Existence and uniqueness of recursive utilities 
without boundedness*

Timothy M. Christensen†

December 23, 2020.

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, D84, 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 or local contractions in the usual sense. In this paper, we derive primitive, easily verifiable sufficient conditions for existence and uniqueness of recursive utilities in 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 parameterizations of 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 statespace, and possibly also bounded per-period utilities.\(^2\) However, most models used in macroeconomics and finance feature unbounded (i.e., non-compact) statespaces 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 statespace is simply to truncate (i.e. compactify) the statespace and apply results for compact statespaces. After all, this truncation occurs implicitly when computing solutions numerically, even though the model may have unbounded statespace. However, truncation of the statespace can materially affect existence and uniqueness. This is highlighted by an empirically relevant example presented in Appendix A, which is based on Bidder and Smith (2018) and Wachter (2013). In this example, there is always a unique solution in the truncated model (irrespective of the truncation level) even when there is no solution or multiple solutions without truncation. This example clearly shows there are important but subtle differences between models with compact and non-compact statespace. Understanding when the original model without truncation has a unique solution therefore remains a pressing issue, especially for reconciling numerical solutions with solutions of the original, un-truncated model.

---

\(^1\)Throughout the paper, by “recursive utility” we mean “stochastic recursive utility”.

\(^2\)Epstein and Zin (1989), Alvarez and Jermann (2005), Marinacci and Montrucchio (2010), Balbus (2015), Guo and He (2017), Bloise and Vailakis (2018), and Borovička and Stachurski (2020).

\(^3\)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), and Schorfheide, Song, and Yaron (2018).
For many of the models we study, the single primitive sufficient condition for both existence and uniqueness is that the distribution of growth in per-period utilities has thin tails, in a sense we make precise below. We verify this condition for robust preferences, models of ambiguity aversion and learning about hidden states, and a common parameterization of Epstein–Zin preferences. 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. As is well known (see, e.g., Marinacci and Montrucchio (2010)), the operators defining recursive utilities are typically not contraction mappings or local contractions in the usual sense. The literature has therefore typically appealed to fixed-point results for positive operators acting on a positive cone of functions. Such arguments often rely on topological properties of the positive cone in the space of bounded functions on a compact set, or other operator-theoretic side conditions.

Our point of departure is to dispense altogether with positivity-based arguments and 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). 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 of the recursion is monotone and its spectral radius is strictly less than one. For many of the models we study, the 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 then 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.

\footnote{An alternative approach taken by Marinacci and Montrucchio (2010) is to establish contractivity in the Thompson metric. This establishes uniqueness in a class of comparable functions, where the set of all functions \( w \) comparable to \( v \) is \( \{ w : a^{-1}w(x) \leq v(x) \leq aw(x) \text{ for some } a > 0 \} \). Such a class can be too restrictive when the statespace is unbounded. For instance, the exponential-affine functions \( v_1(x) = \exp(a_1 + b_1 x) \) and \( v_2(x) = \exp(a_2 + b_2 x) \) are only comparable if \( b_1 = b_2 \).

\footnote{Previously, Hindy and Huang (1992) and Hindy, Huang, and Kreps (1992) used Orlicz classes to define topologies for consumption paths in continuous time. Our use is more in line with the statistics literature, as we use rely on the “thin-tailed” nature of these classes to control the probabilities of tail events.}
To illustrate the usefulness of our results, we then present applications to three classes of models.

Section 3 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, Sargent, Turmuhambetova, and Williams, 2006), as well as under Epstein and Zin (1989) preferences with intertemporal elasticity of substitution (IES) equal to one. There are currently no uniqueness results in the literature for this recursion allowing non-compact statespace and unbounded utilities (see the discussion in Section 3), both of which are important for models in macroeconomics and finance. We fill this gap, establishing existence and uniqueness under a single thin-tail condition on utility growth. We verify the thin-tail condition in canonical linear-Gaussian environments as well as environments featuring regime-switching and stochastic volatility, thereby contributing new existence and uniqueness results for such settings.

Section 4 considers models with learning. The analysis encompasses extensions of multiplier preferences to accommodate both model uncertainty and uncertainty about hidden states following Hansen and Sargent (2007, 2010), some 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. Again, there are currently no existence and uniqueness results in the literature allowing non-compact statespace and unbounded utilities (see the discussion in Section 4). We establish existence and uniqueness under a single thin-tail condition on utility growth. We verify the condition—thereby establishing new 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 5 we examine Epstein–Zin recursive utilities when the IES does not equal one. There are no uniqueness results for models with unbounded statespace when risk aversion and intertemporal substitution are in a range normally encountered in the long-run risks literature. 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. All proofs are in the Appendix B.
2 Preliminaries

This section first describes the classes of “thin-tailed” functions with which we work. A basic existence and uniqueness result is presented. The key condition for uniqueness is a high-level spectral radius condition. Sufficient conditions for this are derived in terms of a thin-tail condition on a change-of-measure. We use this intermediate result to derive more primitive sufficient conditions in the applications in Sections 3-5.

2.1 Orlicz classes

We begin by briefly reviewing some relevant properties of Orlicz classes. We refer the reader to section 10 of Krasnosel’skii and Rutickii (1961) for further details.

Let \((X, \mathcal{X}, \mu)\) be a \(\sigma\)-finite measure space. In most of what follows we will study Markovian environments in which \(X_t\) is a state vector supported on \(X\) and \(\mu\) is the stationary distribution of \(X_t\). Let \(\psi: \mathbb{R}_+ \rightarrow \mathbb{R}_+\) be monotone, continuously differentiable, and strictly convex with \(\psi(0) = 0\) and \(\psi(x)/x \rightarrow +\infty\) as \(x \rightarrow +\infty\). The Luxemburg norm of \(f: X \rightarrow \mathbb{R}\) is defined as

\[
\|f\|_\psi = \inf \left\{ c > 0 : \int \psi(|f(x)|/c) \, d\mu(x) \leq 1 \right\}.
\]

Let \(L\) denote the (equivalence class of) all measurable \(f: X \rightarrow \mathbb{R}\) for which \(\|f\|_\psi < \infty\). Also let

\[
\mathcal{E}_0 = \left\{ f \in L : \int \psi(|f(x)|/c) \, d\mu(x) < \infty \text{ for each } c > 0 \right\},
\]

which is the closure of \(L^\infty\) (the space of all \(\mu\)-essentially bounded functions) in \(L\). Both \(L\) and \(\mathcal{E}_0\) are ordered Banach spaces when equipped with the norm \(\| \cdot \|_\psi\) and partial ordering \(f \geq g\) denoting \(f(x) \geq g(x)\) \(\mu\)-almost everywhere. If

\[
\lim_{x \rightarrow +\infty} \frac{x\psi'(x)}{\psi(x)} < +\infty
\]

then \(L = \mathcal{E}_0\); otherwise, \(\mathcal{E}_0\) is a proper subset of \(L\).

**Example 1: Lp classes.** Let \(\psi(x) = x^p\) for \(p \in (1, \infty)\). Then \(\| \cdot \|_\psi\) is equivalent to the \(L^p\) norm \(\|f\|_p := (\int |f(x)|^p \, d\mu(x))^{1/p}\) and \(L = \mathcal{E}_0 = L^p\).

**Example 2: Exponential-Orlicz classes.** These will play an important role in what follows. We shall use \(L^{\phi_r}\) and \(E^{\phi_r}\) to denote the spaces \(L\) and \(\mathcal{E}_0\) corresponding to \(\psi(x) = \exp(x^r) - 1\) with \(r \geq 1\), and denote the Luxemburg norm by \(\| \cdot \|_{\phi_r}\). Here \(E^{\phi_r}\) is a proper
subset of $L^{\phi_r}$. When $\mu$ is a probability measure, $L^\infty \hookrightarrow E^{\phi_r} \hookrightarrow L^{\phi_r} \hookrightarrow E^{\phi_s} \hookrightarrow L^p$ are continuous embeddings for $1 \leq s < r < \infty$, with $\|f\|_p \leq p!(\log 2)^{1/r-1}\|f\|_{\phi_r}$ for each $1 \leq p < \infty$, and $\|f\|_{\phi_s} \leq (\log 2)^{1/r-1/s}\|f\|_{\phi_r}$ (van der Vaart and Wellner, 1996, p. 95).

2.2 A basic fixed-point result

Let $E$ be a closed linear subspace of $E_0$ and let $\mu$ be a finite measure.\(^6\) Let

$$
\|D\|_E := \sup\{\|Df\|_\psi : f \in E, \|f\|_\psi = 1\} \\
\rho(D;E) := \lim_{n \to \infty} \|D^n\|_E^{1/n}
$$

denote the operator norm and spectral radius of a bounded (i.e., continuous) linear operator $D : E \to E$, where $D^n f$ denotes $D$ applied $n$ times in succession to $f$. The operator $T$ is monotone if $Tf \geq Tg$ whenever $f \geq g$. A bounded linear operator $D_f : E \to E$ is a subgradient of $T$ at $f \in E$ if

$$
Tg - Tf \geq D_f(g - f)
$$

holds for each $g \in E$. We say that a decreasing sequence of functions $\{v_n\}_{n \geq 1} \subset E$ is bounded from below by $\underline{v} \in E$ if $\liminf_{n \to \infty} v_n \geq \underline{v}$; similarly, an increasing sequence of functions $\{v_n\}_{n \geq 1} \subset E$ is bounded from above by $\bar{v} \in E$ if $\limsup_{n \to \infty} v_n \leq \bar{v}$. Let $T^n v$ denote $T$ applied $n$ times in succession to $v$. The following Proposition is a useful starting point for organizing the discussion that follows.

**Proposition 2.1.** (i) Existence: Let $T$ be a continuous and monotone operator on $E$ and let there exist $\underline{v}, \bar{v} \in E$ such that either (a) $T\bar{v} \leq \bar{v}$ and $\{T^n \bar{v}\}_{n \geq 1}$ is bounded from below by $\underline{v}$, or (b) $T\underline{v} \geq \underline{v}$ and $\{T^n \underline{v}\}_{n \geq 1}$ is bounded from above by $\bar{v}$. Then: $T^n \bar{v}$ (if (a) holds) or $T^n \underline{v}$ (if (b) holds) converges to a fixed point $v \in E$, with $\underline{v} \leq v \leq \bar{v}$.

(ii) Uniqueness: Suppose that at each of its each fixed points $v \in E$, $T$ has a subgradient $D_v$ which is monotone with $\rho(D_v;E) < 1$. Then: $T$ has at most one fixed point in $E$.

When uniqueness cannot be guaranteed, we use ordering and stability criteria to refine the set of fixed points. Let $V$ denote the set of fixed points of $T$. Say $v$ is the smallest fixed point of $T$ if $v \leq v'$ for each $v' \in V$. Say $v$ is stable if $\rho(D_v;E) < 1$ (see, e.g., Amann (1976)). Stability of $v$ is a useful property. In many of the examples we consider below, the subgradient is of the form $D_v = \beta \tilde{E}$ with $\beta \in (0,1)$, where $\tilde{E}$ is a distorted probability

\(^6\)Finiteness of the measure $\mu$ is not important for the existence result, but it simplifies some arguments and is relevant for the applications we study in Sections 3-5 where $\mu$ is a probability measure.
measure. Stability ensures that discounted expected utilities under $\tilde{E}$ are finite. Stability of $v$ also ensures that fixed-point iteration on a neighborhood of $v$ will converge to $v$.

**Corollary 2.1.** Let $v$ be a fixed point of $T$ with $\rho(D_v; E) < 1$. Then: $v$ is both the smallest fixed point and the unique stable fixed point of $T$ in $E$.

### 2.3 Verifying the spectral radius condition

In models featuring forward-looking agents, the subgradient is typically a discounted conditional expectation operator. However, nonlinearities of the recursion can introduce a wedge between the probability measure describing the evolution of state variables and the probability measure under which the expectation is taken.

When there is no such wedge (e.g., time-separable preferences and rational expectations), the spectral radius condition is easily seen to hold. Let $X = \{X_t\}_{t \geq 0}$ be a time-homogeneous stationary Markov process with transition kernel $Q$ and stationary (i.e., ergodic) distribution $\mu$. Let $D_v = \beta \mathbb{E}^Q$, where $\mathbb{E}^Q$ denotes conditional expectation under $Q$. Then for any $c > 0$ and $f \in \mathcal{E}$,

$$
\int \psi(|\mathbb{E}^Q f(x)|/c) \, d\mu(x) \leq \int \mathbb{E}^Q[\psi(|f(X_{t+1})|/c)|X_t = x] \, d\mu(x) = \int \psi(|f(x)|/c) \, d\mu(x),
$$

by Jensen’s inequality and the fact that $\mu$ is the stationary distribution associated with $X$. Taking $f$ to be almost-everywhere constant, we obtain $\|D_v\|_E = \beta$ and so $\rho(D_v; E) = \beta$.

This argument breaks down in the settings we study, where $D_v = \beta \tilde{E}$, where $\tilde{E}$ denotes conditional expectation under a distribution different from $Q$. We verify the spectral radius condition under a thin-tail condition on the change-of-measure transforming $\mathbb{E}^Q$ into $\tilde{E}$. Suppose

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

(3)

where $m$ is the (conditional) change-of-measure transforming $\mathbb{E}^Q$ into $\tilde{E}$. Repeatedly applying $D_v$ involves repeatedly multiplying by $m$, taking conditional expectations under $Q$, and discounting. Provided the moments of $m$ don’t grow too quickly, repeatedly applying $D_v$ to thin-tailed functions ensures discounting eventually dominates and the spectral radius condition holds. We now formalize this reasoning. Let $\log m \lor 0$ denote the pointwise maximum of $\log m$ and 0. Also let $\mu \otimes Q$ denote the joint (stationary) distribution of $(X_t, X_{t+1})$.
Lemma 2.1. Let $\mathbb{D} = \beta \tilde{\mathbb{E}}$ where $\beta \in (0, 1)$ and $\tilde{\mathbb{E}}$ is of the form (3) with
\[
\mathbb{E}^{\mu \otimes Q} \left[ \exp(\log m(X_t, X_{t+1}) \vee 0)|r/c) \right] < \infty
\]  
for some $c > 0$ and $r > 1$. Then: $\mathbb{D}$ is a continuous linear operator on $E^{\phi s}$ with $\rho(\mathbb{D}; E^{\phi s}) < 1$ for each $s \geq 1$.

Remark 2.1. Lemma 2.1 does not require stationarity (or any other property) of $X$ under the law of motion corresponding to $\tilde{\mathbb{E}}$.

Remark 2.2. Lemma 2.1 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 $\mathbb{D} = \beta \tilde{\mathbb{E}}$.

3 Application 1: Robust (and related) preferences

3.1 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
\[
V_t = U_t - \beta \theta \log \mathbb{E} \left[ e^{-\theta^{-1}V_{t+1}} \big| \mathcal{F}_t \right], \tag{5}
\]
where $\mathcal{F}_t$ is the date-$t$ information set, $\beta \in (0, 1)$ is a time preference parameter, and $\theta > 0$. Recursion (5) arises in a number of settings. It is the risk-sensitive recursion of Hansen and Sargent (1995), where $\theta$ is interpreted as a risk-sensitivity parameter. The recursion also arises under “robust” preferences which express an aversion to model uncertainty, namely multiplier preferences (Hansen and Sargent, 2001) and constraint preferences (Hansen et al., 2006), in which $\theta$ encodes the agent’s aversion to model uncertainty. Finally, recursion (5) is equivalent to Epstein and Zin (1989) preferences with unit IES, in which case $\theta$ is a transformation of the risk aversion parameter.\(^7\)

We follow much of the literature and consider environments characterized by a stationary Markov state process $X = \{X_t : t \geq 0\}$ supported on a statespace $\mathcal{X} \subseteq \mathbb{R}^d$. The set $\mathcal{F}_t$ will denote the information set generated by the realization of $X$ up to date $t$. Let $Q$ denote

\(^7\)Specifically, $\theta = 1/(\gamma - 1)$ where $\gamma$ is the coefficient of relative risk aversion. See, e.g., Section III in Hansen et al. (2008) for a derivation of recursion (5) from the Epstein–Zin recursion with unit IES.
the Markov transition kernel and \( \mathbb{E}^Q \) denote conditional expectation with respect to \( Q \). 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
\[
v(x_t) = -\frac{1}{\theta} \left( V_t - \frac{1}{1 - \beta} U_t \right), \quad u(x_t, x_{t+1}) = U_{t+1} - U_t.
\]
For instance, this is true when \( U_t = \log(C_t) \) and consumption growth \( \log(C_{t+1}/C_t) \) is a function of \((X_t, X_{t+1})\). Under these conditions, the recursion may be rewritten in terms of the scaled continuation value function \( v \):
\[
v(x) = \beta \log \mathbb{E}^Q \left[ e^{v(x_{t+1}) + \alpha u(x_t, x_{t+1})} \left| X_t = x \right. \right], \quad (6)
\]
where \( \alpha = -\left( \theta(1 - \beta) \right)^{-1} \). Recursion (6) may be expressed in operator notation as \( v = \mathbb{T}v \), where
\[
\mathbb{T}f(x) = \beta \log \mathbb{E}^Q \left[ e^{f(x_{t+1}) + \alpha u(x_t, x_{t+1})} \left| X_t = x \right. \right].
\]

### 3.2 Existing results

Hansen and Scheinkman (2012) and Christensen (2017) studied this recursion in the context of Epstein–Zin preferences with unit IES allowing 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. We establish both existence and uniqueness properties under a primitive tail condition for the stationary distribution of \( u(X_t, X_{t+1}) \).

### 3.3 New results

Given the form of \( \mathbb{T} \), the functions \( f \) and \( u \) must have sufficiently thin tails in order that \( \mathbb{T}f \) be well defined. We therefore work with Orlicz classes of the form \( E^{\phi_r} \) with \( r \geq 1 \) defined relative to the stationary distribution \( \mu \) of \( X \). The single condition we require is that the

---

8Our 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 \((X_t, Y_t)\).
stationary distribution of per-period utility growth has thin tails: for some $r \geq 1$, we have
\[ E^{\mu \otimes Q} \left[ \exp \left( |u(X_t, X_{t+1})|^{r} / c \right) \right] < \infty \quad \text{for all } c > 0. \tag{7} \]
We verify this condition below in three examples. In the second example, we show that uniqueness can fail when condition (7) does not hold.

We shall establish existence and uniqueness by applying Proposition 2.1. The operator $T$ is continuous, monotone, and convex under condition (7); see Lemma B.5. The proof of existence constructs an upper value $\bar{v}$ and shows the sequence of iterates $\{T^n \bar{v}\}_{n \geq 1}$ is bounded from below. For uniqueness, the operator $T$ the subgradient inequality (2) with subgradient
\[ D_v f(x) = \beta E_v f(x), \]
where $E_v$ is a distorted conditional expectation:
\[ E_v f(x) = E^{Q} [m_v(X_t, X_{t+1}) f(X_{t+1}) | X_t = x], \tag{8} \]
\[ m_v(X_t, X_{t+1}) = \frac{e^{v(X_{t+1}) + au(X_t, X_{t+1})}}{E^{Q} [e^{v(X_{t+1}) + au(X_t, X_{t+1})} | X_t]} \tag{9} \]
For robust preferences, $E_v$ may be interpreted as expectation under the agent’s “worst-case” model. The spectral radius condition is verified by applying Lemma 2.1; see Lemma B.6.

**Theorem 3.1.** Let condition (7) hold. Then: $T$ has a fixed point $v \in E^{\phi_r}$. Moreover, if $r > 1$ then: (i) $v$ is the unique fixed point of $T$ in $E^{\phi_s}$ for each $s \in (1, r]$, and (ii) $v$ is both the smallest fixed point and the unique stable fixed point of $T$ in $E^{\phi_1}$.

**Example 1: Linear-Gaussian environments.** Condition (7) 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 $U_t = \log(C_t e^{X_t})$ where $\log(C_{t+1}/C_t)$ is a function of $(X_t, X_{t+1})$ and the process $X$ is a stationary Gaussian VAR(1):
\[ X_{t+1} = \nu + AX_t + u_{t+1}, \quad u_{t+1} \sim N(0, \Sigma), \]
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 $T$ 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

\[
a = \frac{\beta}{1 - \beta} \left( (\alpha \lambda_1 + b)' \nu + \frac{1}{2} (\alpha \lambda_1 + b)' \Sigma' (\alpha \lambda_1 + b) \right).
\]

Theorem 3.1 shows that \( v(x) = a + b'x \) is the unique fixed point in \( E^{\phi_s} \) for all \( s \in (1, 2) \), and the smallest fixed point and unique stable fixed point in \( E^{\phi_1} \).

**Example 2: Fat tails and rare disasters.** This example shows there can exist multiple fixed points when condition (7) is violated. The model features time-varying rare disasters from Bidder and Smith (2018). A similar model is studied in Wachter (2013) in the context of Epstein–Zin preferences with IES = 1. Consumption growth \( g_{t+1} := \log(C_{t+1}/C_t) \) is modeled as

\[
g_{t+1} = \nu_g + w_{g,t+1} + \sigma w_{z,t+1},
\]

with \( w_{g,t+1} \sim N(0, 1), w_{z,t+1} | j_{t+1} \sim N(\nu_j j_{t+1}, \sigma_j^2 j_{t+1}) \) where \( \nu_j \) is Poisson with mean \( h_t \) which follows an autoregressive gamma (ARG) process. Defining \( X_t = (g_t, h_t) \), we see that \( u(X_t, X_{t+1}) = g_{t+1} \). By iterated expectations we may deduce

\[
E^{\phi_s} \left[ e^{cu(X_t, X_{t+1})} \right] = e^{c
u_g + c^2 \sigma^2} \left[ \exp \left( h_t \left( \exp \left( c
u_j + c^2 \sigma_j^2 \right) - 1 \right) \right) \right].
\]

Condition (7) 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. Indeed, it is known that there may exist zero, one, or two fixed points of the form \( v(x) = a + b'x \) under this specification. The precise number of fixed points of this form is determined by the number of real solutions to a particular quadratic equation.

One could modify the above specification so that \( w_{z,t+1} | j_{t+1} \sim N(\mu_j j_{t+1}, \sigma_j^2) \) for some \( \zeta \in \left[ \frac{1}{2}, 1 \right) \). Given the low frequency of jumps, this modification is likely difficult to distinguish empirically from the original specification. Under this modification, one may deduce that condition (7) holds for each \( r \in [1, 1/\zeta) \). Therefore, there is a unique fixed point \( v \in E^{\phi_s} \) for all \( s \in (1, 1/\zeta) \), and a unique stable fixed point in \( E^{\phi_1} \).

**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 \( A_\delta \) are inside the unit circle for each \( s = 1, \ldots, N \). The full state vector is now \((X_t, s_t)\), which is jointly 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 (7) holds for all \( r \in [1, 2) \). It follows by Theorem 3.1 there is a unique fixed point in \( E^{\phi_s} \) for all \( s \in (1, 2) \) (with \( E^{\phi_s} \) defined with respect to the stationary distribution of \((X_t, s_t)\)), and a unique stable fixed point in \( E^{\phi_1} \).

**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

\[
g_{t+1} = \bar{g} + x_t + \sigma_t \eta_{t+1}^g, \\
x_{t+1} = \rho_x x_t + \varphi_x \sigma_t \eta_{t+1}^x, \\
\sigma_{t+1}^2 = \bar{\sigma}^2 + \rho_{\sigma}(\sigma_t^2 - \bar{\sigma}^2) + \varphi_{\sigma} \eta_{t+1}^\sigma,
\]

where \( \eta_{t+1}^g, \eta_{t+1}^x, \) and \( \eta_{t+1}^\sigma \) 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 \( \rho_x = 0 \) though this is not essential to our analysis. Second, to deal with the complications arising when \( \sigma_{t+1}^2 < 0 \) we take absolute values. This leads to the consumption growth process

\[
g_{t+1} = \bar{g} + \sqrt{|v_t|} \eta_{t+1}^g, \\
v_{t+1} = \bar{v} + \rho_v (v_t - v) + \varphi_v \eta_{t+1}^v,
\]

where \( \eta_{t+1}^g \) and \( \eta_{t+1}^v \) are i.i.d. \( N(0, 1) \). Defining \( X_t = v_t \), we see that \( u(X_t, X_{t+1}) = g_{t+1} \) when per-period utility is logarithmic in consumption. To verify condition (7), first note that

\[
E^{\mu \otimes Q}[\exp((|g_{t+1} - \bar{g}|)/c^r)] = E^{\mu'}[E[\exp(|\sqrt{|v_t|} \eta_{t+1}^g/c^r)|v_t]],
\]

where the inner expectation is taken with respect to \( \eta_{t+1}^g \sim N(0, 1) \). The inner expectation is equivalent to \( E[\exp(Y/r'\alpha^r)] \) where \( Y = |Z| \) with \( Z \sim N(0, 1) \) and \( \alpha = c/\sqrt{|v_t|} > 0 \). In Appendix B we derive a crude bound on this expectation (see Lemma B.7) from which we may deduce

\[
E \left[ \exp \left( \left( \frac{1}{c^r} \frac{2\sqrt{|v_t|} \eta_{t+1}^g}{c^r} \right) \frac{1}{r} \exp \left( \frac{(2|v_t|)^{\frac{1}{2r}}}{c^r} \right) + \frac{4\sqrt{|v_t|} \eta_{t+1}^g}{c^r} \right) \right] \leq \sqrt{2 \pi} \left( \frac{2\sqrt{|v_t|} \eta_{t+1}^g}{c^r} \right)^{\frac{1}{2r}} \exp \left( \frac{(2|v_t|)^{\frac{1}{2r}}}{c^r} \right) + \frac{4\sqrt{|v_t|} \eta_{t+1}^g}{c^r} \right)^{\frac{1}{2r}} + 2\sqrt{\pi},
\]

which is valid for \( r \in [1, 2) \). As the stationary distribution of \( v_t \) is Gaussian and the exponent \( \frac{1}{2r} \) of the \(|v_t|\) term appearing in the right-hand side exponential is less than 2 for \( r \in [1, 4/3) \),
it follows that the expectation (10) is finite for all $c > 0$ provided $r \in [1, 4/3)$. As the location shift by $\hat{g}$ does not affect finiteness of the moments, condition (7) holds for all $r \in [1, 4/3)$. Therefore, there is a unique fixed point in $E^{g_s}$ for all $s \in (1, 4/3)$ and a unique stable fixed point in $E^{g_1}$.

4 Application 2: Learning and ambiguity

This section extends the setting from Section 3 to a class of dynamic models where 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).

4.1 Setting

We again consider environments characterized by a Markov state process $X = \{X_t\}_{t \geq 0}$ with transition kernel $Q$. Partition the state as $X_t = (\varphi_t, \xi_t)$ where the agent observes $\varphi_t$ but does not observe $\xi_t$. Let $O_t = \sigma(\varphi_t, \varphi_{t-1}, \ldots, \varphi_0)$ denote the history of the observed state to date $t$. Beliefs about $\xi_t$ are summarized by a posterior distribution $\Pi_t$ conditional on $O_t$. 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

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

(11)

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 ($\Pi_t$), where $\vartheta > 0$ and $\theta > 0$ encode concerns about misspecification of $Q$ and $\Pi_t$, respectively. When $U_t = \log C_t$, recursion (11) also arises under generalized recursive smooth ambiguity preferences of Ju and Miao (2012) with unit IES, where $\theta$ and $\vartheta$ are one-to-one transformations of their ambiguity aversion and risk.
aversion parameters, respectively. When \( \vartheta = \theta \), recursion (11) reduces to

\[
V_t = U_t - \beta \vartheta \log \mathbb{E}^{\Pi_t} \left[ e^{-\theta^{-1}V_{t+1}} \left| O_t, \xi_t \right. \right] O_t .
\]

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 (11) becomes

\[
V_t = U_t - \beta \theta \log \mathbb{E}^{\Pi_t} \left[ e^{-\theta^{-1}\mathbb{E}^Q[V_{t+1}]} \left| O_t, \xi_t \right. \right] O_t . \tag{12}
\]

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

We impose several (standard) conditions to make the problem tractable. First, the state is assumed to have a conventional hidden Markov structure, in which

\[
Q(X_{t+1}|X_t) = Q_\varphi(\varphi_{t+1}|\xi_t)Q_\xi(\xi_{t+1}|\xi_t) .
\]

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 \( \varphi_t \) to influence \( \varphi_{t+1} \), but we maintain this simpler presentation for convenience.

Second, we assume \( \Pi_t \) 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 \( \Pi_\xi \), 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 \( \xi_t \) 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 \( \varphi_t \).

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. Uniqueness of the limiting steady state recursion is necessary for uniqueness of the sequence of backward iterates.

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

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

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 \( \varphi_{t+1} \).

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

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

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 \( X \) 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 \( u^\varphi_t \) and \( u^\xi_t \) are i.i.d. \( N(0, \Sigma_u) \) and \( N(0, \Sigma_w) \), respectively, and all eigenvalues of \( B \) are inside the unit circle. This is the setting studied in Hansen and Sargent (2007, 2010),

\footnote{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.}
Croce et al. (2015), Collard et al. (2018), and several other works. If ξ_0 ∼ N(μ_0, Σ_0) under Π_0 then ξ_t ∼ N(μ_t, Σ_t) under Π_t. The matrix Σ_t will converge to a fixed matrix Σ as \( t \to \infty \). In this steady state, the sufficient statistic for Π_t is \( \hat{\xi}_t \) which is updated according to the rule

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

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

4.2 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 (12)). Their result applies to bounded functions and requires bounded per-period utilities.

4.3 New results

Recursion (11) may be reformulated as the fixed-point equation \( v = Tv \) where

\[
Tf(\hat{\xi}_t) = \beta \log \mathbb{E}^{\Pi_t} \left[ \mathbb{E}^{Q_{\varphi}} \left[ e^{\frac{\hat{\varphi}}{\sigma} f(\Xi(\hat{\xi}_t, \varphi_{t+1}))+\alpha u(\varphi_{t+1})} \bigg| \xi_t, \hat{\xi}_t \right] \right].
\]

Recursion (12) in the limiting case with \( \vartheta = +\infty \) may be reformulated as the fixed-point equation \( v = Tv \) where

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

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

\[
u \in E^{\phi_{\varphi}}
\]

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 $T$ by applying Proposition 2.1. Further details on the form of the subgradient and verification of Lemma 2.1 are deferred to Appendix B.4.

**Theorem 4.1.** Let condition (13) hold. Then: $T$ has a fixed point $v \in E^\phi_{\hat{\xi}}$. Moreover, if $r > 1$, then: (i) $v$ is the unique fixed point of $T$ in $E^\phi_s$ for all $s \in (1, r]$, and (ii) $v$ is both the smallest fixed point and the unique stable fixed point of $T$ in $E^\phi_{1}$.

**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 (13) 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 in $E^\phi_s$ for all $s \in (1, 2)$, and a unique stable fixed point in $E^\phi_{1}$.

**Example 2: Gaussian state-space models (continued).** Here the stationary distribution of $u(\varphi_{t+1})$ is Gaussian, so (13) holds for all $r \in [1, 2)$. Therefore, there is a unique fixed point in $E^\phi_s$ for all $s \in (1, 2)$, and a unique stable fixed point in $E^\phi_{1}$.

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

$$Q(X_{t+1}|X_t) = Q_{\varphi}(\varphi_{t+1}|\xi_t, \varphi_t)Q_{\xi}(\xi_{t+1}|\xi_t).$$

In this case, however, the effective state vector will be $\hat{X}_t$ rather than $\hat{\xi}_t$.

**5 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 statespace. There are currently no uniqueness results for the recursion we study with non-compact statespace. This is a complicated issue and it would be 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.
5.1 Setting

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

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

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 case in which $\rho \neq 1$ in this section; the case with $\rho = 1$ is subsumed in the analysis of Section 3. We consider environments characterized by a stationary Markov state process $X = \{X_t : t \geq 0\}$ supported on a statespace $\mathcal{X} \subseteq \mathbb{R}^d$. 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 $g: \mathcal{X} \times \mathcal{X} \to \mathbb{R}$.

Then $(1 - \rho) \log(V_t/C_t) = v(X_t)$ where the function $v: \mathcal{X} \to \mathbb{R}$ solves

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

with $\kappa = \frac{1-\gamma}{1-\rho}$ (see, e.g., Hansen et al. (2008)). The properties of this recursion differ depending on whether $\kappa < 0$, $\kappa \in (0, 1)$, or $\kappa \in [1, \infty)$. For brevity we focus on the former, as it is the pertinent case in the long-run risks literature where standard typically $\gamma > 1$ and $1/\rho > 1$. Similar arguments can be used to study the other cases.

5.2 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 statespace.

The two most closely related works are Hansen and Scheinkman (2012; HS hereafter) and Borovička and Stachurski (2020; BS hereafter), both of which present conditions for existence (but not uniqueness) when $\mathcal{X}$ is unbounded in the case $\kappa < 0$.

Our results trivially extend to allow $\log(C_{t+1}/C_t) = g(X_t, X_{t+1})$ 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 $(X_t, Y_t)$.

Hansen and Scheinkman (2012) also establish sufficient conditions for uniqueness when $\kappa \geq 1$. 

---

10 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 $(X_t, Y_t)$.

11 Hansen and Scheinkman (2012) also establish sufficient conditions for uniqueness when $\kappa \geq 1$. 

---
sufficient conditions only for existence because the operator does not have a subgradient of the form studied in Section 2.3 when $\rho \neq 1$. The existence conditions in HS require stationarity of the state under a distorted law of motion and restrict the size of a Perron–Frobenius eigenvalue and moments of its eigenfunction. Our first result below imposes the same eigenvalue and stationarity conditions and a slightly stronger integrability condition.\footnote{Specifically condition (18) below, which is stronger than HS’s Assumptions 4 and 5.}

The stronger integrability condition does not seem to bite for models commonly encountered (see the linear-Gaussian example at the end of this section) and also ensures that the stochastic discount factor (SDF)

$$\beta(C_{t+1}/C_t)^{-\rho} \left[ \frac{V_{t+1}^{1-\gamma}}{\mathbb{E}Q[V_{t+1}^{1-\gamma}]} \right]^{\frac{\rho}{1-\gamma}} \equiv \beta(C_{t+1}/C_t)^{-\rho} \left[ \frac{h(X_{t+1})(C_{t+1}/C_t)^{1-\gamma}}{\mathbb{E}Q[h(X_{t+1})(C_{t+1}/C_t)^{1-\gamma}|X_t]} \right]^{\frac{\rho}{1-\gamma}}$$

is well defined provided consumption growth has sufficiently thin tails. Our second result does not require stationarity under a distorted law of motion. BS showed HS’s eigenvalue condition is necessary and sufficient for existence in $L^1$ under additional some additional weak-compactness and irreducibility side conditions on an operator. We require no such operator-theoretic side conditions.

5.3 New results

Under general conditions (see Hansen and Scheinkman (2009) and Christensen (2015, 2017)), there exists a strictly positive function $\iota$ and scalar $\lambda > 0$ solving\footnote{Note the function $\iota$ is defined only up to scale normalization.} the equation

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

Hansen and Scheinkman (2009) use $\iota$ and $\lambda$ to define a distorted conditional expectation operator

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

HS show solving (14) is equivalent to finding a fixed point of

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

(17)
with the $v$ to recursion (14) and the fixed point of $T$ differing additively by $\frac{1}{\kappa} \log \epsilon$.\textsuperscript{14}

For the first two results, we assume $X$ is stationary under the law of motion corresponding to the distorted conditional expectation $\tilde{E}$ (Theorem 5.2 below does not require this). Let $\mu$ denote the stationary distribution induced by $\tilde{E}$. The first result is for Orlicz spaces $\tilde{E}^{\phi_s}$ defined relative to $\mu$ (subsequent results pertain to the true stationary measure $\mu$).

Our first regularity condition requires that $\log \epsilon$ has thin tails, in the sense that $\log \epsilon \in \tilde{E}^{\phi_r}$ for some $r \geq 1$. (18)

Under this condition, Lemma B.10 shows that $\tilde{T}$ is a continuous, monotone operator on $\tilde{E}^{\phi_s}$ for each $1 \leq s \leq r$. It is clear that $\tilde{T}v \geq \log((1 - \beta)\epsilon(x)^{-\frac{1}{\kappa}})$. Therefore, should there exist a $\tilde{v} \in \tilde{E}^{\phi_r}$ for which $\tilde{T}\tilde{v} \leq \tilde{v}$, the sequence of iterates $\tilde{T}^n\tilde{v}$ must be bounded from below. The remainder of the proof shows that the inequality $\tilde{T}\tilde{v} \leq \tilde{v}$ holds for the function

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

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

$$
\beta \lambda \frac{1}{\kappa} < 1.
$$

\textbf{Remark 5.1.} Although $\tilde{T}$ is not contractive, it follows from Proposition 2.1(i) that the sequence of iterates $\tilde{v}, \tilde{T}\tilde{v}, \tilde{T}^2\tilde{v}, \ldots$ will converge to a fixed point of $\tilde{T}$ under the conditions of any of Theorems 5.1 and 5.2 below.

\textbf{Theorem 5.1.} Let $X$ be stationary under the law of motion corresponding to the distorted conditional expectation $\tilde{E}$, $\kappa < 0$, and conditions (18) and (19) hold. Then: $\tilde{T}$ has a fixed point in $\tilde{E}^{\phi_s}$ and therefore the recursion (14) has a solution $v \in \tilde{E}^{\phi_s}$ for all $s \in [1, r]$.

Theorem 5.1 has implications for existence results in spaces defined relative to the stationary distribution $\mu$ of $X$. Suppose that $\mu$ is absolutely continuous with respect to the true stationary distribution $\tilde{\mu}$ of $X$, and that $\mu$ is absolutely continuous with respect to $\mu$. If so, we let $\Delta = \frac{d\tilde{\mu}}{d\mu}$ denote the change of measure of $\tilde{\mu}$ with respect to $\mu$. Consider the

\textsuperscript{14}The version of recursion (16) above appears on p. 11968 of HS. In our notation, their recursion is $\hat{U}g(x) = (1 - \beta)\epsilon(x)^{-\frac{1}{\kappa}} + \beta \lambda \frac{1}{\kappa} \hat{E}[g(X_{t+1})]\epsilon|X_t = x]^{\frac{1}{\kappa}}$. Recursion (16) is obtained by setting $\tilde{T}f = \log(\hat{U}(\exp(f)))$. 

20
following thin-tail condition on $\Delta$:

$$E^\mu[\Delta(X_t)^{1+\varepsilon}] < \infty \quad \text{and} \quad E^\mu[\Delta(X_t)^{-\varepsilon}] < \infty \quad \text{for some } \varepsilon > 0. \quad (20)$$

A sufficient condition for (20) is that $\log \Delta \in L^{\phi_1}$. The spaces $E^{\phi_r}$ and $\tilde{E}^{\phi_r}$ are equivalent under condition (20); see Lemma B.3. We may therefore restate condition (18) as

$$\log \iota \in E^{\phi_r} \quad \text{for some } r \geq 1. \quad (21)$$

We have the following version of Theorem 5.1 restated for the space $E^{\phi_r}$.

**Corollary 5.1.** Let $X$ be stationary under the law of motion corresponding to the distorted conditional expectation $\tilde{E}$, $\kappa < 0$, and conditions (19), (20), and (21) hold. Then: $T$ has a fixed point in $E^{\phi_s}$ and therefore the recursion (14) has a solution $v \in E^{\phi_s}$ for all $s \in [1, r]$.

Unlike the preceding two results and those in Hansen and Scheinkman (2012), the final existence result does not require stationarity of $X$ under $\tilde{E}$. However, this result requires a further eigenvalue condition:

$$\beta \lambda^{\frac{1}{r} - 1} < 1. \quad (22)$$

As $\lambda > 1$ in standard parameterizations, this condition is typically not binding in view of condition (19).

**Theorem 5.2.** Let $\kappa < 0$ and conditions (19), (21), and (22) hold. Then: $T$ has a fixed point in $E^{\phi_s}$ and therefore the recursion (14) 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. Let $X$ evolve as a stationary Gaussian VAR(1):

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

with all eigenvalues of $A$ are inside the unit circle and $\log(C_{t+1}/C_t) = \delta'X_{t+1}$ for some vector $\delta$, which is trivially true if log consumption growth is itself a component of $X_t$. Solving (16),

$$\iota(x) = e^{(1-\gamma)\delta'(I-A)^{-1}x}, \quad \lambda = e^{\frac{(1-\gamma)^2}{2}\delta'(I-A)^{-1}\Sigma(I-A')^{-1}\delta+(1-\gamma)\delta'(I-A)^{-1}\nu}.$$
To apply Corollary 5.1 we must verify conditions (19), (20), and (21). To verify condition (20), first note
\[
\frac{\ell(X_{t+1})(C_{t+1}/C_t)^{1-\gamma}}{\lambda_t(X_t)} = e^{(1-\gamma)\delta'(I-A)^{-1}u_{t+1} - \frac{(1-\gamma)^2}{2}\delta'(I-A)^{-1}\Sigma(I-A^\prime)^{-1}\delta}
\]
so the \(u_t\) are i.i.d. \(N((1 - \gamma)\delta'(I - A)^{-1}\Sigma, \Sigma)\) under \(\tilde{E}\). Equivalently, under \(\tilde{E}\) 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 \(\mu\) 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 (20) holds for any \(\varepsilon > 0\). As \(\log \ell(x)\) is also affine in \(x\), we have that \(\log \ell \in E^{\phi_r}\) for all \(r \in [1,2)\), which verifies condition (21). It follows that the single condition one needs to verify for existence of recursive utilities in linear-Gaussian environments is the eigenvalue condition
\[
\beta e^{(1-\rho)(1-\gamma)\frac{(1-\gamma)^2}{2}\delta'(I-A)^{-1}\Sigma(I-A^\prime)^{-1}\delta + (1-\rho)\delta'(I-A)^{-1}\nu} < 1.
\]
Note also that here \(\log(C_{t+1}/C_t) \in E^{\phi_r}\) for all \(r \in [1,2)\). The SDF (15) is therefore well defined and all of its moments exist.

## A Truncation affects existence and uniqueness

This appendix provides an example to show truncating the statespace affects existence and uniqueness.

Consider the model of Bidder and Smith (2018) and Wachter (2013) from Section 3. In that example, condition (7) fails and it is known that there can exist zero, one, or two fixed points of the form \(v(x) = a + b'x\), where the multiplicity of fixed points depends on the number of real solutions to a particular quadratic equation. We will now see that truncating the statespace always results in a unique fixed point, irrespective of non-existence or non-uniqueness in the original, un-truncated model.

The state variable \(h_t\), representing the intensity at which disasters arrive, is supported on \([0, \infty)\). Suppose that its support is truncated to \([0, \bar{h}]\) for some \(\bar{h} < \infty\). Note that the only relevant conditioning variable is \(h\), so it suffices to confine our attention to functions that depend on \(h\) only. A truncated transition kernel \(Q_h(x'|x) \equiv Q_h(x'|h)\) may be constructed...
naturally by restricting $Q$ to the truncated space and renormalizing:

$$Q_h(x'|x) = \frac{Q(x'|x)1\{0 \leq h' \leq \bar{h}\}}{\mathbb{E}^Q[1\{0 \leq h_{t+1} \leq \bar{h}\}|X_t = x]}$$

for each $x = (g, h) \in \mathbb{R} \times [0, \bar{h}]$. Under this truncation,

$$\mathbb{E}^\bar{h}[e^{\alpha u(X_t, X_{t+1})} | X_t = x] = \exp \left( c \nu_g + \frac{c^2 \sigma^2}{2} + \bar{h} \left( \exp \left( c \nu_j + \frac{c^2 \sigma^2}{2} \right) - 1 \right) \right)$$

which is bounded between $\exp(c \nu_g + \frac{c^2 \sigma^2}{2})$ and $\exp(c \nu_g + \frac{c^2 \sigma^2}{2} + \bar{h}(\exp(c \nu_j + \frac{c^2 \sigma^2}{2}) - 1))$.

Let $T_{\bar{h}}$ denote the operator $T$ with the true transition kernel $Q$ replaced by the truncated transition kernel $Q_{\bar{h}}$. It is straightforward to verify that $T_{\bar{h}}$ satisfies Blackwell’s conditions for a contraction mapping on the space $\mathcal{B}([0, \bar{h}])$ of bounded functions on $[0, \bar{h}]$ equipped with the sup-norm. Therefore, $T_{\bar{h}}$ has a unique, globally attracting fixed point in $\mathcal{B}([0, \bar{h}])$ for all $\bar{h} < \infty$. Existence and uniqueness of a fixed point of $T_{\bar{h}}$ in $\mathcal{B}([0, \bar{h}])$ holds irrespective of the existence or uniqueness of fixed points of the original, un-truncated operator $T$.

B Proofs

Remark B.1. Several of the proofs below require showing that a function $f$ is an element of $E_{\psi}^s$ 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

$$\mathbb{E}^\mu[\exp(|f(X_t)/\bar{c}|^s)] = \mathbb{E}^\mu[\exp(|f(X_t)/\bar{c}|^s)]^{(\bar{c}/c)^s} \leq \mathbb{E}^\mu[\exp(|f(X_t)/c|^s)]^{(\bar{c}/c)^s}$$

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

B.1 Ancillary results

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

Lemma B.1 (Pollard (2015)). Let $\mathbb{E}^\mu[\psi(|f(X_t)|/C)] \leq C''$ for finite constants $C > 0$ and $C'' \geq 1$. Then: $\|f\|_{\psi} \leq CC''$.
Proof of Lemma B.1. Take \( \tau \in [0,1] \). By convexity of \( \psi \):

\[
\mathbb{E}^\mu[\psi(\tau|f(X_t)|/C)] \leq \tau \mathbb{E}^\mu[\psi(|f(X_t)|/C)] + (1 - \tau)\psi(0) = \tau \mathbb{E}^\mu[\psi(|f(X_t)|/C)].
\]

The result follows by setting \( \tau = 1/C' \).

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 \( \| \prod_{i=1}^{\infty} f_i \|_1 \leq \prod_{i=1}^{\infty} \|f_i\|_{p_i} \).

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

Lemma B.3. Let \( \mu \ll \nu \) and \( \int \Delta^p \, d\nu < \infty \) 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 \( r \geq 1 \).

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:

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

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

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

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

B.2 Proofs for Section 2

Proof of Proposition 2.1. Existence: we prove this for case (a); similar arguments apply for (b). The sequence \( \{\tilde{v}_n\}_{n \geq 1} \) with \( \tilde{v}_n = \mathbb{T}^n \bar{v} \) is monotone: \( \underline{u} \leq \ldots \leq \bar{v}_{n+1} \leq \bar{v}_n \leq \ldots \leq \bar{v} \) with \( \underline{v}, \bar{v} \in \mathcal{E} \). The sequence is therefore bounded in \( \mathcal{E} \) and hence in \( L^1 \) because \( \mathcal{E} \hookrightarrow L^1 \). It follows by Beppo Levi’s monotone convergence theorem (Malliavin, 1995, Theorem I.7.1) that there exists \( v \in L^1 \) such that \( \lim_{n \to \infty} \bar{v}_n = v \) (almost everywhere) and \( \lim_{n \to \infty} \|\bar{v}_n - v\|_1 = 0 \).

To strengthen convergence in \( \| \cdot \|_1 \) to convergence in \( \| \cdot \|_\psi \), first observe that \( \underline{v} \leq v \leq \bar{v} \)
and hence \( v \in \mathcal{E} \). To establish a contradiction, suppose that \( \limsup_{n \to \infty} \| \bar{v}_n - v \|_\psi \geq 2\varepsilon \) for some \( \varepsilon > 0 \). Then:

\[
\limsup_{n \to \infty} \int \psi(|\bar{v}_n - v|/\varepsilon) \, d\mu \geq 1.
\]

(23)

Note that \( \{f_n\}_{n \in \mathbb{N}} \) with \( f_n = \psi(|\bar{v}_n - v|/\varepsilon) \) is a monotone sequence of non-negative functions with \( \limsup_{n \to \infty} f_n = 0 \) (almost everywhere). Moreover, each \( f_n \leq \psi((|\bar{v}| + |\bar{u}| + |v|)/\varepsilon) \) where \( \int \psi((|\bar{v}| + |\bar{u}| + |v|)/\varepsilon) \, d\mu < \infty \) for each \( \varepsilon > 0 \) because \( \bar{v}, \bar{u} \) and \( v \) all belong to \( \mathcal{E} \). Therefore, by reverse Fatou:

\[
\limsup_{n \to \infty} \int \psi(|\bar{v}_n - v|/\varepsilon) \, d\mu \leq \int \limsup_{n \to \infty} \psi(|\bar{v}_n - v|/\varepsilon) \, d\mu = 0
\]

contradicting (23). It follows that \( \| \bar{v}_n - v \|_\psi \to 0 \). Finally,

\[
\| Tv - v \|_\psi \leq \| Tv - T\bar{v}_n \|_\psi + \| T\bar{v}_n - v \|_\psi = \| Tv - T\bar{v}_n \|_\psi + \| \bar{v}_{n+1} - v \|_\psi \to 0
\]

by continuity of \( T \), hence \( Tv = v \).

Uniqueness: Suppose that \( v, v' \in \mathcal{E} \) are fixed points of \( T \). By the subgradient inequality

\[
v' - v = Tv' - Tv \geq \mathbb{D}_v(v' - v)
\]

which implies that

\[
(\mathbb{I} - \mathbb{D}_v)(v' - v) \geq 0.
\]

(24)

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 (24) yields \( v' - v \geq 0 \). A parallel argument yields \( v - v' \geq 0 \). Therefore, \( v = v' \).

\( \square \)

**Proof of Corollary 2.1.** By the subgradient inequality, for \( v, v' \in \mathcal{V} \):

\[
v' - v = Tv' - Tv \geq \mathbb{D}_v(v' - v)
\]

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 \( T \).

Suppose any other \( v' \in \mathcal{V} \) distinct from \( v \) were also stable. Then we could apply an identical argument to obtain the reverse inequality \( v - v' \geq 0 \), a contradiction.

\( \square \)
Before proceeding, we present an intermediate result used to prove Lemma 2.1. Note that condition (4) implies that \((\log m \lor 0) \in L^{\phi} (\mu \otimes Q)\), 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 \(\| (\log m \lor 0) \|_{\phi_r}\) denote the corresponding Orlicz norm of \((\log m \lor 0)\).

**Lemma B.4.** Let \(\tilde{E}\) be of the form (3) and let \(m\) satisfy condition (4). Then for any \(p \in (1, \infty)\) and \(n \geq 1:\)

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

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

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

holds for each \(n \geq 1\).

**Proof of Lemma B.4.** First note \(\mathbb{E}^{\mu \otimes Q}[m(X_t, X_{t+1})^{np}] \leq \mathbb{E}^{\mu \otimes Q}[e^{np|\log m(X_t, X_{t+1})\lor 0}]\). To simplify notation, let \(Y_t = (X_t, X_{t+1})\), \(a = \log m \lor 0\), and \(\|a\|_{\phi_r} = \| (\log m \lor 0) \|_{\phi_r}\). 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 \(a_+ = |a| \mathbb{I}\{|a| \leq A\}\) and \(a_- = |a| \mathbb{I}\{|a| > A\}\). For any \(z > 0\), we have

\[
\Pr(e^{np|a(Y_t)|} \geq z) = \Pr(a_+(Y_t) \geq \frac{\log z}{2np}) + \Pr(a_-(Y_t) \geq \frac{\log z}{2np}).
\]

By Markov’s inequality and definition of \(\| \cdot \|_{\phi_r}\), we have

\[
\Pr(a_-(Y_t) \geq \frac{\log z}{2np}) \leq \Pr \left( |a(Y_t)|^{r} \geq \frac{A^{-1} \log z}{2np} \right) = \Pr \left( \exp \left( \frac{|a(Y_t)|^{r}}{\|a\|_{\phi_r}^{r}} \right) \geq \exp \left( \frac{1}{\|a\|_{\phi_r}^{r}} A^{-1} \log z \right) \right) \leq \frac{\mathbb{E}^{\mu \otimes Q}[\exp(|a(Y_t)|^{r})]}{\mathbb{E}^{\mu \otimes Q}[\exp(|a(Y_t)/\|a\|_{\phi_r}^{r}|^{r})]} \leq 2 \exp \left( - \frac{1}{\|a\|_{\phi_r}^{r}} \frac{A^{-1} \log z}{2np} \right).
\]
Setting $A = (\|a\|_{\phi_n}^r, 4np)^{\frac{1}{r-1}}$, 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
\[
\int_0^\infty \Pr \left( a_-(Y_t) \geq \frac{\log z}{2np} \right) \, dz \leq \sqrt{2} + 2 \int_{\sqrt{2}}^\infty z^{-2} \, dz = 2^\frac{3}{2}.
\]
For the first term on the right-hand side of (25), as $a_+ \leq A$ we have
\[
\Pr \left( a_+(Y_t) \geq \frac{\log z}{2np} \right) = 0 \text{ if } z > e^{2npA}.
\]
Note $2npA = (2np\|a\|_{\phi_n})^{\frac{1}{r-1}} \cdot 2^\frac{1}{r-1}$. Using the fact that $\mathbb{E}[Z] = \int_0^\infty \Pr(Z \geq z) \, dz$ for a non-negative random variable $Z$, we may deduce from (25), (26), and (27) that
\[
\mathbb{E}^{\mu \otimes Q}[m(X_t, X_{t+1})^{np}] \leq \int_0^\infty \Pr(e^{np|a|Y_t}) \geq z) \, dz \\
\leq e^{(2np\|a\|_{\phi_n})^{\frac{1}{r-1}}} + 2^\frac{3}{2}.
\]
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 $n^{\frac{1}{r-1}} = o((\beta + c)^{-n})$ for any $\beta \in (0, 1)$ and $c \in (0, 1 - \beta)$. \hfill \Box

**Proof of Lemma 2.1.** We first show $\mathbb{D}$ is a bounded linear operator on $L^{\phi_s}$ 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 $q \in (0, 1)$. By applying Jensen’s inequality, definition of $\tilde{\mathbb{E}}$ from (3), and Hölder’s inequality with $p^{-1} + q^{-1} = 1$, we obtain
\[
\mathbb{E}^{\mu} \left[ e^{\left| f(X_t)/(q^{\frac{1}{p}}\|f\|_{\phi_s})^\alpha \right|} \right] = \mathbb{E}^{\mu} \left[ e^{q^{-1}\tilde{\mathbb{E}}f(X_t)/\|f\|_{\phi_s}} \right] \\
\leq \mathbb{E}^{\mu \otimes Q} \left[ m(X_t, X_{t+1})e^{q^{-1}|f(X_{t+1})/\|f\|_{\phi_s}|} \right] \\
\leq \mathbb{E}^{\mu \otimes Q} \left[ m(X_t, X_{t+1})^p \right]^\frac{1}{p} \mathbb{E}^{\mu} \left[ e^{\left| f(X_t)/\|f\|_{\phi_s} \right|^\alpha} \right] \frac{1}{p} \\
\leq 2^\frac{1}{p} \mathbb{E}^{\mu \otimes Q} \left[ m(X_t, X_{t+1})^p \right]^\frac{1}{p},
\]
where the final line uses definition of $\|\cdot\|_{\phi_s}$. Note all moments of $m$ are finite under condition
(4). It follows by Lemma B.1 and definition of the operator norm $\|D\|_{L^\phi}$ that

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

That $D : E^\phi \rightarrow E^\phi$ may be deduced similarly. Boundedness of $D$ on $E^\phi$ now follows because $E^\phi$ is a closed linear subspace of $L^\phi$.

We use Lemma B.4 to establish the spectral radius condition. We prove the result for the spaces $L^\phi$; the results for $E^\phi$ follow because $E^\phi$ is a closed linear subspace of $L^\phi$. First consider the case with $s > 1$. Fix $p, q \in (1, \infty)$ with $p^{-1} + q^{-1} = 1$. For any $f \in L^\phi$ with $\|f\|_{\phi} > 0$, by two applications of Jensen’s inequality we have

$$E^\mu \left[e^{\|f(X_t)\|/(q^{\frac{1}{p}}(\beta^{\frac{s-1}{p}})^n)\}}\right] = E^\mu \left[e^{\beta^n q^{-1}[E^{\mu}(f(X_t))]^{\frac{1}{p}}\|f\|_{\phi}^n}\right]$$

$$\leq E^\mu \left[e^{\beta^n q^{-1}[E^{\mu}(f(X_t))]^{\frac{1}{p}}\|f\|_{\phi}^n}\right]^{\beta^n} \leq E^\mu \left[E^\mu g(X_t)\right]^{\beta^n},$$

where $g(x) = \exp(q^{-1}[f(x)]^{\|f\|_{\phi}^n})$. By Hölder’s inequality,

$$E^\mu \left[E^\mu \left[E^\mu g(X_t)\right]^{\beta^n}\right] = E^{\mu \otimes Q}[m(X_t, X_{t+1}) \cdots m(X_{t+n-1}, X_t)g(X_t)]$$

$$\leq E^{\mu \otimes Q}[m(X_t, X_{t+1}) \cdots m(X_{t+n-1}, X_t)]^{\frac{1}{p}} E^\mu \left[g(X_t)\right]^{\frac{1}{q}}$$

$$\leq E^{\mu \otimes Q}[m(X_t, X_{t+1})^{np}]^{\frac{1}{np}} \cdots E^{\mu \otimes Q}[m(X_{t+n-1}, X_t)^{np}]^{\frac{1}{np}} E^\mu \left[g(X_t)\right]^{\frac{1}{q}}$$

$$= E^{\mu \otimes Q}[m(X_t, X_{t+1})^{np}]^{\frac{1}{p}} E^\mu \left[g(X_t)\right]^{\frac{1}{q}}.$$

It follows by Lemma B.4, and definition of $g$ and $\|\cdot\|_{\phi}$ that

$$E^\mu \left[E^\mu g(X_t)\right] \leq E^{\mu \otimes Q}[m(X_t, X_{t+1})^{np}]^{\frac{1}{p}} E^\mu \left[e^{f(X_t)\|f\|_{\phi}^n}\right]^{\frac{1}{q}}$$

$$\leq 2^{\frac{1}{p}} C e^{(\beta+c)^n}$$

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

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

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

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

and therefore $\rho(D; L^\phi) \equiv \lim_{n \to \infty} \|D^n\|_{L^\phi}^{\frac{1}{n}} \leq \beta^{\frac{s-1}{s}} < 1.$
Now consider the case with $s = 1$. Let $c$ be as in Lemma B.4. Fix any $\epsilon \in (0, 1)$ and note that $\beta < \beta + \epsilon c < \beta + c < 1$. For any $f \in L^\phi_1$ with $\|f\|_\phi > 0$, we have:

$$
\mathbb{E}^\mu \left[ e^{\|D^n f(X_t)\|/(q\beta^n(\beta+\epsilon c)^{-n} \|f\|_\phi)} \right] = \mathbb{E}^\mu \left[ e^{(\beta+\epsilon c)^n q^{-1} \mathbb{E}^\mu f(X_t)/\|f\|_\phi)} \right]
\leq \mathbb{E}^\mu \left[ e^{q^{-1} \mathbb{E}^\mu f(X_t)/\|f\|_\phi)} \right]^{(\beta+\epsilon c)^n}
\leq \mathbb{E}^\mu \left[ \mathbb{E}^\mu g(X_t) \right]^{(\beta+\epsilon c)^n},
$$

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

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

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

$$
\|D^n\|_\phi \leq \left( \left( (2^{\frac{1}{q}}Ce^{(\beta+c)^{-n}})^{(\beta+\epsilon c)^n} - 1 \right) \lor 1 \right) q \left( \frac{\beta}{\beta + \epsilon c} \right)^n,
$$

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

B.3 Proofs for Section 3

Proof of Theorem 3.1. We verify the conditions of Proposition 2.1. For existence, Lemma B.5 shows $T$ is a continuous, monotone, and convex operator on $E^\phi_s$ 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 $h(x) = \mathbb{E}^Q e^{\frac{\alpha}{\beta} u(X_t, X_{t+1}) | X_t = x}$. 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 $x \mapsto e^{x/c}$ and $x \mapsto e^{(\log x)/c}$ for $c \in (0, 1)$, we obtain

$$
\mathbb{E}^\mu \left[ e^{\bar{v}(X_t)/(\beta c)\tau} \right] = \mathbb{E}^\mu \left[ \exp \left( (1 - \beta) \sum_{n=0}^\infty \beta^n \log \left( \left| (\mathbb{E}^Q)^n h(X_t) / c \right| \right) \right) \right] \\
\leq (1 - \beta) \sum_{n=0}^\infty \beta^n \mathbb{E}^\mu \left[ \exp \left( \log \left( (\mathbb{E}^Q)^n h(x) / c \right) \right) \right] \\
\leq (1 - \beta) \sum_{n=0}^\infty \beta^n \mathbb{E}^\mu \mathbb{Q} \left[ e^{\bar{v}(X_{t+n+1})} \right] \\
= \mathbb{E}^\mu \mathbb{Q} \left[ e^{\bar{v}(X_{t+n+1})} \right] < \infty.
$$

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

We now show that $T\bar{v} \leq \bar{v}$. By Holder’s inequality we first have

$$
T\bar{v}(X_t) \leq \beta \log \left( \mathbb{E}^Q \left[ e^{\bar{v}(X_{t+1})/\beta} | X_t \right] \beta \mathbb{E}^Q \left[ e^{\bar{v}(X_{t+1})/\beta} | X_t \right]^{1-\beta} \right) \\
= \beta^2 \log \mathbb{E}^Q \left[ e^{\bar{v}(X_{t+1})/\beta} | X_t \right] + (1 - \beta) \beta \log h(X_t).
$$

By Lemma B.2, we may deduce

$$
\log \mathbb{E}^Q \left[ e^{\bar{v}(X_{t+1})/\beta} | X_t \right] = \log \mathbb{E}^Q \left[ \prod_{n=0}^{\infty} \left( \mathbb{E}^Q h(X_{t+1}) \right)^{(1-\beta)\beta^n} | X_t \right] \\
\leq \log \left( \prod_{n=0}^{\infty} \mathbb{E}^Q \left[ \left( \mathbb{E}^Q h(X_{t+1}) \right)^{(1-\beta)\beta^n} | X_t \right] \right) \\
= (1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} \log \left( \mathbb{E}^Q \left[ h(X_t) \right] \right).
$$

Substituting (29) into (28) yields $T\bar{v} \leq \bar{v}$.

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

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

Therefore,

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

for each $n \geq 1$, where $h_1(x) = \beta \mathbb{E}^Q [\alpha u(X_t, X_{t+1}) | X_t = x]$. Note also that $\|\beta \mathbb{E}^Q\|_{E^{\phi_r}} = \beta$ and $\rho(\beta \mathbb{E}^Q; E^{\phi_r}) = \beta$ (see Section 2.3), and so we obtain

$$
\liminf_{n \to \infty} T^n \bar{v} \geq (1 - \beta \mathbb{E}^Q)^{-1} h_1 \in E^{\phi_r}.
$$
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.5 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.6. Uniqueness in $E^{\phi_s}$ with $s \in (1, r]$ follows by Proposition 2.1(ii). That $v$ is the smallest and unique stable fixed point in $E^{\phi_s}$ follows by Corollary 2.1.

**Lemma B.5.** Let condition (7) hold. Then: $\mathbb{T}$ is a continuous, monotone and convex operator on $E^{\phi_s}$ for each $1 \leq s \leq r$.

**Proof of Lemma B.5.** Fix any $1 \leq s \leq r$. 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:

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

which is finite for any $f \in E^{\phi_s}$ under (7). 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 $E_f$ denote the distorted conditional expectation operator from (8) 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,

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

because $c = 2^{1/s}\|g\|_{\phi_s}$. Finiteness of $\mathbb{E}^\mu\mathbb{E}_f[m_f(X_t, X_{t+1})^2]$ holds for any $f \in E^{\phi_s}$ under (7).
To see this, by several applications of the Cauchy–Schwarz and Jensen inequalities, we have

\[ E^{\mu \otimes Q}[m_f(X_t, X_{t+1})^2] = E^{\mu \otimes Q} \left[ \left( \frac{e^{f(X_{t+1})+\alpha u(X_t, X_{t+1})}}{E^{\phi_Q}[e^{f(X_{t+1})+\alpha u(X_t, X_{t+1})}|X_t]} \right)^2 \right] \]

\[ \leq E^{\mu \otimes Q} \left[ e^{4f(X_{t+1})+\alpha u(X_t, X_{t+1})} \right] \]

\[ \leq E^{\mu} \left[ e^{8|\alpha u(X_t, X_{t+1})|} \right]^{1/2} E^{\mu \otimes Q} \left[ e^{8|\alpha u(X_t, X_{t+1})|} \right]^{1/2}, \]

which is finite for any \( f \in E^{\phi_s} \) under (7). 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

\[ E^{\phi_Q} \left[ e^{r(v_1(X_{t+1})+\alpha u(X_t, X_{t+1}))+|1-\tau)(v_2(X_{t+1})+\alpha u(X_t, X_{t+1}))|} \right]_{X_t = x} \]

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

**Lemma B.6.** Let condition (7) hold with \( r > 1 \) and fix any \( v \in E^{\phi_{r'}} \) with \( r' > 1 \). Then: for each \( s \geq 1 \), \( D_v \) is a continuous linear operator on \( E^{\phi_s} \) with \( \rho(D_v; E^{\phi_s}) < 1 \).

**Proof of Lemma B.6.** We verify condition (4) from Lemma 2.1. The log change-of-measure is

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

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

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

by Jensen’s inequality. The right-hand side is finite by condition (7). Therefore,

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

for any \( c \in (0, 1] \) and hence for any \( c > 0 \) (see Remark B.1), verifying (4). \( \square \)

**Lemma B.7.** Let \( Y = |Z| \) with \( Z \sim N(0, 1) \). Then for \( a > 0 \) and \( r \in [1, 2) \), we have

\[ E \left[ \exp \left( \frac{Y^r}{a^r} \right) \right] \leq \frac{\sqrt{2}}{\sqrt{\pi}} \left( \frac{2}{a^r} \right)^{\frac{1}{2r}} \exp \left( 2 \frac{1}{a^r} \right) + \left( 4 \frac{2^{\frac{1}{r}}}{a^r} \right)^{\frac{1}{2r}} + 2\sqrt{\pi}. \]
Proof of Lemma B.7. First write

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

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

B.4 Proofs for Section 4

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 $\mu$ is the stationary distribution of $\hat{X}_t$ under $\hat{Q}$. For $v \in E^\varphi_{\hat{\xi}}$, define

$$
m^\Pi_{\varphi} (\xi_t, \hat{\xi}_t) = \frac{E^{Q_\varphi} \left[ e^{\frac{\varphi}{\sigma^2} v(\Xi(\hat{\xi}_t, \varphi_{t+1}) + \alpha u(\varphi_{t+1}))} \right]_{\xi_t, \hat{\xi}_t}^{\hat{\xi}_t}}{E^{\Pi_\xi} \left[ E^{Q_\varphi} \left[ e^{\frac{\varphi}{\sigma^2} v(\Xi(\hat{\xi}_t, \varphi_{t+1}) + \alpha u(\varphi_{t+1}))} \right]_{\xi_t, \hat{\xi}_t}^{\hat{\xi}_t} \right]},
$$

$$
m^Q_{\varphi} (\xi_t, \hat{\xi}_t, \varphi_{t+1}) = \frac{E^{Q_\varphi} \left[ e^{\frac{\varphi}{\sigma^2} v(\Xi(\hat{\xi}_t, \varphi_{t+1}) + \alpha u(\varphi_{t+1}))} \right]_{\xi_t, \hat{\xi}_t}^{\hat{\xi}_t}}{E^{Q_\varphi} \left[ e^{\frac{\varphi}{\sigma^2} v(\Xi(\hat{\xi}_t, \varphi_{t+1}) + \alpha u(\varphi_{t+1}))} \right]_{\xi_t, \hat{\xi}_t}^{\hat{\xi}_t}}.
$$

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

$$
E^\Pi_\xi f(\hat{\xi}) = E^{\Pi_\xi} \left[ m^\Pi_{\varphi}(\xi_t, \hat{\xi}_t) f(\xi_t, \hat{\xi}_t) \right]_{\xi_t = \hat{\xi}_t},
$$

$$
E^Q_{\varphi} f(\hat{\xi}) = E^{\Pi_\xi} \left[ m^Q_{\varphi}(\xi_t, \hat{\xi}_t, \varphi_{t+1}) f(\xi_t, \hat{\xi}_t, \varphi_{t+1}) \right]_{\xi_t = \hat{\xi}_t, \hat{\xi}_t = \hat{\xi}_t}.
$$
The subgradient of $T$ at $v$ is the composition of these two distorted conditional expectations, discounted by $\beta$:

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

(30)

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

**Proof of Theorem 4.1.** We verify the conditions of Proposition 2.1. Lemma B.8 shows that $T$ is a continuous, monotone, and convex operator on $E^\phi_{\hat{\xi}}$ 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 $g_1(X_t) = \exp(\frac{\alpha u}{(1-\beta)\vartheta} u(\varphi_t))$. For any $c > 0$, by Jensen’s inequality we may deduce

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

where $g_1^r(X_t) = \exp(\frac{\alpha u}{(1-\beta)\vartheta} u(\varphi_t))$. As $u \in E^\phi_{\hat{\xi}}$, the right-hand side of the preceding display is finite and so $\bar{v} \in E^\phi_{\hat{\xi}}$.

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

$$T\bar{v}(\hat{\xi}) = \beta \log \mathbb{E}^{\hat{Q}} \left[ e^{\frac{\alpha u}{(1-\beta)\vartheta} u(\varphi_{t+1})} \Big| \xi_t, \hat{\xi}_t \Big] \hat{\xi}_t = \hat{\xi} \right]$$

$$\leq \beta \log \mathbb{E}^{\hat{Q}} \left[ e^{\frac{\alpha u}{(1-\beta)\vartheta} u(\varphi_{t+1})} \Big| \hat{\xi}_t = \hat{\xi} \right]$$

$$\leq \beta^2 \log \mathbb{E}^{\hat{Q}} \left[ e^{\frac{\alpha u}{(1-\beta)\vartheta} u(\varphi_{t+1})} \Big| \hat{\xi}_t = \hat{\xi} \right] + \beta (1 - \beta) \log \mathbb{E}^{\hat{Q}} \left[ e^{\frac{\alpha u}{(1-\beta)\vartheta} u(\varphi_{t+1})} \Big| \hat{\xi}_t = \hat{\xi} \right].$$

By Lemma B.2, we may deduce

$$\log \mathbb{E}^{\hat{Q}} \left[ e^{\frac{\alpha u}{(1-\beta)\vartheta} u(\varphi_{t+1})} \Big| \hat{\xi}_t = \hat{\xi} \right] \leq (1 - \beta) \sum_{n=1}^{\infty} \beta^{n-1} \log \left( \left( \mathbb{E}^{\hat{Q}} \right)^{n+1} g_1(\hat{\xi}) \right),$$

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

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

\[
T \bar{v}(\hat{\xi}) = \beta \log E^{\Pi_{\xi}} \left[ E^{Q_{r}} \left[ e^{\frac{\theta \Pi_{\xi}(\hat{\xi}_{t},\varphi_{t+1})}{\beta} + \alpha u(\varphi_{t+1})} \right] \bar{v} \bigg| \hat{\xi}_{t} = \hat{\xi} \right] \\
\leq \frac{\theta}{\beta} \log E^{Q} \left[ e^{\frac{\theta \Pi_{\xi}(\hat{\xi}_{t+1}) + \alpha u(\varphi_{t+1})}{\beta}} \right] \\
\leq \frac{\theta}{\beta} \log E^{Q} \left[ e^{\frac{\theta \Pi_{\xi}(\hat{\xi}_{t})}{\beta}} \right] + \frac{\theta}{\beta}(1 - \beta) \log E^{Q} \left[ e^{\frac{\alpha u(\varphi_{t+1})}{\beta}} \right] \hat{\xi}_{t} = \hat{\xi}.
\]

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

To show that the sequence of iterates \( T^{n} \bar{v} \) is bounded from below, first note that for any \( f \in E_{\xi}^{\phi_{r}} \), we have

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

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

\[
T^{n} \bar{v}(\hat{\xi}) \geq \left( \beta E^{Q} \right)^{n} \bar{v}(\hat{\xi}) + \sum_{i=0}^{n-1} \left( \beta E^{Q} \right)^{i} g_{3}(\hat{\xi})
\]

where \( g_{3}(\hat{\xi}) = \beta E^{Q}[\alpha \frac{\theta}{\beta} u(\varphi_{t+1})|\hat{\xi}_{t} = \hat{\xi}] \in E_{\xi}^{\phi_{r}} \). Note also that \( \rho(\beta E^{Q}, E^{\phi_{r}}) = \beta \) (see Section 2.3), hence \( \liminf_{n \to \infty} T^{n} \bar{v} \geq (1 - \beta E^{Q})^{-1} g_{3} \in E^{\phi_{r}} \). This completes the proof of existence.

For uniqueness, \( v \) is necessarily a fixed point of \( T : E_{\xi}^{\phi_{s}} \to E_{\xi}^{\phi_{s}} \) for each \( 1 \leq s \leq r \). The subgradient \( \mathbb{D}_{v} \) is monotone. Lemma B.9 shows \( \mathbb{D}_{v} : E_{\xi}^{\phi_{s}} \to E_{\xi}^{\phi_{s}} \) is bounded and \( \rho(\mathbb{D}_{v}, E_{\xi}^{\phi_{s}}) < 1 \) for \( s \in [1, r] \). Uniqueness follows by Proposition 2.1(ii) and Corollary 2.1.

**Lemma B.8.** Let condition (13) hold. Then: \( T \) is a continuous, monotone, and convex operator on \( E_{\xi}^{\phi_{s}} \) for each \( 1 \leq s \leq r \).

**Proof of Lemma B.8.** Fix \( s \in [1, r] \). We first show \( E^{\mu}[\exp(|T f(\hat{\xi}_{t})|/(\beta c)^{s})] < \infty \) holds for each \( f \in E_{\xi}^{\phi_{s}} \) and \( c \in (0, \frac{\theta}{\beta} \land 1] \). By convexity of \( x \mapsto e^{((\log x)/c)^{s}} \) for \( c \in (0, 1] \) and Jensen’s
inequality,
\[
\mathbb{E}^\mu \left[ \exp \left( \frac{T_f(\hat{\xi}_t)}{\beta c} \right)^s \right] = \mathbb{E}^\mu \left[ \exp \left( \frac{1}{c^s} \log \mathbb{E}^{\Pi \xi} \left[ \mathbb{E}^{Q \varphi} \left[ e^{\frac{\theta}{c} f(\Xi(\hat{\xi}_t, \varphi_{t+1})) + \alpha u(\varphi_{t+1})} \bigg| \xi_t, \hat{\xi}_t \right]^s \right] \right) \right] 
\leq \mathbb{E}^\mu \left[ \mathbb{E}^{\Pi \xi} \left[ \exp \left( \frac{1}{c^s} \log \mathbb{E}^{Q \varphi} \left[ e^{\frac{\theta}{c} f(\Xi(\hat{\xi}_t, \varphi_{t+1})) + \alpha u(\varphi_{t+1})} \bigg| \xi_t, \hat{\xi}_t \right]^s \right] \right] \right] 
\leq \mathbb{E}^\mu \left[ \mathbb{E}^{\Pi \xi} \left[ \mathbb{E}^{Q \varphi} \left[ \exp \left( \frac{1}{c^s} \log e^{\frac{\theta}{c} f(\Xi(\hat{\xi}_t, \varphi_{t+1})) + \alpha u(\varphi_{t+1})} \bigg| \xi_t, \hat{\xi}_t \right)^s \right] \right] \right] 
= \mathbb{E}^{\mu \otimes \Pi \xi \otimes Q \varphi} \left[ \exp \left( \frac{1}{c^s} \log e^{\frac{\theta}{c} f(\Xi(\hat{\xi}_t, \varphi_{t+1})) + \alpha u(\varphi_{t+1})} \bigg| \xi_t, \hat{\xi}_t \right)^s \right]
\]

which is finite because \( f \in E_\xi^{\phi_s} \) and \( u \in E_\varphi^{\phi_s} \). It follows by Remark B.1 that \( \mathbb{T} : E_\xi^{\phi_s} \to E_\xi^{\phi_s} \).

For continuity, fix \( f \in E_\xi^{\phi_s} \). Take \( g \in E_\xi^{\phi_s} \) with \( 0 < \|g\|_{\phi_s} \leq 2^{-1/s}(1 + \frac{\theta}{c}) \) and set \( c = 2^{1/s}\|g\|_{\phi_s} \). Note
\[
\mathbb{T}(f + g)(\hat{\xi}) - T_f(\hat{\xi}) = \beta \log \left( \mathbb{E}^{\Pi \xi} \left[ \mathbb{E}^{Q \varphi} \left[ e^{\frac{\theta}{c} g(\Xi(\hat{\xi}_t, \varphi_{t+1}))} \bigg| \xi_t, \hat{\xi}_t \right] \right] \right) 
\]

By similar arguments to the above, we may deduce
\[
\mathbb{E}^\mu \left[ \exp \left( \frac{\mathbb{T}(f + g)(\hat{\xi}_t) - T_f(\hat{\xi}_t)}{\beta c} \right)^s \right] \leq \mathbb{E}^\mu \left[ \mathbb{E}^{\Pi \xi} \left[ \mathbb{E}^{Q \varphi} \left[ \exp \left( \frac{1}{c^s} \log e^{\frac{\theta}{c} g(\Xi(\hat{\xi}_t, \varphi_{t+1}))} \bigg| \xi_t, \hat{\xi}_t \right)^s \right] \right] \right] 
\leq \mathbb{E}^{\mu \otimes Q} \left[ m_f(\xi_t, \hat{\xi}_t, \varphi_{t+1}) \exp \left( \frac{1}{c^s} \log e^{\frac{\theta}{c} g(\Xi(\hat{\xi}_t, \varphi_{t+1}))} \bigg| \xi_t, \hat{\xi}_t \right)^s \right] \left[ m_f(\xi_t, \hat{\xi}_t, \varphi_{t+1}) \right]^{1/2} \mathbb{E}^\mu \left[ \exp \left( 2\|g(\hat{\xi}_t)\|_{\phi_s} \right)^{1/2} \right] \left[ m_f(\xi_t, \hat{\xi}_t, \varphi_{t+1}) \right]^{1/2} \left[ m_f(\xi_t, \hat{\xi}_t, \varphi_{t+1}) \right]^{1/2},
\]

because \( c = 2^{1/s}\|g\|_{\phi_s} \). The expectation on the right-hand side is finite because \( f \in E_\xi^{\phi_s} \) and \( u \in E_\varphi^{\phi_s} \). It follows by Lemma B.1 that \( \|\mathbb{T}(f + g) - T_f\|_{\phi_s} \to 0 \) 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.9.** Let condition (13) hold. Fix any \( v \in E_\xi^{\phi_{s'} \otimes \Phi_{s'}} \) with \( r' > 1 \). Then: for each \( s \geq 1 \), \( \mathbb{D}_v \) is a continuous linear operator on \( E_\xi^{\phi_s} \) with \( \rho(\mathbb{D}_v; E_\xi^{\phi_s}) < 1 \).

**Proof of Lemma B.9.** It suffices to verify the conditions of Lemma 2.1. Note that the process
\[ \hat{\xi} = \{\hat{\xi}_t\}_{t \in \mathcal{T}} \] is a stationary Markov process (this follows from our maintained assumptions that learning is in a steady state and the conventional hidden Markov structure on \( X \)). By iterated expectations, we may rewrite the subgradient from (30) as

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

where \( \hat{m}_v(\hat{\xi}_t, \hat{\xi}_{t+1}) \) denotes the conditional expectation of \( m_v(\xi_t, \xi_{t+1}) \) 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.6 for any \( v \in E_{\hat{\xi}}^{\phi, \iota} \) with \( \iota' > 1 \).

\[ \square \]

### B.5 Proof for Section 5

**Proof of Theorem 5.1.** In view of the discussion preceding Theorem 5.1 and Lemma B.10, it suffices to show that \( \tilde{v} \in \hat{E}^{\phi, \iota} \) and that \( \mathbb{T}\tilde{v} \leq \tilde{v} \). By (19), convexity of \( x \mapsto e^{(|\log x|/c)^\iota'} \) for \( c \in (0, 1] \), and two applications of Jensen’s inequality, for any \( c \in (0, 1] \) we have

\[
\mathbb{E}^{\tilde{\mu}} \left[ e^{v(X_t)/c} \right] = \mathbb{E}^{\tilde{\mu}} \left[ e^{\left| \log \left( (1-\beta) \sum_{n=0}^{\infty} \left( \beta \lambda \frac{1}{n} \frac{\hat{\mu}}{e} \left( (\nu - \frac{1}{c})(X_t) \right) \right) \right) /c} \right]
\]

\[
\leq (1 - \beta \lambda \frac{1}{c}) \sum_{n=0}^{\infty} \left( \beta \lambda \frac{1}{n} \frac{\hat{\mu}}{e} \right) \mathbb{E}^{\tilde{\mu}} \left[ e^{\left( \log \left( (1-\beta) \sum_{n=0}^{\infty} \left( \beta \lambda \frac{1}{n} \frac{\hat{\mu}}{e} \left( (\nu - \frac{1}{c})(X_t) \right) \right) \right) /c} \right]
\]

\[
= \mathbb{E}^{\tilde{\mu}} \left[ e^{\left( \log \left( \log \left( (1-\beta) \sum_{n=0}^{\infty} \left( \beta \lambda \frac{1}{n} \frac{\hat{\mu}}{e} \left( (\nu - \frac{1}{c})(X_t) \right) \right) \right) /c \right)} \right] < \infty
\]

by condition (18), with the final equality because \( \tilde{\mu} \) is the stationary distribution corresponding to \( \tilde{E} \). It follows by Remark B.1 that \( \tilde{v} \in E^{\phi, \iota} \).

To see that \( \mathbb{T}\tilde{v} \leq \tilde{v} \), first observe that by Jensen’s inequality (as \( \kappa < 0 \), we have

\[ \mathbb{T}\tilde{v}(x) = \log \left( (1 - \beta) \nu(x)^{-\frac{1}{\kappa}} + \beta \lambda \frac{1}{\kappa} \mathbb{E}^{\tilde{\mu}} \left[ (1 - \beta) \sum_{n=0}^{\infty} \left( \beta \lambda \frac{1}{n} \frac{\hat{\mu}}{e} \left( (\nu - \frac{1}{c})(X_{t+1}) \right) \right)^{\kappa} \bigg| X_t = x \right] \right) \]

\[ \leq \log \left( (1 - \beta) \nu(x)^{-\frac{1}{\kappa}} + \beta \lambda \frac{1}{\kappa} \mathbb{E}^{\tilde{\mu}} \left[ (1 - \beta) \sum_{n=0}^{\infty} \left( \beta \lambda \frac{1}{n} \frac{\hat{\mu}}{e} \left( (\nu - \frac{1}{c})(X_{t+1}) \right) \bigg| X_t = x \right] \right) \]

\[ = \log \left( (1 - \beta) \nu(x)^{-\frac{1}{\kappa}} + (1 - \beta) \sum_{n=1}^{\infty} \left( \beta \lambda \frac{1}{n} \frac{\hat{\mu}}{e} \left( (\nu - \frac{1}{c})(x) \right) \right) \right) \]

\[ = \tilde{v}(x). \]

Existence now follows by Proposition 2.1(i).  

\[ \square \]
Lemma B.10. Let condition (18) hold. Then the operator $\mathbb{T}$ from (17) is a continuous, monotone operator on $\hat{E}^{\phi_s}$ for each $1 \leq s \leq r$.

Proof of Lemma B.10. Fix any $s \in [1, r]$. We first show that $\mathbb{E}^\mu[e^{T f(X_t)/c_s}] < \infty$ holds for any $f \in \hat{E}^{\phi_s}$ 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

$$
\mathbb{E}^{\hat{\mu}} \left[ e^{T f(X_t)/c_s} \right] = \mathbb{E}^{\hat{\mu}} \left[ \exp \left( \log \left( (1 - \beta) \mu(X_t) - \frac{1}{2} + \beta \frac{1}{\kappa} \mathbb{E}^{\hat{\mu}}[e^{\kappa f(X_{t+1})}|X_t] \right)/c_s \right) \right]
\leq \mathbb{E}^{\hat{\mu}} \left[ (1 - \beta) e^{(\log \mu(X_t)/(\kappa c_s))} + \beta \exp \left( \log \left( \mathbb{E}^{\hat{\mu}}[e^{\kappa f(X_{t+1})}|X_t] \right)/c_s \right) \right]
\leq \mathbb{E}^{\hat{\mu}} \left[ (1 - \beta) e^{(\log \mu(X_t)/(\kappa c_s))} + \beta \mathbb{E}^{\hat{\mu}} \left[ \exp \left( \log \left( \mathbb{E}^{\hat{\mu}}[e^{\kappa f(X_{t+1})}|X_t] \right)/c_s \right) \right] X_t \right]
= (1 - \beta) \mathbb{E}^{\hat{\mu}} \left[ e^{(\log \mu(X_t)/(\kappa c_s))} \right] + \beta \mathbb{E}^{\hat{\mu}} \left[ e^{(\log \mu(X_t)/(\kappa c_s))} \right] X_t
$$

where the right-hand side is finite under condition (18), and the final equality is because $\hat{\mu}$ is the stationary distribution under $\hat{\mathbb{E}}$. It follows by Remark B.1 that $\mathbb{T} : \hat{E}^{\phi_s} \to \hat{E}^{\phi_s}$.

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

$$
\mathbb{T}(f + h)(x) - \mathbb{T} f(x) = \log \left\{ \frac{(1 - \beta)\mu(x)^{-2} + \beta \frac{1}{\kappa} w(x) \mathbb{E}_f[e^{\kappa h(X_{t+1})}|X_t = x]}{(1 - \beta)\mu(x)^{-2} + \beta \frac{1}{\kappa} w(x)} \right\}
$$

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

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

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
\[
\mathbb{E} \tilde{\mu} \left[ e^{(\mathbb{T}(f+h)(X_t) - Tf(X_t))/c^*} \right] \\
= \mathbb{E} \tilde{\mu} \left[ \exp \left( \frac{1}{c} \log \left( \frac{1 - \beta}{} + \beta \lambda \frac{h}{X_t} w(X_t) \mathbb{E} f \left[ e^{\kappa h(X_{t+1})} | X_t \right] \right)^{\frac{1}{\kappa}} \right) \right] \\
\leq \mathbb{E} \tilde{\mu} \left[ \left( 1 - \beta \right) (X_t) - \frac{1}{\kappa} + \beta \lambda \frac{h}{X_t} w(X_t) \mathbb{E} f \left[ e^{\kappa h(X_{t+1})} | X_t \right] \right] \\
\leq \mathbb{E} \tilde{\mu} \left[ \left( 1 - \beta \right) (X_t) - \frac{1}{\kappa} + \beta \lambda \frac{h}{X_t} w(X_t) \right] \\
\leq \mathbb{E} \tilde{\mu} \left[ e^{4 \kappa f(X_t)} \right] \frac{1}{2} \mathbb{E} \tilde{\mu} \left[ e^{\frac{4 h(X_t)}{c}} \right] \frac{1}{2}.
\]

For \( h \in E^{b*} \) with \( \|h\|_{\phi_2} \leq \frac{1}{2} (1 \land |\kappa|^{-1}) \), setting \( c = 2 \|h\|_{\phi_2} \) we therefore have
\[
\mathbb{E} \mu \left[ e^{(\mathbb{T}(f+h)(X_t) - Tf(X_t))/(2\|h\|_{\phi_2})^*} \right] \leq \left( 2 \mathbb{E} \tilde{\mu} \left[ e^{4 \kappa f(X_t)} \right] \right) \frac{1}{2}.
\]
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 5.1. Immediate from Theorem 5.1 and Lemma B.3.

Proof of Theorem 5.2. The proof follows by similar arguments to Theorem 5.1, we list only the modifications here. First use (22) and the definition of \( \tilde{E} \) to rewrite \( \tilde{v} \) as
\[
\tilde{v}(x) = \log \left( (1 - \beta) \sum_{n=0}^{\infty} (\beta \lambda \frac{h}{X_t})^{n} \mathbb{E} Q \left[ e^{\kappa f(X_{t+n})} | X_t = x, t(x) \right] \right).
\]
Let \( b = (1 - \beta)(1 - \beta \lambda \frac{h}{X_t})^{-1} \). By (22), convexity of \( x \mapsto e^{|\log x|/c^*} \) for \( c \in (0, 1] \), and two applications of Jensen’s inequality, for any \( c \in (0, 1] \) we have
\[
\mathbb{E} \mu \left[ e^{\tilde{v}(X_t)/c^*} \right] \leq \mathbb{E} \mu \left[ (1 - \beta \lambda \frac{h}{X_t})^{\infty} \sum_{n=0}^{\infty} (\beta \lambda \frac{h}{X_t})^{n} e^{\log \left( \mathbb{E} Q \left[ e^{\kappa f(X_{t+n})} | X_t = x, t(x) \right] /c^* \right)} \right] \\
\leq (1 - \beta \lambda \frac{h}{X_t})^{\infty} \mathbb{E} \mu \left[ \mathbb{E} Q \left[ \exp \left( \log \left( b(1 - \beta \lambda \frac{h}{X_t})^{\infty} \right)^{\frac{1}{\kappa}} t(X_t) \right) /c^* \right) \right] .
\]
As \( \log \left( b(1 - \beta \lambda \frac{h}{X_t})^{\infty} \right)^{\frac{1}{\kappa}} \leq 4 \left(|(\log b)/c^*| + |\frac{n}{\kappa} - 1 | \log \left( 1 - \beta \lambda \frac{h}{X_t} \right) /c^*| \right) \),
we may use Hölder’s inequality, stationarity of $X$, and condition (21) to deduce that the right-hand side of the above display is finite. It follows by Remark B.1 that $\bar{v} \in E^{\phi_r}$.

The only other modification we require is to show that $T$ is a continuous operator on $E^{\phi_s}$ for each $1 \leq s \leq r$. By the first chain of inequalities in the proof of Lemma B.10, for any $f \in E^{\phi_s}$ and $c \in (0, 1 \land |\kappa|^{-1})$, we obtain

$$
E^\mu \left[ e^{\left| Tf(X_t)/c \right|^s} \right] \leq (1 - \beta)E^\mu \left[ e^{\left| \log (X_t)/(\kappa c) \right|^s} \right] + \beta E^\mu \left[ \exp \left( \left| \log \left( \lambda e^{\kappa f(X_{t+1})}/(\kappa c) \right) \right|^s \right) \right] X_t \right] \\
= (1 - \beta)E^\mu \left[ e^{\left| \log (X_t)/(\kappa c) \right|^s} \right] + \frac{\beta}{\lambda} E^\mu \otimes Q \left[ e^{\left| \log (X_{t+1})/(\kappa c) \right|^s} \right] e^{\left| (\log \lambda)/(\kappa c) + f(X_{t+1})/c \right|^s} .
$$

Finiteness of the right-hand side then follows by applying Hölder’s inequality and using stationarity of $X$ and condition (21). It follows by Remark B.1 that $T : E^{\phi_s} \to E^{\phi_s}$. The proof of continuity follows by a similar modification.

References

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.

Balbus, L. (2015). On non-negative recursive utilities in dynamic programming with non-linear aggregator and CES. SSRN working paper no. 2703975.

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.

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

Bliose, 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.

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.

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. and Ya. B. Rutickii (1961). *Convex Functions and Orlicz Spaces*. P. Noordhoff Ltd., Groningen.

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

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/.*

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.