mathcal{E}) < 1$, the operator $(\mathbb{I} - \mathbb{D}_v)$ is invertible on $\mathcal{E}$ with $(\mathbb{I} - \mathbb{D}_v)^{-1} = \sum_{n=0}^{\infty} \mathbb{D}_v^n$. As $\mathbb{D}_v$ is monotone, so too is $(\mathbb{I} - \mathbb{D}_v)^{-1}$. Applying $(\mathbb{I} - \mathbb{D}_v)^{-1}$ to both sides of the above display yields $v' - v \geq 0$, so v is the smallest fixed point of $\mathbb{T}$.

Suppose any other $v' \in \mathcal{V}$ distinct from v was also stable. Then we could apply an identical argument to obtain the reverse inequality $v - v' \geq 0$, a contradiction. The proof when $\mathbb{T}$ satisfies (7) at each fixed point follows similarly. □

Lemma B.4. *Let $v \in \mathcal{E}$ be a stable fixed point of $\mathbb{T}$, and let there exist a neighborhood N of v for which*

$$\mathbb{T}f - \mathbb{T}v = \mathbb{D}_v(f - v) + o(\|f - v\|) \tag{31}$$

for all $f \in N$. Then: there exists a neighborhood N' of v for which $\lim_{n\to\infty} \mathbb{T}^n f = v$ for all $f \in N'$

Proof of Lemma B.4. As $\rho(\mathbb{D}_v; \mathcal{E}) < 1$, there exists $n_0 \in \mathbb{N}$ and $\epsilon > 0$ for which $\|(\mathbb{D}_v)^{n_0} f\| \leq e^{-\epsilon n_0}\|f\|$ for all $f \in \mathcal{E}$. Recursively applying condition (31), we may deduce that there exists a sufficiently small neighborhood N' of v upon which

$$\mathbb{T}^n f - v = (\mathbb{D}_v)^n (f - v) + o(\|f - v\|)\,, \qquad \text{for all } 1 \leq n \leq n_0\,, \tag{32}$$

and hence

$$\|\mathbb{T}^{n_0} f - v\| \leq e^{-\epsilon n_0}\|f - v\| + o(\|f - v\|)\,.$$

Making N' smaller if necessary, we may therefore deduce that there is a $\varrho \in (0, 1)$ for which $\|\mathbb{T}^{n_0} f - v\| \leq \varrho\|f - v\|$ holds for all $f \in N'$. For any $f \in N'$ and $k \in \mathbb{N}$, we therefore have that $\|\mathbb{T}^{kn_0} f - v\| \leq \varrho^k\|f - v\|$. Moreover, for any $n \in \mathbb{N}$ that is not an integer multiple of n_0, it follows by (32) with $k = \lfloor n/n_0 \rfloor$ that $\mathbb{T}^n f - v = \mathbb{T}^{n-kn_0}(\mathbb{T}^{kn_0} f) - v = (\mathbb{D}_v)^{n-kn_0}(\mathbb{T}^{kn_0} f - v) + o(\|\mathbb{T}^{kn_0} f - v\|) = O(\|\mathbb{T}^{kn_0} f - v\|) = O(\varrho^k)$. $\square$

Proof of Corollary 3.2. Suppose condition (a) holds. Fix $w \in \mathcal{E}$ with $w \leq \bar{v}$, let $w_0 = w$, and let $w_n = \mathbb{T}^n w$ for $n \in \mathbb{N}$. Also let $\bar{v}_n = \mathbb{T}^n \bar{v}$. By Proposition 3.1 we know that there is a unique fixed point $v \in \mathcal{E}$. Then by monotonicity of $\mathbb{T}$ and the subgradient inequality (6), for every $n \in \mathbb{N}$ we have

$$\bar{v}_n - v \geq w_n - v = \mathbb{T}w_{n-1} - \mathbb{T}v \geq \mathbb{D}_v(w_{n-1} - v) \geq (\mathbb{D}_v)^n (w - v)\,,$$

where the final inequality is by monotonicity of $\mathbb{D}_v$. The left-hand side term $\bar{v}_n - v \to 0$ as $n \to \infty$ by Proposition 3.1. Moreover, as $\rho(\mathbb{D}_v; \mathcal{E}) < 1$, there exists $n_0 \in \mathbb{N}$ and $\epsilon > 0$ for which $\|(\mathbb{D}_v)^{n_0} f\| \leq e^{-\epsilon n_0}\|f\|$ for all $f \in \mathcal{E}$, from which we may deduce that the right-hand side term $(\mathbb{D}_v)^n(w - v) \to 0$ as $n \to \infty$. As $\|\cdot\|$ is a lattice norm, it follows that $w_n \to v$ as $n \to \infty$. The proof when (b) holds and $\mathbb{T}$ satisfies (7) follows similarly. $\square$

Lemma B.5. *Let μ be a probability measure on $(\mathcal{X}, \mathscr{X})$. Then: for any $r \geq 1$, the space E^{ϕ_r} has the monotone convergence property.*

Proof of Lemma B.5. Let $\{f_n\}_{n\geq 1} \subset E^{\phi_r}$ be an increasing sequence of functions bounded above by some $g \in E^{\phi_r}$. As $E^{\phi_r} \hookrightarrow L^1(\mu)$, the sequence $\{f_n\}_{n\geq 1}$ is uniformly bounded in $L^1(\mu)$ and so it follows by Beppo Levi's monotone convergence theorem (Malliavin, 1995, Theorem I.7.1) that there exists $f \in L^1(\mu)$ for which $\lim_{n\to\infty} f_n = f$ (μ-almost everywhere) and $\lim_{n\to\infty} \|f_n - f\|_1 = 0$, where $\|\cdot\|_1$ denotes the $L^1(\mu)$ norm. As $f_1 \leq f \leq g$, we have $|f| \leq |f_1| + |g|$. Moreover, as $f_1, g \in E^{\phi_r}$, for any $c > 0$ we have

$$\begin{aligned}\mathbb{E}^\mu[\exp(|f(X)/c|^r)] &\leq \mathbb{E}^\mu[\exp(((|f_1(X)| + |g(X)|)/c)^r)] \\ &\leq \frac{1}{2}\mathbb{E}^\mu[\exp(|2f_1(X)/c|^r)] + \frac{1}{2}\mathbb{E}^\mu[\exp(|2g(X)/c|^r)] < \infty\,,\end{aligned}$$

from which it follows that $f \in E^{\phi_r}$.

To establish convergence in $\|\cdot\|_{\phi_r}$, suppose that $\limsup_{n\to\infty} \|f_n - f\|_{\phi_r} \geq 2\varepsilon$ for some $\varepsilon > 0$. Then

$$\limsup_{n\to\infty} \mathbb{E}^\mu[\exp(|(f_n(X) - f(X))/\varepsilon|^r)] \geq 2\,. \tag{33}$$

Note that $\{g_n\}_{n\geq 1}$ with $g_n = \exp(|(f_n - f)/\varepsilon|^r)$ is a monotone sequence of non-negative functions with $\limsup_{n\to\infty} g_n = 0$ (μ-almost everywhere). Moreover, for each $n \geq 1$ we have that

$$g_n \leq \exp\left(((|f_1| + |g| + |f|)/\varepsilon)^r\right),$$

and the right-hand side is μ-integrable because $f_1, g, f \in E^{\phi_r}$. Therefore, by reverse Fatou:

$$\limsup_{n\to\infty} \mathbb{E}^\mu[\exp(|(f_n(X) - f(X))/\varepsilon|^r)] \leq \mathbb{E}^\mu\big[\limsup_{n\to\infty} \exp(|(f_n(X) - f(X))/\varepsilon|^r)\big] = 0\,,$$

contradicting (33). It follows that $\|f_n - f\|_{\phi_r} \to 0$. $\square$

Remark B.2. *It follows by identical arguments to Lemma B.5 that the Orlicz heart* $E^\phi := \{f \in L^0 : \mathbb{E}^\mu[\phi(f(X)/c)] < \infty$ *for all* $c > 0\}$ *defined using any monotone, strictly convex* $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ *with* $\phi(0) = 0$ *and* $\lim_{x\to\infty} \phi(x)/x \to +\infty$ *has the monotone convergence property when equipped with the Luxemburg norm* $\|f\|_\phi := \inf\{c > 0 : \mathbb{E}^\mu[\psi(|f(X)/c|)] \leq 1\}$.

We next present an intermediate result used to prove Lemma 3.1. Note that condition (11) implies that $(\log m \vee 0) \in L^{\phi_r}(\mu \otimes Q)$, the Orlicz class of functions $f : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ defined relative to the stationary distribution $\mu \otimes Q$ of (X_t, X_{t+1}). With slight abuse of notation, let $\|(\log m \vee 0)\|_{\phi_r}$ denote the corresponding Orlicz norm of $(\log m \vee 0)$.

Lemma B.6. *Let* $\tilde{\mathbb{E}}$ *be of the form (10) and let* m *satisfy condition (11). Then for any*

$p \in (1, \infty)$ and $n \geq 1$:

$$\mathbb{E}^{\mu \otimes Q}[m(X_t, X_{t+1})^{np}]^{1/p} \leq e^{(2n \|(\log m \vee 0)\|_{\phi_r})^{\frac{r}{r-1}} (2p)^{\frac{1}{r-1}}} + 2^{\frac{3}{2p}} .$$

Moreover, for any $\beta \in (0,1)$ there exists $C \in (0, \infty)$ and $c \in (0, 1-\beta)$ depending only on β, r, $\|(\log m \vee 0)\|_{\phi_r}$, and p such that the inequality

$$\mathbb{E}^{\mu \otimes Q}[m(X_t, X_{t+1})^{np}]^{1/p} \leq C e^{(\beta + c)^{-n}}$$

holds for each $n \geq 1$.

Proof of Lemma B.6. First note $\mathbb{E}^{\mu \otimes Q}[m(X_t, X_{t+1})^{np}] \leq \mathbb{E}^{\mu \otimes Q}[e^{np|\log m(X_t, X_{t+1}) \vee 0|}]$. To simplify notation, let $Y_t = (X_t, X_{t+1})$, $a = \log m \vee 0$, and $\|a\|_{\phi_r} = \|(\log m \vee 0)\|_{\phi_r}$. In what follows, all probabilities (denoted $\Pr(\cdot)$) are taken with respect to $\mu \otimes Q$. Let A be a positive constant (specified below) and set $|a| = a_+ + a_-$ with $a_+ = |a| 1\!\!1\{|a| \leq A\}$ and $a_- = |a| 1\!\!1\{|a| > A\}$. For any $z > 0$, we have

$$\Pr\left(e^{np|a(Y_t)|} \geq z\right) \leq \Pr\left(a_+(Y_t) \geq \frac{\log z}{2np}\right) + \Pr\left(a_-(Y_t) \geq \frac{\log z}{2np}\right) . \tag{34}$$

By Markov's inequality and definition of $\|\cdot\|_{\phi_r}$, we have

$$\begin{aligned}
\Pr\left(a_-(Y_t) \geq \frac{\log z}{2np}\right) &\leq \Pr\left(|a(Y_t)|^r \geq \frac{A^{r-1} \log z}{2np}\right) \\
&= \Pr\left(\exp\left(\frac{|a(Y_t)|^r}{\|a\|_{\phi_r}^r}\right) \geq \exp\left(\frac{1}{\|a\|_{\phi_r}^r} \frac{A^{r-1} \log z}{2np}\right)\right) \\
&\leq \frac{\mathbb{E}^{\mu \otimes Q}\left[\exp\left(|a(Y_t)/\|a\|_{\phi_r}|^r\right)\right]}{\exp\left(\frac{1}{\|a\|_{\phi_r}^r} \frac{A^{r-1} \log z}{2np}\right)} \\
&\leq 2 \exp\left(-\frac{1}{\|a\|_{\phi_r}^r} \frac{A^{r-1} \log z}{2np}\right) .
\end{aligned}$$

Setting $A = (\|a\|_{\phi_r}^r 4np)^{\frac{1}{r-1}}$, we obtain

$$\Pr\left(a_-(Y_t) \geq \frac{\log z}{2np}\right) \leq 2z^{-2} .$$

As $2z^{-2} \geq 1$ if $z \leq \sqrt{2}$, we therefore have

$$\int_0^\infty \Pr\left(a_-(Y_t) \geq \frac{\log z}{2np}\right) \mathrm{d}z \leq \sqrt{2} + 2\int_{\sqrt{2}}^\infty z^{-2}\,\mathrm{d}z = 2^{\frac{3}{2}}\,. \tag{35}$$

For the first term on the right-hand side of (34), as $a_+ \leq A$ we have

$$\Pr\left(a_+(Y_t) \geq \frac{\log z}{2np}\right) = 0 \text{ if } z > e^{2npA}\,. \tag{36}$$

Note $2npA = (2np\|a\|_{\phi_r})^{\frac{r}{r-1}} 2^{\frac{1}{r-1}}$. Using the fact that $\mathbb{E}[Z] = \int_0^\infty \Pr(Z \geq z)\,\mathrm{d}z$ for a non-negative random variable Z, we may deduce from (34), (35), and (36) that

$$\mathbb{E}^{\mu\otimes Q}[m(X_t, X_{t+1})^{np}] \leq \int_0^\infty \Pr(e^{np|a(Y)|} \geq z)\,\mathrm{d}z \leq e^{(2np\|a\|_{\phi_r})^{\frac{r}{r-1}} 2^{\frac{1}{r-1}}} + 2^{\frac{3}{2}}\,.$$

The first assertion follows because $(x+y)^{1/p} \leq x^{1/p} + y^{1/p}$ for $x, y \geq 0$ and $p \geq 1$. The second assertion follows as $n^{\frac{r}{r-1}} = o((\beta + c)^{-n})$ for any $\beta \in (0,1)$ and $c \in (0, 1-\beta)$. $\square$

Proof of Lemma 3.1. We first show $\mathbb{D}$ is a bounded linear operator on L^{ϕ_s} for any $s \geq 1$. Linearity follows by inspection. For boundedness, fix any $s \geq 1$ and take any $f \in L^{\phi_s}$ with $\|f\|_{\phi_s} > 0$ and any $q \in (1, \infty)$. By applying Jensen's inequality, definition of $\tilde{\mathbb{E}}$ from (10), and Hölder's inequality with $p^{-1} + q^{-1} = 1$, we obtain

$$\begin{aligned}
\mathbb{E}^\mu\left[e^{|\mathbb{D}f(X_t)/(q^{\frac{1}{s}}\beta\|f\|_{\phi_s})|^s}\right] &= \mathbb{E}^\mu\left[e^{q^{-1}|\tilde{\mathbb{E}}f(X_t)/\|f\|_{\phi_s}|^s}\right] \\
&\leq \mathbb{E}^{\mu\otimes Q}\left[m(X_t, X_{t+1})e^{q^{-1}|f(X_{t+1})/\|f\|_{\phi_s}|^s}\right] \\
&\leq \mathbb{E}^{\mu\otimes Q}\left[m(X_t, X_{t+1})^p\right]^{\frac{1}{p}} \mathbb{E}^\mu\left[e^{|f(X_t)/\|f\|_{\phi_s}|^s}\right]^{\frac{1}{q}} \\
&\leq 2^{\frac{1}{q}}\mathbb{E}^{\mu\otimes Q}\left[m(X_t, X_{t+1})^p\right]^{\frac{1}{p}}\,,
\end{aligned}$$

where the final line uses definition of $\|\cdot\|_{\phi_s}$. Note all moments of m are finite under condition (11). Let $\|\mathbb{D}\|_{L^{\phi_s}}$ denote the operator norm of $\mathbb{D}$ on L^{ϕ_s}. It follows by Lemma B.1 and definition of the operator norm that

$$\|\mathbb{D}\|_{L^{\phi_s}} \leq \left(\left(2^{\frac{1}{q}}\mathbb{E}^{\mu\otimes Q}\left[m(X_t, X_{t+1})^p\right]^{\frac{1}{p}} - 1\right) \vee 1\right) q^{\frac{1}{s}}\beta < \infty\,.$$

One may similarly deduce that $\mathbb{D}$ maps E^{ϕ_s} into E^{ϕ_s}. Boundedness of $\mathbb{D}$ on E^{ϕ_s} now follows from $\|\mathbb{D}\|_{L^{\phi_s}} < \infty$ because E^{ϕ_s} is a closed linear subspace of L^{ϕ_s}.

We use Lemma B.6 to establish the spectral radius condition. We prove the result for the spaces L^{ϕ_s}; the results for E^{ϕ_s} follow because E^{ϕ_s} is a closed linear subspace of L^{ϕ_s}.

Suppose $s > 1$. Fix $p, q \in (1, \infty)$ with $p^{-1} + q^{-1} = 1$. For any $f \in L^{\phi_s}$ with $\|f\|_{\phi_s} > 0$, by two applications of Jensen's inequality we have

$$\begin{aligned}\mathbb{E}^\mu\left[e^{|\mathbb{D}^n f(X_t)/(q^{\frac{1}{s}}(\beta^{\frac{s-1}{s}})^n \|f\|_{\phi_s})|^s}\right] &= \mathbb{E}^\mu\left[e^{\beta^n q^{-1}|\tilde{\mathbb{E}}^n f(X_t)/\|f\|_{\phi_s}|^s}\right] \\ &\leq \mathbb{E}^\mu\left[e^{q^{-1}|\tilde{\mathbb{E}}^n f(X_t)/\|f\|_{\phi_s}|^s}\right]^{\beta^n} \leq \mathbb{E}^\mu\left[\tilde{\mathbb{E}}^n g(X_t)\right]^{\beta^n},\end{aligned}$$

where $g(x) = \exp(q^{-1}|f(x)/\|f\|_{\phi_s}|^s)$. By Hölder's inequality,

$$\begin{aligned}\mathbb{E}^\mu\left[\tilde{\mathbb{E}}^n g(X_t)\right] &= \mathbb{E}^{\mu\otimes Q}\left[m(X_t, X_{t+1})\cdots m(X_{t+n-1}, X_{t+n}) g(X_{t+n})\right] \\ &\leq \mathbb{E}^{\mu\otimes Q}\left[(m(X_t, X_{t+1})\cdots m(X_{t+n-1}, X_{t+n}))^p\right]^{\frac{1}{p}} \mathbb{E}^\mu\left[|g(X_t)|^q\right]^{\frac{1}{q}} \\ &\leq \mathbb{E}^{\mu\otimes Q}\left[m(X_t, X_{t+1})^{np}\right]^{\frac{1}{np}} \cdots \mathbb{E}^{\mu\otimes Q}\left[m(X_{t+n-1}, X_{t+n})^{np}\right]^{\frac{1}{np}} \mathbb{E}^\mu\left[|g(X_t)|^q\right]^{\frac{1}{q}} \\ &= \mathbb{E}^{\mu\otimes Q}\left[m(X_t, X_{t+1})^{np}\right]^{\frac{1}{p}} \mathbb{E}^\mu\left[|g(X_t)|^q\right]^{\frac{1}{q}},\end{aligned}$$

where we have slightly abused notation by letting $\mathbb{E}^{\mu\otimes Q}$ denote expectation with respect to the stationary distribution of $(X_t, \ldots, X_{t+n})$. It follows by Lemma B.6, and definition of g and $\|\cdot\|_{\phi_s}$ that

$$\mathbb{E}^\mu\left[\tilde{\mathbb{E}}^n g(X_t)\right] \leq \mathbb{E}^{\mu\otimes Q}\left[m(X_t, X_{t+1})^{np}\right]^{\frac{1}{p}} \mathbb{E}^\mu\left[e^{|f(X_t)/\|f\|_{\phi_s}|^s}\right]^{\frac{1}{q}} \leq 2^{\frac{1}{q}} C e^{(\beta+c)^{-n}}$$

for constants $C \in (0, \infty)$ and $c \in (0, 1-\beta)$ not depending on f. Therefore,

$$\mathbb{E}^\mu\left[e^{|\mathbb{D}^n f(X_t)/(q^{\frac{1}{s}}(\beta^{\frac{s-1}{s}})^n \|f\|_{\phi_s})|^s}\right] \leq \left(2^{\frac{1}{q}} C e^{(\beta+c)^{-n}}\right)^{\beta^n}.$$

It follows by Lemma B.1 and definition of the operator norm $\|\mathbb{D}^n\|_{L^{\phi_s}}$ that

$$\|\mathbb{D}^n\|_{L^{\phi_s}} \leq \left(\left(\left(2^{\frac{1}{q}} C e^{(\beta+c)^{-n}}\right)^{\beta^n} - 1\right) \vee 1\right) q^{\frac{1}{s}} (\beta^{\frac{s-1}{s}})^n$$

and therefore $\rho(\mathbb{D}; L^{\phi_s}) \equiv \lim_{n\to\infty} \|\mathbb{D}^n\|_{L^{\phi_s}}^{1/n} \leq \beta^{\frac{s-1}{s}} < 1$.

Now suppose $s = 1$. Fix $\varepsilon \in (0, 1)$ and note that $\beta < \beta + \varepsilon c < \beta + c < 1$ where c is as in

Lemma B.6. For any $f \in L^{\phi_1}$ with $\|f\|_{\phi_1} > 0$, we have:

$$\begin{aligned}\mathbb{E}^{\mu}\left[e^{|\mathbb{D}^n f(X_t)/(q\beta^n(\beta+\varepsilon c)^{-n}\|f\|_{\phi_1})|}\right] &= \mathbb{E}^{\mu}\left[e^{(\beta+\varepsilon c)^n q^{-1}|\tilde{\mathbb{E}}^n f(X_t)/\|f\|_{\phi_1}|}\right] \\ &\leq \mathbb{E}^{\mu}\left[e^{q^{-1}|\tilde{\mathbb{E}}^n f(X_t)/\|f\|_{\phi_1}|}\right]^{(\beta+\varepsilon c)^n} \\ &\leq \mathbb{E}^{\mu}\left[\tilde{\mathbb{E}}^n g(X_t)\right]^{(\beta+\varepsilon c)^n},\end{aligned}$$

where $g(x) = \exp(q^{-1}|f(x)|/\|f\|_{\phi_1})$. By similar arguments to above, we obtain

$$\mathbb{E}^{\mu}\left[e^{|\mathbb{D}^n f(X_t)/(q\beta^n(\beta+\varepsilon c)^{-n}\|f\|_{\phi_1})|}\right] \leq (2^{\frac{1}{q}} C e^{(\beta+c)^{-n}})^{(\beta+\varepsilon c)^n}.$$

By Lemma B.1 and definition of the operator norm $\|\mathbb{D}^n\|_{L^{\phi_1}}$, we may deduce that

$$\|\mathbb{D}^n\|_{L^{\phi_1}} \leq \left(\left((2^{\frac{1}{q}} C e^{(\beta+c)^{-n}})^{(\beta+\varepsilon c)^n} - 1\right) \vee 1\right) q \left(\frac{\beta}{\beta+\varepsilon c}\right)^n,$$

from which it follows similarly that $\rho(\mathbb{D}; L^{\phi_1}) \equiv \lim_{n\to\infty} \|\mathbb{D}^n\|_{L^{\phi_1}}^{1/n} \leq \frac{\beta}{\beta+\varepsilon c} < 1$. □

B.4 Proofs for Section 4

Proof of Theorem 4.1. We verify the conditions of Proposition 3.1. For existence, Lemma B.7 shows $\mathbb{T}$ is a continuous, monotone, and convex operator on E^{ϕ_s} for each $1 \leq s \leq r$. Let

$$\bar{v}(x) = (1-\beta)\sum_{n=0}^{\infty} \beta^{n+1} \log\left((\mathbb{E}^Q)^n h(x)\right),$$

where $h(x) = \mathbb{E}^Q[e^{\frac{\alpha}{1-\beta}u(X_t,X_{t+1})}|X_t = x]$. We first show that $\mathbb{E}^{\mu}[\exp(|\bar{v}(X_t)/(\beta c)|^r)] < \infty$ holds for each $c \in (0,1]$. By Jensen's inequality (using the fact that $\sum_{n=1}^{\infty}(1-\beta)\beta^n = 1$ and convexity of $x \mapsto e^{|x/c|^r}$ and $x \mapsto e^{|(\log x)/c|^r}$ for $c \in (0,1]$), we obtain

$$\begin{aligned}\mathbb{E}^{\mu}\left[e^{|\bar{v}(X_t)/(\beta c)|^r}\right] &= \mathbb{E}^{\mu}\left[\exp\left(\left|(1-\beta)\sum_{n=0}^{\infty}\beta^n \log\left((\mathbb{E}^Q)^n h(X_t)\right)/c\right|^r\right)\right] \\ &\leq (1-\beta)\sum_{n=0}^{\infty}\beta^n \mathbb{E}^{\mu}\left[\exp\left(\left|\log\left((\mathbb{E}^Q)^n h(x)\right)/c\right|^r\right)\right] \\ &\leq (1-\beta)\sum_{n=0}^{\infty}\beta^n \mathbb{E}^{\mu\otimes Q}\left[e^{|\frac{\alpha}{c(1-\beta)}u(X_{t+n},X_{t+n+1})|^r}\right] \\ &= \mathbb{E}^{\mu\otimes Q}\left[e^{|\frac{\alpha}{c(1-\beta)}u(X_{t+n},X_{t+n+1})|^r}\right] < \infty.\end{aligned}$$

It follows by Remark B.1 that $\bar{v} \in E^{\phi_r}$.

We now show that $\mathbb{T}\bar{v} \le \bar{v}$. By Holder's inequality we first have

$$\begin{aligned}\mathbb{T}\bar{v}(X_t) &\le \beta \log \left(\mathbb{E}^Q \left[e^{\bar{v}(X_{t+1})/\beta} \middle| X_t \right]^{\beta} \mathbb{E}^Q \left[e^{\frac{\alpha}{1-\beta} u(X_t, X_{t+1})} \middle| X_t \right]^{1-\beta} \right) \\ &= \beta^2 \log \mathbb{E}^Q [e^{\bar{v}(X_{t+1})/\beta} | X_t] + (1-\beta)\beta \log h(X_t) \,. \end{aligned} \tag{37}$$

By Lemma B.2, we may deduce

$$\begin{aligned}\log \mathbb{E}^Q \left[e^{\bar{v}(X_{t+1})/\beta} \middle| X_t \right] &= \log \mathbb{E}^Q \left[\prod_{n=0}^{\infty} \left((\mathbb{E}^Q)^n h(X_{t+1}) \right)^{(1-\beta)\beta^n} \middle| X_t \right] \\ &\le \log \left(\prod_{n=0}^{\infty} \mathbb{E}^Q \left[\left((\mathbb{E}^Q)^n h(X_{t+1}) \right) \middle| X_t \right]^{(1-\beta)\beta^n} \right) \\ &= (1-\beta) \sum_{n=1}^{\infty} \beta^{n-1} \log \left((\mathbb{E}^Q)^n h(X_t) \right) \,. \end{aligned} \tag{38}$$

Substituting (38) into (37) yields $\mathbb{T}\bar{v} \le \bar{v}$.

We now show $\{\mathbb{T}^n \bar{v}\}_{n\ge 1}$ is bounded from below, first observe that

$$\mathbb{T}f(x) = \beta \log \mathbb{E}^Q[e^{f(X_{t+1})+\alpha u(X_t,X_{t+1})} | X_t = x] \ge \beta \mathbb{E}^Q[f(X_{t+1}) + \alpha u(X_t, X_{t+1}) | X_t = x] \,.$$

Therefore,

$$\mathbb{T}^n \bar{v} \ge (\beta \mathbb{E}^Q)^n \bar{v} + \sum_{s=0}^{n-1} (\beta \mathbb{E}^Q)^s (h_1)$$

for each $n \ge 1$, where $h_1(x) = \beta \mathbb{E}^Q[\alpha u(X_t, X_{t+1}) | X_t = x]$. Note also that $\|\beta \mathbb{E}^Q\|_{\phi_r} = \beta$ and $\rho(\beta \mathbb{E}^Q; E^{\phi_r}) = \beta$ (see Section 3.3), and so we obtain $\liminf_{n\to\infty} \mathbb{T}^n \bar{v} \ge (\mathbb{I} - \beta \mathbb{E}^Q)^{-1} h_1 \in E^{\phi_r}$.

Uniqueness: v is a fixed point of $\mathbb{T} : E^{\phi_s} \to E^{\phi_s}$ for each $s \in [1, r]$. Moreover, $\mathbb{T} : E^{\phi_s} \to E^{\phi_s}$ is convex by Lemma B.7 and $\mathbb{D}_v$ is a bounded, monotone linear operator with $\rho(\mathbb{D}_v; E^{\phi_s}) < 1$ for $s \in [1, r]$ by Lemma B.8. Uniqueness in E^{ϕ_s} with $s \in (1, r]$ follows by Proposition 3.1(ii). That v is the smallest and unique stable fixed point in E^{ϕ_1} follows by Corollary 3.1. $\square$

Lemma B.7. *Let condition (14) hold. Then: $\mathbb{T}$ is a continuous, monotone and convex operator on E^{ϕ_s} for each $1 \le s \le r$.*

Proof of Lemma B.7. Fix any $1 \le s \le r$. Take any $f \in E^{\phi_s}$ and $c \in (0, 1]$. By convexity of

$x \mapsto e^{|(\log x)/c|^s}$ for $c \in (0,1]$ and Jensen's inequality:

$$\begin{aligned}
\mathbb{E}^\mu[\exp(|\mathbb{T}f(X_t)/(\beta c)|^s)] &= \mathbb{E}^\mu\left[\exp\left(\left|\frac{1}{c}\log\mathbb{E}^Q\left[e^{f(X_{t+1})+\alpha u(X_t,X_{t+1})}\middle| X_t\right]\right|^s\right)\right] \\
&\leq \mathbb{E}^\mu\left[\mathbb{E}^Q\left[\exp\left(\left|\frac{1}{c}\log e^{f(X_{t+1})+\alpha u(X_t,X_{t+1})}\right|^s\right)\middle| X_t\right]\right] \\
&= \mathbb{E}^{\mu\otimes Q}\left[\exp\left(\left|\frac{f(X_{t+1})+\alpha u(X_t,X_{t+1})}{c}\right|^s\right)\right] < \infty
\end{aligned}$$

which is finite for any $f \in E^{\phi_s}$ under (14). It follows by Remark B.1 that $\mathbb{T} : E^{\phi_s} \to E^{\phi_s}$.

Continuity: Fix any $f \in E^{\phi_s}$. Take $g \in E^{\phi_s}$ with $\|g\|_{\phi_s} \in (0, 2^{-1/s}]$ and set $c = 2^{1/s}\|g\|_{\phi_s}$. Let $\mathbb{E}_f$ denote the distorted conditional expectation operator from (15) with f in place of v. By convexity of $x \mapsto e^{|(\log x)/c|^s}$ for $c \in (0,1]$ and the Jensen and Cauchy-Schwarz inequalities,

$$\begin{aligned}
\mathbb{E}^\mu\left[\phi_s(|\mathbb{T}(f+g)(X_t) - \mathbb{T}f(X_t)|/(\beta c))\right] + 1 &= \mathbb{E}^\mu\left[\exp\left(\left|\frac{1}{c}\log\mathbb{E}_f\left[e^{g(X_{t+1})}\middle| X_t\right]\right|^s\right)\right] \\
&\leq \mathbb{E}^\mu\left[\mathbb{E}_f\left[\exp\left(\left|\frac{1}{c}\log e^{g(X_{t+1})}\right|^s\right)\middle| X_t\right]\right] \\
&= \mathbb{E}^{\mu\otimes Q}\left[m_f(X_t,X_{t+1})\exp\left(\left|\frac{g(X_{t+1})}{c}\right|^s\right)\right] \\
&\leq \mathbb{E}^\mu\left[e^{2|g(X_t)/c|^s}\right]^{1/2}\mathbb{E}^{\mu\otimes Q}[m_f(X_t,X_{t+1})^2]^{1/2} \\
&= \sqrt{2\mathbb{E}^{\mu\otimes Q}[m_f(X_t,X_{t+1})^2]}
\end{aligned}$$

because $c = 2^{1/s}\|g\|_{\phi_s}$. Finiteness of $\mathbb{E}^{\mu\otimes Q}[m_f(X_t,X_{t+1})^2]$ holds for any $f \in E^{\phi_s}$ under (14). To see this, by several applications of the Cauchy–Schwarz and Jensen inequalities, we have

$$\begin{aligned}
\mathbb{E}^{\mu\otimes Q}[m_f(X_t,X_{t+1})^2] &= \mathbb{E}^{\mu\otimes Q}\left[\left(\frac{e^{f(X_{t+1})+\alpha u(X_t,X_{t+1})}}{\mathbb{E}^Q[e^{f(X_{t+1})+\alpha u(X_t,X_{t+1})}|X_t]}\right)^2\right] \\
&\leq \mathbb{E}^{\mu\otimes Q}\left[e^{4|f(X_{t+1})+\alpha u(X_t,X_{t+1})|}\right] \\
&\leq \mathbb{E}^\mu\left[e^{8|f(X_t)|}\right]^{1/2}\mathbb{E}^{\mu\otimes Q}\left[e^{8|\alpha u(X_t,X_{t+1})|}\right]^{1/2},
\end{aligned}$$

which is finite for any $f \in E^{\phi_s}$ under (14). Continuity now follows by Lemma B.1. Monotonicity follows from monotonicity of $\exp(\cdot)$, $\log(\cdot)$, and conditional expectations. Convexity follows by applying Hölder's inequality to the conditional expectation

$$\mathbb{E}^Q\left[e^{\tau(v_1(X_{t+1})+\alpha u(X_t,X_{t+1}))+(1-\tau)(v_2(X_{t+1})+\alpha u(X_t,X_{t+1}))}\Big| X_t = x\right]$$

with $p = \tau^{-1}$ and $q = (1-\tau)^{-1}$. $\square$

Lemma B.8. *Let condition (14) hold with $r > 1$ and fix any $v \in E^{\phi_{r'}}$ with $r' > 1$. Then: for each $s \geq 1$, $\mathbb{D}_v$ is a continuous linear operator on E^{ϕ_s} with $\rho(\mathbb{D}_v; E^{\phi_s}) < 1$.*

Proof of Lemma B.8. We verify condition (11) from Lemma 3.1. The log change-of-measure is

$$\log m_v(X_t, X_{t+1}) = v(X_{t+1}) + \alpha u(X_t, X_{t+1}) - \log \mathbb{E}^Q[e^{v(X_{t+1})+\alpha u(X_t,X_{t+1})}|X_t]\,.$$

For any $v \in E^{\phi_{r'}}$ with $r' > 1$, setting $\underline{r} = (r \wedge r') > 1$ and taking any $c \in (0,1]$,

$$\mathbb{E}^\mu\left[e^{|\log \mathbb{E}^Q[e^{v(X_{t+1})+\alpha u(X_t,X_{t+1})}|X_t]/c|^{\underline{r}}}\right] \leq \mathbb{E}^{\mu\otimes Q}\left[e^{|(v(X_{t+1})+\alpha u(X_t,X_{t+1}))/c|^{\underline{r}}}\right]$$

by Jensen's inequality. The right-hand side is finite by condition (14). Therefore,

$$\mathbb{E}^{\mu\otimes Q}\left[e^{|\log m_v(X_t,X_{t+1})/c|^{\underline{r}}}\right] < \infty$$

for any $c \in (0,1]$ and hence for any $c > 0$ (see Remark B.1), verifying condition (11). $\square$

Lemma B.9. *Let $Y = |Z|$ with $Z \sim N(0,1)$. Then for $a > 0$ and $r \in [1,2)$, we have*

$$\mathbb{E}\left[\exp\left(\frac{Y^r}{a^r}\right)\right] \leq \frac{\sqrt{2}}{\sqrt{\pi}}\left(\left(\frac{2}{a^r}\right)^{\frac{1}{2-r}}\exp\left(\frac{2^{\frac{r}{2-r}}}{a^{\frac{2r}{2-r}}}\right) + \left(\frac{4}{a^r}\right)^{\frac{1}{2-r}} + \sqrt{\pi}\right).$$

Proof of Lemma B.9. First write

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\frac{Y^r}{a^r}\right)\right] &= \frac{\sqrt{2}}{\sqrt{\pi}}\int_0^\infty \exp\left(\frac{y^r}{a^r} - \frac{1}{2}y^2\right)\mathrm{d}y \\
&\leq \frac{\sqrt{2}}{\sqrt{\pi}}\left(\int_0^{(\frac{2}{a^r})^{\frac{1}{2-r}}} \exp\left(\frac{y^r}{a^r}\right)\mathrm{d}y + \int_{(\frac{2}{a^r})^{\frac{1}{2-r}}}^{(\frac{4}{a^r})^{\frac{1}{2-r}}} \mathrm{d}y + \int_{(\frac{4}{a^r})^{\frac{1}{2-r}}}^\infty \exp\left(-\frac{1}{4}y^2\right)\mathrm{d}y\right) \\
&\leq \frac{\sqrt{2}}{\sqrt{\pi}}\left(\left(\frac{2}{a^r}\right)^{\frac{1}{2-r}}\exp\left(\frac{2^{\frac{r}{2-r}}}{a^{\frac{2r}{2-r}}}\right) + \left(\frac{4}{a^r}\right)^{\frac{1}{2-r}} + \sqrt{\pi}\right).
\end{aligned}$$

The first inequality follows by noting that $\frac{y^r}{a^r} - \frac{1}{2}y^2 \leq \frac{y^r}{a^r}$ (for the first integral), $\frac{y^r}{a^r} - \frac{1}{2}y^2 \leq 0$ over $[(\frac{2}{a^r})^{\frac{1}{2-r}}, \infty)$ (for the second integral), and $\frac{y^r}{a^r} - \frac{1}{2}y^2 \leq -\frac{1}{4}y^2$ over $[(\frac{4}{a^r})^{\frac{1}{2-r}}, \infty)$ (for the third integral). For the three integrals on the second line, the first is bounded using the inequality $\int_0^b \exp(\frac{y^r}{a^r})dy \leq b\exp(\frac{b^r}{a^r})$ (valid for $b \geq 0$); the second and third are trivial. $\square$

Proof of Proposition 4.1. The boundedness condition in the statement of the lemma ensures $\mathbb{T}_{\mathcal{C}}$ is a self map on $B(\mathcal{C})$. It is then straightforward to verify that $\mathbb{T}_{\mathcal{C}}$ satisfies Blackwell's sufficient conditions, and therefore has a unique fixed point $v_{\mathcal{C}} \in B(\mathcal{C})$.

To relate v and $v_{\mathcal{C}}$, let $v_|$ denote the restriction of v to $\mathcal{C}$. Then for $x \in \mathcal{C}$, we have

$$\begin{aligned}
v(x) - v_{\mathcal{C}}(x) &= \beta \log \mathbb{E}^Q[e^{v(X_{t+1})+\alpha u(X_t,X_{t+1})}|X_t = x] - \mathbb{T}_{\mathcal{C}} v_{\mathcal{C}}(x) \\
&\geq \beta \log \mathbb{E}^Q[e^{v_|(X_{t+1})+\alpha u(X_t,X_{t+1})} 1\!\!1\{X_{t+1} \in \mathcal{C}\}|X_t = x] - \mathbb{T}_{\mathcal{C}} v_{\mathcal{C}}(x) \\
&= \beta \log Q(\mathcal{C}, x) + \mathbb{T}_{\mathcal{C}} v_|(x) - \mathbb{T}_{\mathcal{C}} v_{\mathcal{C}}(x) \\
&\geq \beta \log Q(\mathcal{C}, x) + \beta \mathbb{E}^Q \left[m_{\mathcal{C},v_{\mathcal{C}}}(X_t, X_{t+1})(v_|(X_{t+1}) - v_{\mathcal{C}}(X_{t+1})) \middle| X_t = x \right] \\
&\geq \beta \log Q(\mathcal{C}, x) + \beta \inf_{x \in \mathcal{C}} (v(x) - v_{\mathcal{C}}(x)) \ ,
\end{aligned}$$

where the first inequality is by monotonicity of expectations, the second equality is because $\inf_{x\in\mathcal{C}} Q(\mathcal{C}, x) > 0$, and the second inequality is by Jensen's inequality with

$$m_{\mathcal{C},v_{\mathcal{C}}}(X_t, X_{t+1}) = \frac{e^{v_{\mathcal{C}}(X_{t+1})+\alpha u(X_t,X_{t+1})} 1\!\!1\{X_{t+1} \in \mathcal{C}\}}{\mathbb{E}^Q[e^{v_{\mathcal{C}}(X_{t+1})+\alpha u(X_t,X_{t+1})} 1\!\!1\{X_{t+1} \in \mathcal{C}\}|X_t]} \,.$$

The result follows by taking the infimum of both sides with respect to $x \in \mathcal{C}$. $\square$

B.5 Proofs for Section 5

Recall $\hat{X}_t = (\hat{\xi}_t, \varphi_t)$. The conditional distribution $\hat{Q}$ of $(\xi_t, \hat{X}_{t+1})$ given $\hat{X}_t$ may be represented by

$$\mathbb{E}^{\hat{Q}}[h(\xi_t, \hat{X}_{t+1})|\hat{X}_t] = \mathbb{E}^{\hat{Q}}[h(\xi_t, \hat{X}_{t+1})|\hat{\xi}_t] = \mathbb{E}^{\Pi_\xi \otimes Q_\varphi}[h(\xi_t, \varphi_{t+1}, \Xi(\hat{\xi}_t, \varphi_{t+1}))|\hat{\xi}_t] \,.$$

Recall that μ is the stationary distribution of $\hat{X}_t$ under $\hat{Q}$. For $v \in E_{\hat{\xi}}^{\phi_1}$, define

$$m_v^{\Pi_\xi}(\xi_t, \hat{\xi}_t) = \frac{\mathbb{E}^{Q_\varphi}\left[e^{\frac{\theta}{\vartheta} v(\Xi(\hat{\xi}_t, \varphi_{t+1})) + \alpha u(\varphi_{t+1})} \middle| \xi_t, \hat{\xi}_t \right]^{\frac{\vartheta}{\theta}}}{\mathbb{E}^{\Pi_\xi}\left[\mathbb{E}^{Q_\varphi}\left[e^{\frac{\theta}{\vartheta} v(\Xi(\hat{\xi}_t, \varphi_{t+1})) + \alpha u(\varphi_{t+1})} \middle| \xi_t, \hat{\xi}_t \right]^{\frac{\vartheta}{\theta}} \middle| \hat{\xi}_t \right]}$$

$$m_v^{Q_\varphi}(\xi_t, \hat{\xi}_t, \varphi_{t+1}) = \frac{e^{\frac{\theta}{\vartheta} v(\Xi(\hat{\xi}_t, \varphi_{t+1})) + \alpha u(\varphi_{t+1})}}{\mathbb{E}^{Q_\varphi}\left[e^{\frac{\theta}{\vartheta} v(\Xi(\hat{\xi}_t, \varphi_{t+1})) + \alpha u(\varphi_{t+1})} \middle| \xi_t, \hat{\xi}_t \right]} .$$

The quantity $m_v^{\Pi_\xi}$ distorts the posterior distribution for ξ_t given $\hat{X}_t$ whereas $m_v^{Q_\varphi}$ distorts the conditional distribution Q_φ. To simplify notation, define the distorted conditional expectations $\mathbb{E}_v^{\Pi_\xi}$ and $\mathbb{E}_v^{Q_\varphi}$ by

$$\mathbb{E}_v^{\Pi_\xi} f(\hat{\xi}) = \mathbb{E}^{\Pi_\xi}\left[m_v^{\Pi_\xi}(\xi_t, \hat{\xi}_t) f(\xi_t, \hat{\xi}_t) \middle| \hat{\xi}_t = \hat{\xi} \right] ,$$

$$\mathbb{E}_v^{Q_\varphi} f(\xi, \hat{\xi}) = \mathbb{E}^{\Pi_\xi}\left[m_v^{Q_\varphi}(\xi_t, \hat{\xi}_t, \varphi_{t+1}) f(\xi_t, \hat{\xi}_t, \varphi_{t+1}) \middle| \xi_t = \xi, \hat{\xi}_t = \hat{\xi} \right] .$$

The subgradient of $\mathbb{T}$ at v is the composition of these two distorted conditional expectations, discounted by β:

$$\mathbb{D}_v f(\hat{\xi}) = \beta \mathbb{E}^{\hat{Q}}\left[m_v(\xi_t, \hat{\xi}_t, \varphi_{t+1}) f(\hat{\xi}_{t+1}) \middle| \hat{\xi}_t = \hat{\xi} \right] \tag{39}$$

where $m_v(\xi_t, \hat{\xi}_t, \varphi_{t+1}) = m_v^{\Pi_\xi}(\xi_t, \hat{\xi}_t) m_v^{Q_\varphi}(\xi_t, \hat{\xi}_t, \varphi_{t+1})$.

Proof of Theorem 5.1. We verify the conditions of Proposition 3.1. Lemma B.10 shows that $\mathbb{T}$ is a continuous, monotone, and convex operator on $E_{\hat{\xi}}^{\phi_s}$ for each $1 \leq s \leq r$. If $\theta < \vartheta$, let

$$\bar{v}(\hat{\xi}) = (1-\beta) \sum_{n=0}^{\infty} \beta^{n+1} \log \left(\left(\mathbb{E}^{\hat{Q}} \right)^{n+1} g_1(\hat{\xi}) \right) ,$$

where $g_1(\hat{X}_t) = \exp(\frac{\alpha\vartheta}{(1-\beta)\theta} u(\varphi_t))$. For any $c > 0$, by Jensen's inequality we may deduce

$$\mathbb{E}^{\mu}[e^{|\bar{v}(\hat{\xi}_t)/(\beta c)|^r}] \leq (1-\beta) \sum_{n=0}^{\infty} \beta^n \mathbb{E}^{\mu} \left[\left(\left(\mathbb{E}^{\hat{Q}} \right)^{n+1} g_1^r(\hat{\xi}_t) \right) \right] ,$$

where $g_1^r(\hat{X}_t) = \exp(|\frac{\alpha\vartheta}{(1-\beta)\theta c} u(\varphi_t)|^r)$. As $u \in E_\varphi^{\phi_r}$, the right-hand side of the preceding display is finite and so $\bar{v} \in E_{\hat{\xi}}^{\phi_r}$.

To show $\mathbb{T}\bar{v} \leq \bar{v}$, first by the Jensen and Hölder inequalities,

$$
\begin{aligned}
\mathbb{T}\bar{v}(\hat{\xi}) &= \beta \log \mathbb{E}^{\Pi_\xi}\left[\mathbb{E}^{Q_\varphi}\left[e^{\frac{\theta}{\vartheta}\bar{v}(\Xi(\hat{\xi}_t,\varphi_{t+1}))+\alpha u(\varphi_{t+1})} \Big| \xi_t, \hat{\xi}_t \right]^{\vartheta/\theta} \Big| \hat{\xi}_t = \hat{\xi} \right] \\
&\leq \beta \log \mathbb{E}^{\hat{Q}}\left[e^{\bar{v}(\hat{\xi}_{t+1})+\alpha\frac{\vartheta}{\theta}u(\varphi_{t+1})} \Big| \hat{\xi}_t = \hat{\xi} \right] \\
&\leq \beta^2 \log \mathbb{E}^{\hat{Q}}\left[e^{\bar{v}(\hat{\xi}_{t+1})/\beta} \Big| \hat{\xi}_t = \hat{\xi} \right] + \beta(1-\beta)\log \mathbb{E}^{\hat{Q}}\left[e^{\frac{\alpha\vartheta}{(1-\beta)\theta}u(\varphi_{t+1})} \Big| \hat{\xi}_t = \hat{\xi} \right] .
\end{aligned}
$$

By Lemma B.2, we may deduce

$$
\log \mathbb{E}^{\hat{Q}}\left[e^{\bar{v}(\hat{\xi}_{t+1})/\beta} \Big| \hat{\xi}_t = \hat{\xi} \right] \leq (1-\beta)\sum_{n=1}^{\infty} \beta^{n-1} \log\left(\left(\mathbb{E}^{\hat{Q}}\right)^{n+1} g_1(\hat{\xi}) \right) ,
$$

hence $\mathbb{T}\bar{v} \leq \bar{v}$.

On the other hand, if $\vartheta \leq \theta$, let $\bar{v}(\hat{\xi}) = \frac{\vartheta}{\theta}(1-\beta)\sum_{n=0}^{\infty}\beta^{n+1}\log((\mathbb{E}^{\hat{Q}})^{n+1}g_2(\hat{\xi}))$ where $g_2(\hat{X}_t) = e^{\frac{\alpha}{1-\beta}u(\varphi_t)}$. By similar arguments to above, we may use the condition $u \in E^{\phi_r}_{\varphi}$ to deduce $\bar{v} \in E^{\phi_r}_{\hat{\xi}}$. Again by the Jensen and Hölder inequalities,

$$
\begin{aligned}
\mathbb{T}\bar{v}(\hat{\xi}) &= \beta \log \mathbb{E}^{\Pi_\xi}\left[\mathbb{E}^{Q_\varphi}\left[e^{\frac{\theta}{\vartheta}\bar{v}(\Xi(\hat{\xi}_t,\varphi_{t+1}))+\alpha u(\varphi_{t+1})} \Big| \hat{\xi}_t, \xi_t \right]^{\frac{\vartheta}{\theta}} \Big| \hat{\xi}_t = \hat{\xi} \right] \\
&\leq \frac{\vartheta}{\theta}\beta \log \mathbb{E}^{\hat{Q}}\left[e^{\frac{\theta}{\vartheta}\bar{v}(\hat{\xi}_{t+1})+\alpha u(\varphi_{t+1})} \Big| \hat{\xi}_t = \hat{\xi} \right] \\
&\leq \frac{\vartheta}{\theta}\beta^2 \log \mathbb{E}^{\hat{Q}}\left[e^{\frac{\theta}{\vartheta}\bar{v}(\hat{\xi}_{t+1})/\beta} \Big| \hat{\xi}_t = \hat{\xi} \right] + \frac{\vartheta}{\theta}\beta(1-\beta)\log \mathbb{E}^{\hat{Q}}\left[e^{\frac{\alpha}{1-\beta}u(\varphi_{t+1})} \Big| \hat{\xi}_t = \hat{\xi} \right] .
\end{aligned}
$$

The inequality $\mathbb{T}\bar{v} \leq \bar{v}$ now follows by similar arguments to the previous case.

To show that the sequence of iterates $\mathbb{T}^n\bar{v}$ is bounded from below, first note that for any $f \in E^{\phi_r}_{\hat{\xi}}$, we have

$$
\mathbb{T}f(\hat{\xi}) \geq \beta\mathbb{E}^{\hat{Q}}\left[f(\hat{\xi}_{t+1}) + \alpha\frac{\vartheta}{\theta}u(\varphi_{t+1}) \Big| \hat{\xi}_t = \hat{\xi} \right]
$$

which follows by several applications of Jensen's inequality. It follows that

$$
\mathbb{T}^n\bar{v}(\hat{\xi}) \geq \left(\beta\mathbb{E}^{\hat{Q}}\right)^n \bar{v}(\hat{\xi}) + \sum_{i=0}^{n-1}\left(\beta\mathbb{E}^{\hat{Q}}\right)^i g_3(\hat{\xi})
$$

where $g_3(\hat{\xi}) = \beta\mathbb{E}^{\hat{Q}}[\alpha\frac{\vartheta}{\theta}u(\varphi_{t+1})|\hat{\xi}_t = \hat{\xi}] \in E^{\phi_r}_{\hat{\xi}}$. Note also that $\rho(\beta\mathbb{E}^{\hat{Q}}; E^{\phi_r}) = \beta$ (see Section 3.3), hence $\liminf_{n\to\infty}\mathbb{T}^n\bar{v} \geq (\mathbb{I}-\beta\mathbb{E}^{\hat{Q}})^{-1}g_3 \in E^{\phi_r}$. This completes the proof of existence.

For uniqueness, v is necessarily a fixed point of $\mathbb{T} : E^{\phi_s}_{\hat{\xi}} \to E^{\phi_s}_{\hat{\xi}}$ for each $1 \leq s \leq r$.

The subgradient $\mathbb{D}_v$ is monotone. Lemma B.11 shows $\mathbb{D}_v : E^{\phi_s}_{\hat{\xi}} \to E^{\phi_s}_{\hat{\xi}}$ is bounded and $\rho(\mathbb{D}_v; E^{\phi_s}_{\hat{\xi}}) < 1$ for $s \in [1, r]$. Uniqueness follows by Proposition 3.1(ii) and Corollary 3.1. $\square$

Lemma B.10. *Let condition (20) hold. Then: $\mathbb{T}$ is a continuous, monotone, and convex operator on $E^{\phi_s}_{\hat{\xi}}$ for each $1 \le s \le r$.*

Proof of Lemma B.10. Fix $s \in [1, r]$. We first show $\mathbb{E}^\mu[\exp(|\mathbb{T}f(\hat{\xi}_t)/(\beta c)|^s)] < \infty$ holds for each $f \in E^{\phi_s}_{\hat{\xi}}$ and $c \in (0, \frac{\vartheta}{\theta} \wedge 1]$. By convexity of $x \mapsto e^{|(\log x)/c|^s}$ for $c \in (0, 1]$ and Jensen's inequality,

$$
\begin{aligned}
\mathbb{E}^\mu\left[\exp\left(\left|\frac{\mathbb{T}f(\hat{\xi}_t)}{\beta c}\right|^s\right)\right] &= \mathbb{E}^\mu\left[\exp\left(\frac{1}{c^s}\left|\log \mathbb{E}^{\Pi_\xi}\left[\mathbb{E}^{Q_\varphi}\left[e^{\frac{\theta}{\vartheta}f(\Xi(\hat{\xi}_t,\varphi_{t+1}))+\alpha u(\varphi_{t+1})}\middle|\, \xi_t, \hat{\xi}_t\right]^{\frac{\vartheta}{\theta}}\middle|\, \hat{\xi}_t\right]\right|^s\right)\right] \\
&\le \mathbb{E}^\mu\left[\mathbb{E}^{\Pi_\xi}\left[\exp\left(\frac{1}{c^s}\left|\log \mathbb{E}^{Q_\varphi}\left[e^{\frac{\theta}{\vartheta}f(\Xi(\hat{\xi}_t,\varphi_{t+1}))+\alpha u(\varphi_{t+1})}\middle|\, \xi_t, \hat{\xi}_t\right]^{\frac{\vartheta}{\theta}}\right|^s\right)\middle|\, \hat{\xi}_t\right]\right] \\
&\le \mathbb{E}^\mu\left[\mathbb{E}^{\Pi_\xi}\left[\mathbb{E}^{Q_\varphi}\left[\exp\left(\frac{1}{c^s}\left|\frac{\vartheta}{\theta}\log e^{\frac{\theta}{\vartheta}f(\Xi(\hat{\xi}_t,\varphi_{t+1}))+\alpha u(\varphi_{t+1})}\right|^s\right)\middle|\, \xi_t, \hat{\xi}_t\right]\middle|\, \hat{\xi}_t\right]\right] \\
&= \mathbb{E}^{\mu\otimes\Pi_\xi\otimes Q_\varphi}\left[\exp\left(\frac{1}{c^s}\left|f(\Xi(\hat{\xi}_t,\varphi_{t+1})) + \frac{\vartheta}{c\theta}\alpha u(\varphi_{t+1})\right|^s\right)\right]
\end{aligned}
$$

which is finite because $f \in E^{\phi_s}_{\hat{\xi}}$ and $u \in E^{\phi_r}_{\varphi}$. It follows by Remark B.1 that $\mathbb{T} : E^{\phi_s}_{\hat{\xi}} \to E^{\phi_s}_{\hat{\xi}}$.

For continuity, fix $f \in E^{\phi_s}_{\hat{\xi}}$. Take $g \in E^{\phi_s}_{\hat{\xi}}$ with $0 < \|g\|_{\phi_s} \le 2^{-1/s}(1 \wedge \frac{\vartheta}{\theta})$ and set $c = 2^{1/s}\|g\|_{\phi_s}$. Note

$$
\mathbb{T}(f+g)(\hat{\xi}) - \mathbb{T}f(\hat{\xi}) = \beta \log\left(\mathbb{E}^{\Pi_\xi}_f\left[\mathbb{E}^{Q_\varphi}_f\left[e^{\frac{\theta}{\vartheta}g(\Xi(\hat{\xi}_t,\varphi_{t+1}))}\middle|\, \xi_t, \hat{\xi}_t\right]^{\frac{\vartheta}{\theta}}\middle|\, \hat{\xi}_t = \hat{\xi}\right]\right).
$$

By similar arguments to the above, we may deduce

$$
\begin{aligned}
\mathbb{E}^\mu\left[\exp\left(\left|\frac{\mathbb{T}(f+g)(\hat{\xi}_t) - \mathbb{T}f(\hat{\xi}_t)}{\beta c}\right|^s\right)\right] &\le \mathbb{E}^\mu\left[\mathbb{E}^{\Pi_\xi}_f\left[\mathbb{E}^{Q_\varphi}_f\left[\exp\left(\left|\frac{1}{c}g(\Xi(\hat{\xi}_t,\varphi_{t+1}))\right|^s\right)\middle|\, \xi_t, \hat{\xi}_t\right]\middle|\, \hat{\xi}_t\right]\right] \\
&= \mathbb{E}^\mu\left[\mathbb{E}^{\hat{Q}}\left[m_f(\xi_t, \hat{\xi}_t, \varphi_{t+1})\exp\left(\left|\frac{1}{c}g(\Xi(\hat{\xi}_t,\varphi_{t+1}))\right|^s\right)\middle|\, \hat{\xi}_t\right]\right] \\
&\le \mathbb{E}^{\mu\otimes\hat{Q}}\left[m_f(\xi_t, \hat{\xi}_t, \varphi_{t+1})^2\right]^{1/2} \mathbb{E}^\mu\left[\exp(2|g(\hat{\xi}_{t+1})/c|^s\right]^{1/2} \\
&\le \left(2\mathbb{E}^{\mu\otimes\hat{Q}}\left[m_f(\xi_t, \hat{\xi}_t, \varphi_{t+1})^2\right]\right)^{1/2},
\end{aligned}
$$

because $c = 2^{1/s}\|g\|_{\phi_s}$. The expectation on the right-hand side is finite because $f \in E^{\phi_s}_{\hat{\xi}}$ and $u \in E^{\phi_r}_{\varphi}$. It follows by Lemma B.1 that $\|\mathbb{T}(f+g) - \mathbb{T}f\|_{\phi_s} \to 0$ as $\|g\|_{\phi_s} \to 0$.

Finally, monotonicity follows from monotonicity of the exponential and logarithm functions and monotonicity of conditional expectations. Convexity follows by Hölder's inequality. $\square$

Lemma B.11. *Let condition (20) hold. Fix any $v \in E^{\phi_{r'}}_{\hat{\xi}}$ with $r' > 1$. Then: for each $s \geq 1$, $\mathbb{D}_v$ is a continuous linear operator on $E^{\phi_s}_{\hat{\xi}}$ with $\rho(\mathbb{D}_v; E^{\phi_s}_{\hat{\xi}}) < 1$.*

Proof of Lemma B.11. It suffices to verify the conditions of Lemma 3.1. By iterated expectations, we may rewrite the subgradient from (39) as

$$\mathbb{D}_v f(\hat{\xi}) = \beta \mathbb{E}^{\hat{Q}} \left[\bar{m}_v(\hat{\xi}_t, \hat{\xi}_{t+1}) f(\hat{\xi}_{t+1}) \middle| \hat{\xi}_t = \hat{\xi} \right]$$

where $\bar{m}_v(\hat{\xi}_t, \hat{\xi}_{t+1})$ denotes the conditional expectation of $m_v(\xi_t, \hat{\xi}_t, \varphi_{t+1})$ given $\hat{\xi}_t, \hat{\xi}_{t+1}$ under $\hat{Q}$. The thin-tail condition on m_v then follows by similar arguments to the proof of Lemma B.8 for any $v \in E^{\phi_{r'}}_{\hat{\xi}}$ with $r' > 1$. $\square$

B.6 Proof for Section 6

Proof of Theorem 6.1. In view of the discussion preceding Theorem 6.1 and Lemma B.12, it suffices to show that $\bar{v} \in \tilde{E}^{\phi_r}$ and that $\mathbb{T}\bar{v} \leq \bar{v}$. By (25), convexity of $x \mapsto e^{|(\log x)/c|^r}$ for $c \in (0, 1]$, and two applications of Jensen's inequality, for any $c \in (0, 1]$ we have

$$\begin{aligned}
\mathbb{E}^{\tilde{\mu}} \left[e^{|\bar{v}(X_t)/c|^r} \right] &= \mathbb{E}^{\tilde{\mu}} \left[e^{\left| \log \left((1-\beta) \sum_{n=0}^{\infty} (\beta \lambda^{\frac{1}{\kappa}})^n \tilde{\mathbb{E}}^n (\iota^{-\frac{1}{\kappa}})(X_t) \right) / c \right|^r} \right] \\
&\leq (1 - \beta \lambda^{\frac{1}{\kappa}}) \sum_{n=0}^{\infty} (\beta \lambda^{\frac{1}{\kappa}})^n \mathbb{E}^{\tilde{\mu}} \left[\tilde{\mathbb{E}}^n \exp \left(\left| \log \left((1-\beta)(1-\beta\lambda^{\frac{1}{\kappa}})^{-1} (\iota(X_t))^{-\frac{1}{\kappa}} \right) / c \right|^r \right) \right] \\
&= \mathbb{E}^{\tilde{\mu}} \left[\exp \left(\left| \log \left((1-\beta)(1-\beta\lambda^{\frac{1}{\kappa}})^{-1} (\iota(X_t))^{-\frac{1}{\kappa}} \right) / c \right|^r \right) \right] < \infty
\end{aligned}$$

by condition (24), with the final equality because $\tilde{\mu}$ is the stationary distribution corresponding to $\tilde{\mathbb{E}}$. It follows by Remark B.1 that $\bar{v} \in E^{\phi_r}$.

To see that $\mathbb{T}\bar{v} \leq \bar{v}$, first note by Jensen's inequality that $\mathbb{E}[Z^\kappa]^{1/\kappa} \leq \mathbb{E}[Z]$ holds when $\kappa < 0$

for any random variable Z that is (strictly) positive with probability 1. Therefore,

$$\begin{aligned}
\mathbb{T}\bar{v}(x) &= \log\left((1-\beta)\iota(x)^{-\frac{1}{\kappa}} + \beta\lambda^{\frac{1}{\kappa}}\tilde{\mathbb{E}}\left[\left((1-\beta)\sum_{n=0}^{\infty}(\beta\lambda^{\frac{1}{\kappa}})^n\tilde{\mathbb{E}}^n(\iota^{-\frac{1}{\kappa}})(X_{t+1})\right)^{\kappa}\middle| X_t = x\right]^{\frac{1}{\kappa}}\right) \\
&\leq \log\left((1-\beta)\iota(x)^{-\frac{1}{\kappa}} + \beta\lambda^{\frac{1}{\kappa}}\tilde{\mathbb{E}}\left[(1-\beta)\sum_{n=0}^{\infty}(\beta\lambda^{\frac{1}{\kappa}})^n\tilde{\mathbb{E}}^n(\iota^{-\frac{1}{\kappa}})(X_{t+1})\middle| X_t = x\right]\right) \\
&= \log\left((1-\beta)\iota(x)^{-\frac{1}{\kappa}} + (1-\beta)\sum_{n=1}^{\infty}(\beta\lambda^{\frac{1}{\kappa}})^n\tilde{\mathbb{E}}^n(\iota^{-\frac{1}{\kappa}})(x)\right) = \bar{v}(x)\,.
\end{aligned}$$

Existence now follows by Proposition 3.1(i). □

Lemma B.12. *Let condition (24) hold. Then for any $\kappa \neq 0$, the operator $\mathbb{T}$ from (23) is a continuous, monotone operator on $\tilde{E}^{\phi_s}$ for each $1 \leq s \leq r$.*

Proof of Lemma B.12. Fix any $s \in [1, r]$. We first show that $\mathbb{E}^{\mu}[e^{|\mathbb{T}f(X_t)/c|^s}] < \infty$ holds for any $f \in \tilde{E}^{\phi_s}$ and c sufficiently small. By convexity of $x \mapsto e^{|(\log x)/c|^s}$ for $c \in (0,1]$ and two applications of Jensen's inequality and iterated expectations, for any $c \in (0, 1 \wedge |\kappa|^{-1}]$ we obtain

$$\begin{aligned}
\mathbb{E}^{\tilde{\mu}}\left[e^{|\mathbb{T}f(X_t)/c|^s}\right] &= \mathbb{E}^{\tilde{\mu}}\left[\exp\left(\left|\log\left((1-\beta)\iota(X_t)^{-\frac{1}{\kappa}} + \beta\lambda^{\frac{1}{\kappa}}\tilde{\mathbb{E}}[e^{\kappa f(X_{t+1})}|X_t]^{\frac{1}{\kappa}}\right)/c\right|^s\right)\right] \\
&\leq \mathbb{E}^{\tilde{\mu}}\left[(1-\beta)e^{|\log\iota(X_t)/(\kappa c)|^s} + \beta\exp\left(\left|\log\left(\tilde{\mathbb{E}}[\lambda e^{\kappa f(X_{t+1})}|X_t]\right)/(\kappa c)\right|^s\right)\right] \\
&\leq \mathbb{E}^{\tilde{\mu}}\left[(1-\beta)e^{|\log\iota(X_t)/(\kappa c)|^s} + \beta\tilde{\mathbb{E}}\left[\exp\left(\left|\log\left(\lambda e^{\kappa f(X_{t+1})}\right)/(\kappa c)\right|^s\right)\middle| X_t\right]\right] \\
&= (1-\beta)\mathbb{E}^{\tilde{\mu}}\left[e^{|\log\iota(X_t)/(\kappa c)|^s}\right] + \beta\mathbb{E}^{\tilde{\mu}}\left[e^{|(\log\lambda)/(\kappa c) + f(X_t)/c|^s}\right],
\end{aligned}$$

where the right-hand side is finite under condition (24), and the final equality is because $\tilde{\mu}$ is the stationary distribution under $\tilde{\mathbb{E}}$. It follows by Remark B.1 that $\mathbb{T} : \tilde{E}^{\phi_s} \to \tilde{E}^{\phi_s}$.

For continuity, fix $f \in \tilde{E}^{\phi_s}$ and take any $h \in \tilde{E}^{\phi_s}$ for which $\|h\|_{\phi_s}$ is sufficiently small in a sense we make precise below (the norm should be understood to be defined relative to the measure $\tilde{\mu}$). Then

$$\mathbb{T}(f+h)(x) - \mathbb{T}f(x) = \log\left\{\frac{(1-\beta)\iota(x)^{-\frac{1}{\kappa}} + \beta\lambda^{\frac{1}{\kappa}}w(x)\tilde{\mathbb{E}}_f[e^{\kappa h(X_{t+1})}|X_t = x]^{\frac{1}{\kappa}}}{(1-\beta)\iota(x)^{-\frac{1}{\kappa}} + \beta\lambda^{\frac{1}{\kappa}}w(x)}\right\}$$

where $w(x) = \tilde{\mathbb{E}}\left[e^{\kappa f(X_{t+1})}\middle| X_t = x\right]^{1/\kappa}$ and $\tilde{\mathbb{E}}_f$ denotes the distorted conditional expectation operator $\tilde{\mathbb{E}}_f g(x) := \tilde{\mathbb{E}}[m_f(X_t, X_{t+1}) g(X_{t+1}) | X_t = x]$ where

$$m_f(X_t, X_{t+1}) = \frac{e^{\kappa f(X_{t+1})}}{\tilde{\mathbb{E}}[e^{\kappa f(X_{t+1})} | X_t]} \,.$$

Take any $c \in (0, 1 \wedge |\kappa|^{-1}]$. By convexity of $x \mapsto e^{|(\log x)/c|^s}$ for $c \in (0,1]$, two applications of Jensen's inequality, and the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}
&\mathbb{E}^{\tilde{\mu}}\left[e^{|(\mathbb{T}(f+h)(X_t) - \mathbb{T}f(X_t))/c|^s}\right] \\
&= \mathbb{E}^{\tilde{\mu}}\left[\exp\left(\left|\frac{1}{c}\log\left\{\frac{(1-\beta)\iota(X_t)^{-\frac{1}{\kappa}} + \beta\lambda^{\frac{1}{\kappa}} w(X_t) \tilde{\mathbb{E}}_f\left[e^{\kappa h(X_{t+1})}\middle| X_t\right]^{\frac{1}{\kappa}}}{(1-\beta)\iota(X_t)^{-\frac{1}{\kappa}} + \beta\lambda^{\frac{1}{\kappa}} w(X_t)}\right\}\right|^s\right)\right] \\
&\leq \mathbb{E}^{\tilde{\mu}}\left[\frac{(1-\beta)\iota(X_t)^{-\frac{1}{\kappa}} + \beta\lambda^{\frac{1}{\kappa}} w(X_t) e^{\left|\frac{1}{c\kappa}\log \tilde{\mathbb{E}}_f\left[e^{\kappa h(X_{t+1})}|X_t\right]\right|^s}}{(1-\beta)\iota(X_t)^{-\frac{1}{\kappa}} + \beta\lambda^{\frac{1}{\kappa}} w(X_t)}\right] \\
&\leq \mathbb{E}^{\tilde{\mu}}\left[e^{\left|\frac{1}{c\kappa}\log \tilde{\mathbb{E}}_f\left[e^{\kappa h(X_{t+1})}|X_t\right]\right|^s}\right] \\
&\leq \mathbb{E}^{\tilde{\mu}}\left[\tilde{\mathbb{E}}_f\left[e^{|h(X_{t+1})/c|^s}\middle| X_t\right]\right] \\
&\leq \mathbb{E}^{\tilde{\mu}}\left[e^{4|\kappa f(X_t)|}\right]^{\frac{1}{2}} \mathbb{E}^{\tilde{\mu}}\left[e^{2|h(X_t)/c|^s}\right]^{\frac{1}{2}} .
\end{aligned}$$

For $h \in E^{\phi_s}$ with $\|h\|_{\phi_s} \leq 2^{-1/s}(1 \wedge |\kappa|^{-1})$, setting $c = 2^{1/s}\|h\|_{\phi_s}$ we therefore have

$$\mathbb{E}^{\mu}\left[e^{|(\mathbb{T}(f+h)(X_t) - \mathbb{T}f(X_t))/(2\|h\|_{\phi_s})|^s}\right] \leq \left(2\mathbb{E}^{\tilde{\mu}}\left[e^{4|\kappa f(X_t)|}\right]\right)^{\frac{1}{2}} .$$

Continuity now follows by Lemma B.1. Monotonicity of $\mathbb{T}$ follows form monotonicity of conditional expectations and monotonicity of the log and exp functions. □

Proof of Corollary 6.1. Immediate from Theorem 6.1 and Lemma B.3. □

References

Aliprantis, C. D. and K. C. Border (1999). *Infinite Dimensional Analysis: A Hitchhiker's Guide* (2 ed.). Springer.

Alvarez, F. and U. J. Jermann (2005). Using asset prices to measure the persistence of the marginal utility of wealth. *Econometrica 73*(6), 1977–2016.

Amann, H. (1976). Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces. *SIAM Review 18*(4), 620–709.

Ang, A. and A. Timmermann (2012). Regime changes and financial markets. *Annual Review of Financial Economics 4*(1), 313–337.

Backus, D., M. Chernov, and S. Zin (2014). Sources of entropy in representative agent models. *Journal of Finance 69*(1), 51–99.

Balbus, L. (2020). On recursive utilities with non-affine aggregator and conditional certainty equivalent. *Economic Theory 70*, 551?577.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014). Volatility, the macroeconomy, and asset prices. *The Journal of Finance 69*(6), 2471–2511.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance 59*(4), 1481–1509.

Barillas, F., L. P. Hansen, and T. J. Sargent (2009). Doubts or variability? *Journal of Economic Theory 144*(6), 2388 – 2418.

Becker, R. A. and J. P. Rincon-Zapatero (2017). Recursive utiity and thompson aggregators. *CAEPR Working Paper 2017-007*.

Bidder, R. and M. Smith (2018). Doubts and variability: A robust perspective on exotic consumption series. *Journal of Economic Theory 175*, 689–712.

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

Borovička, J. and J. Stachurski (2020). Necessary and sufficient conditions for existence and uniqueness of recursive utilities. *The Journal of Finance 75*(3), 1457–1493.

Boyd, J. H. (1990). Recursive utility and the Ramsey problem. *Journal of Economic Theory 50*(2), 326–345.

Cecchetti, S. G., P.-S. Lam, and N. C. Mark (1990). Mean reversion in equilibrium asset prices. *The American Economic Review 80*(3), 398–418.

Cecchetti, S. G., P.-S. Lam, and N. C. Mark (2000). Asset pricing with distorted beliefs: Are equity returns too good to be true? *American Economic Review 90*(4), 787–805.

Chen, W., L. B. Jia, and Y. Jiao (2016). Hölder's inequalities involving the infinite product and their applications in martingale spaces. *Analysis Mathematica 42*(2), 121–141.

Christensen, T. M. (2015). Nonparametric identification of positive eigenfunctions. *Econometric Theory 31*(6), 1310–1330.

Christensen, T. M. (2017). Nonparametric stochastic discount factor decomposition. *Econometrica 85*(5), 1501–1536.

Collard, F., S. Mukerji, K. Sheppard, and J.-M. Tallon (2018). Ambiguity and the historical equity premium. *Quantitative Economics 9*(2), 945–993.

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2016). Parameter learning in general equilibrium: The asset pricing implications. *American Economic Review 106*(3), 664–698.

Croce, M. M., M. Lettau, and S. C. Ludvigson (2015). Investor information, long-run risk, and the term structure of equity. *The Review of Financial Studies 28*(3), 706–742.

Durán, J. (2003). Discounting long run average growth in stochastic dynamic programs. *Economic Theory 22*(2), 395–413.

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

Guo, J. and X. D. He (2017). Recursive utility with investment gains and losses: Existence, uniqueness, and convergence. SSRN working paper no. 2790768.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica 57*(2), 357–384.

Hamilton, J. D. (1994). State-space models. Volume 4 of *Handbook of Econometrics*, Chapter 50, pp. 3039 – 3080. Elsevier.

Hansen, L. P., J. C. Heaton, and N. Li (2008). Consumption strikes back? Measuring long-run risk. *Journal of Political Economy 116*, 260–302.

Hansen, L. P. and T. J. Sargent (1995). Discounted linear exponential quadratic gaussian control. *IEEE Transactions on Automatic Control 40*(5), 968–971.

Hansen, L. P. and T. J. Sargent (2001). Robust control and model uncertainty. *The American Economic Review 91*(2), 60–66.

Hansen, L. P. and T. J. Sargent (2007). Recursive robust estimation and control without commitment. *Journal of Economic Theory 136*(1), 1–27.

Hansen, L. P. and T. J. Sargent (2010). Fragile beliefs and the price of uncertainty. *Quantitative Economics 1*(1), 129–162.

Hansen, L. P., T. J. Sargent, G. Turmuhambetova, and N. Williams (2006). Robust control and model misspecification. *Journal of Economic Theory 128*(1), 45–90.

Hansen, L. P. and J. A. Scheinkman (2009). Long-term risk: An operator approach. *Econometrica 77*(1), 177–234.

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.

Hindy, A. and C. Huang (1992). Intertemporal preferences for uncertain consumption: A continuous time approach. *Econometrica 60*(4), 781–801.

Hindy, A., C. Huang, and D. Kreps (1992). On intertemporal preferences in continuous time: The case of certainty. *Journal of Mathematical Economics 21*(5), 401 – 440.

Ju, N. and J. Miao (2012). Ambiguity, learning, and asset returns. *Econometrica 80*(2), 559–591.

Karakostas, G. L. (2008). An extension of Hölder's inequality and some results on infinite products. *Indian Journal of Mathematics 50*(2), 303–307.

Klibanoff, P., M. Marinacci, and S. Mukerji (2009). Recursive smooth ambiguity preferences. *Journal of Economic Theory 144*(3), 930–976.

Krasnosel'skii, M. A. (1964). *Positive Solution of Operator Equations.* P. Noordhoff, Groningen.

Krasnosel'skii, M. A. and Ya. B. Rutickii (1961). *Convex Functions and Orlicz Spaces.* P. Noordhoff Ltd., Groningen.

Kress, R. (2014). *Linear Integral Equations.* Springer.

Le Van, C. and Y. Vailakis (2005). Recursive utility and optimal growth with bounded or unbounded returns. *Journal of Economic Theory 123*(2), 187–209.

Malliavin, P. (1995). *Integration and Probability.* Springer.

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

Pollard, D. (2015). Mini-empirical. *Manuscript. http://www.stat.yale.edu/˜pollard/Books/Mini/.*

Ren, G. and J. Stachurski (2020). Dynamic programming with recursive preferences: Optimality and applications. *arXiv preprint arXiv:1812.05748 [econ.GN]*.

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

van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag.

Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge University Press.

Wachter, J. A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility? *The Journal of Finance 68*(3), 987–1035.