Analysis of Forward Looking Models using Integrability Conditions[1]

Jaroslav Borovička,[a] Thomas J. Sargent[b] and John Stachurski[c]

[a, b] Department of Economics, New York University

[c] Research School of Economics, Australian National University

September 20, 2016

Abstract. In this paper we study existence and uniqueness of equilibria in forward looking models, focusing on fundamental solutions. The theory presented below can accommodate several classes of nonlinear models, bounded or unbounded solutions and a variety of applications. It covers some known results as special cases and opens the door to new ones. As well as sufficient conditions for existence and uniqueness of equilibria, we discuss methods of computation.

JEL Classifications: D81, G11
Keywords: Forward looking dynamics, asset pricing, time reversibility, projection

1. Introduction

One of the defining features of economic models—as opposed to, say, physical or biological models—is the prevalence of forward looking restrictions. Economic models have forward looking dynamics because economic agents have preferences over future outcomes. These concerns influence current actions, which in turn affect current outcomes. Forward looking models came to the fore with the work of authors such as Cagan (1956), Muth (1961) and Lucas (1972, 1976). They are central to the

[1]Much of this research was completed while the third author spent time at New York University and Singapore Management University. These visits were partially funded by the Alfred P. Sloan Foundation. We also thank Ippei Fujiwara, John Rust and Sean Meyn for helpful comments and suggestions.

modeling of a vast range of economic phenomena, from the determination of asset and commodity prices to exchange rates, inflation, interest rates, saving, borrowing, investment and consumption.

This modeling methodology gives rise to the problem of characterizing equilibria that satisfy the restrictions of forward looking models, as well as clarifying conditions under which they exist and are unique. The early literature focused on linear systems or linearized systems, with well-cited references including Taylor (1977), Blanchard and Kahn (1980), Hansen and Sargent (1980), Uhlig (1999), Klein (2000), Christiano (2002) and Sims (2002). A recent treatment based on Weiner-Hopf factorizations can be found in Al-Sadoon (2016).

Regarding existence and uniqueness of solutions for more general models, the current state of the field is that we have many specific results but still lack general theory. While some special classes of models have analytical solutions, moving beyond these models typically requires sophisticated fixed point arguments. Two seminal contributions are Lucas (1972) and Lucas (1978). More recent examples includes Labadie (1986), Lucas and Stokey (1987), Calin et al. (2005) and Brogueira and Schütze (2015). In these papers the fixed point arguments are tailored directly to the application in question.

Here we set down a first pass at a more general approach, including tools to analyze equilibria of forward looking systems encompassing many applications. We focus throughout on so-called fundamental solutions (as opposed to "bubble" solutions, which are also of interest but set aside in what follows). The theory developed below admits linear models, nonlinear models, bounded models, unbounded models and so on. It covers a number of known existence and uniqueness results across different applications and opens the door to new ones. As well as sufficient conditions for existence and uniqueness of equilibria, we provide some means of characterizing equilibria and methods of computation.

At the first stage, we restrict attention to discrete time models that take the form $Y_t = \mathbb{E}_t [A_{t+1} Y_{t+1}] + \varphi_t$, where the vector-valued stochastic process $\{Y_t\}$ is the endogenous object that we wish to solve for and $\{\varphi_t\}$ and A_{t+1} are exogenous random vector and matrix valued processes respectively. We call these models *random coefficient models*. If A_{t+1} is nonrandom, then the problem reduces to a linear rational expectations model of the type discussed above. A large variety of commonly used models can be expressed in this framework.

Despite the seemingly minimal departure from linearity, general results on existence and uniqueness of solutions to such models are still thin on the ground. It is of course well known that one can potentially "solve" the model by iterating forwards, but, apart from some special cases, the question of whether the resulting sequence converges is nontrivial and insufficiently addressed.

Here we tackle the problem from two complementary angles. First, we study convergence of the random series produced by forward iteration. We do so by adapting the stability theory of Kesten random coefficient processes with stationary and ergodic coefficients (see Kesten (1973) and Brandt (1986)) to forward looking models.[1] This provides relatively simple and general sufficient conditions for convergence of the random series.

Unfortunately, the conditions discussed above do not imply that the limiting sums have finite expectation. To address integrability we use an operator theoretic approach. Because of the need to handle applications that include unbounded solutions, we resist the traditional method of embedding our equilibrium problem in a space of bounded functions. Instead we replace boundedness with the weaker condition that at least one moment is finite.[2]

Working in a setting of integrable functions satisfying moment conditions turns out to have several advantages beyond the ability to accommodate unbounded solutions. For example, if we have finite second moments we can embed our problem in a Hilbert space of square integrable functions. This allows us to analyze contractions using relatively weak sufficient conditions that can be checked numerically, using all the analytical machinery provided by complete (and, in our case, separable) inner product spaces.

In addition, the space of square integrable functions supports a powerful approximation theory based around orthogonal projections that can be exploited in the discussion of computation. As we demonstrate below, this theory is particularly effective when the orthonormal basis used to represent and decompose functions in the space is precisely matched with the state process that drives stochastic outcomes in the model. In fact, in many quantitative studies, economists choose Markov processes

[1]The theory of traditional backward looking Kesten processes has been applied productively to a several economic problems, including the tail properties of wealth distributions (see, e.g., Benhabib et al. (2011)).

[2]This condition cannot be weakened further, since the conditional expectation in the statement of the forward looking restriction is not well defined without at least one finite moment.

for the state that have the property of time reversibility, and, if we match orthonormal bases with these state processes, we can greatly simplifying a range of numerical computations.

Having treated random coefficient models, we then turn to more general models of the form $Y_t = \mathbb{E}_t \, G(X_t, X_{t+1}, Y_{t+1})$ where G is a given function and Y_t is now allowed to be either finite or infinite dimensional. Once again our approach is based around integrability conditions. We develop sufficient conditions for existence and uniqueness of solutions and discuss computation. It turns out that many of the advantages of the integrability based approach continue to hold for these more general models.

In terms of existing literature, the theoretical and computational components of parts of our work are related to Tauchen and Hussey (1991), who study a class of asset pricing problems that can be expressed in operator form as Fredholm integral equations. The problems in section 2 have this property. Unlike Tauchen and Hussey (1991), we frame our problem in a space of integrable functions, and hence permit unbounded solutions and unbounded state spaces. We focus on existence and uniqueness of solutions, while Tauchen and Hussey (1991) focus on computation.

There have been a variety of studies attempting to solve the kinds of functional equations that arise in economic modeling in the setting where solutions can be unbounded. One line of attack is based around weighted supremum norms with an appropriately chosed weighting function. Important examples can be found in Epstein and Zin (1989), Boyd (1990) and Alvarez and Stokey (1998). For some highly nonlinear problems, this approach is idea. All of these papers consider specific applications, and the weighting function is tailored to the structure of the problem.[3]

The paper is structured as follows: Section 2 treats random coefficient models. Section 3 discusses sufficient conditions for the results in section 2. Section 4 covers computation. More general nonlinear models are treated in section 5, while section 6 covers further extensions. Some proofs are deferred to the appendix.

[3]Another line of research tackles unboundedness functional equations via the local contraction approach. See, for example, Rincón-Zapatero and Rodríguez-Palmero (2003), Martins-da Rocha and Vailakis (2010) and Matkowski and Nowak (2011). One of the motivations of this line of work is to deal with dynamic programming problems that are unbounded both above and below. The integrability based methods used here can also handle functions that are unbounded above and below, although the class of problems we treat are different.

2. Random Coefficient Models

In some models, interaction between the endogenous and exogenous processes comes only through random additive and multiplicative components. This is the setting where it is potentially feasible to solve the model by iterating forward in time. In this section we provide sufficient conditions for convergence of both the random series and the current expectation of that series.

2.1. **Set Up.** Consider a random coefficient model of the form

$$Y_t = \mathbb{E}_t\left[A_{t+1} Y_{t+1}\right] + \varphi_t, \qquad t \in \mathbb{Z}, \tag{1}$$

where $\{Y_t\}$ is an endogenous sequence of random vectors and $\{\varphi_t\}$ and $\{A_t\}$ are random and exogenous. The random sequences are defined on a fixed probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and adapted to a filtration $\{\mathscr{F}_t\}$. The operator $\mathbb{E}_t := \mathbb{E}[\,\cdot \mid \mathscr{F}_t]$ is vector-valued expectation conditional on time t information $\mathscr{F}_t$, and the equality in (39) is understood as holding $\mathbb{P}$-almost surely.

Since many recent studies deal with infinite dimensional state variables (distributions of assets or productivity, location, etc.), and since accommodating such states causes no additional complications to theory or notation, we allow for this possibility. Hence in what follows we assume that $\{Y_t\}$ and $\{\varphi_t\}$ take values in a separable Banach space $(\mathsf{Y}, \|\cdot\|)$ and that $\{A_t\}$ takes values in the set of bounded linear operators from Y into itself. In the standard case, and in the majority of our applications, $\mathsf{Y} = \mathbb{R}^m$ and each A_t should be understood as a random $m \times m$ matrix.[4]

Example 2.1. In the consumption-based asset pricing model of Lucas (1978), the price process $\{P_t\}$ of a claim to the dividend stream satisfies

$$P_t = \mathbb{E}_t\left[\beta \frac{u'(C_{t+1})}{u'(C_t)}(D_{t+1} + P_{t+1})\right], \tag{2}$$

where C_t is consumption, β is a discount factor, and u is utility. In equilibrium, $C_t = D_t$ for all t, where $\{D_t\}$ is an exogenous endowment process. The model (2) is a version of (1) with $A_{t+1} = \beta \frac{u'(C_{t+1})}{u'(C_t)}$ and $\varphi_t = \beta \mathbb{E}_t \frac{u'(C_{t+1})}{u'(C_t)} D_{t+1}$.

---

[4]Since Y is allowed to be any separable Banach space, the conditional expectation in (1) should be understood as a Bochner integral. When Y_t is an ordinary finite dimensional vector, this reduces to the usual notion of element-by-element integration. See section 11.8 of Aliprantis and Border (2007) for details. In the finite dimensional case, vectors are understood as column vectors unless otherwise stated.

Example 2.2. When dividends are nonstationary, one can study the price dividend ratio rather than the price. For example, suppose that $\mathbb{E}_t [M_{t+1} R_{t+1}] = 1$, where M_{t+1} is the pricing kernel and $R_{t+1} = (D_{t+1} + P_{t+1})/P_t$ is gross returns on an asset with stochastic cash flow $\{D_t\}$ and time t price P_t. $\{M_t\}$ and $\{D_t\}$ are exogenous. Rearranging and defining $Y_t := P_t/D_t$ as the price-dividend ratio, we can express this restriction as

$$Y_t = \mathbb{E}_t \left[M_{t+1} \frac{D_{t+1}}{D_t} (1 + Y_{t+1}) \right]. \tag{3}$$

This is a version of (1) with $A_{t+1} = M_{t+1} \frac{D_{t+1}}{D_t}$ and $\varphi_t = \mathbb{E}_t A_{t+1}$. Many specifications for the pricing kernel have been proposed in the literature and we discuss several below.

Example 2.3. Consider the simple new-Keynesian model

$$\begin{aligned} \pi_t &= \beta \mathbb{E}_t \pi_{t+1} + \theta y_t + u_t^S \\ y_t &= \mathbb{E}_t y_{t+1} - \gamma (i_t - \mathbb{E}_t \pi_{t+1}) + u_t^D, \end{aligned}$$

where y_t is the output gap, π_t is inflation and i_t is the interest rate (see, e.g., Farmer et al. (2009)). The shocks $\{u_t^S\}$ and $\{u_t^D\}$ are exogenous AR(1) processes. If $i_t = \alpha y_t + \delta \pi_t$ and we substitute this into the second and define $Y_t = (y_t, \pi_t)$ and $X_t = (u_t^S, u_t^D)$, we can express this as the vector system $Y_t = A \, \mathbb{E}_t [Y_{t+1}] + B Y_t + X_t$. If $r(B) < 1$, then $I - B$ is invertible, and we can rearrange to obtain

$$Y_t = \mathbb{E}_t \left[(I - B)^{-1} A \, Y_{t+1} \right] + (I - B)^{-1} X_t.$$

This is also a version of (1).

Remark 2.1. In example 2.3, the entire system is linear and the dynamics are already well understood. Our main interest will be in solving models with some degree of nonlinearity.

2.2. **Convergence of the Random Sum.** Returning to the model in (11), we have the option to iterate forwards, producing the fundamental "solution"

$$Y_t^* = \mathbb{E}_t \left[\sum_{j=0}^{\infty} \prod_{i=1}^{j} A_{t+i} \, \varphi_{t+j} \right] \tag{4}$$

where $\prod_{i=1}^{0} A_{t+i} = 1$. This expression often has a natural economic interpretation. For example, in the asset pricing model in example 2.1, the right hand side of (4) is the current expectation of a future flow of dividends, discounted to present value using time and risk preferences.

There are two potential problems with the expression (4). First, the sum might fail to converge, in which case $\sum_{j=0}^{\infty} \prod_{i=1}^{j} A_{t+i} \, \varphi_{t+j}$ is not a well defined random vector. Second, even if the sum does converge, its expectation might not be finite. In this section we tackle the first problem. In stating our results, we let

$$w_n := \sum_{j=0}^{n} \prod_{i=1}^{j} A_{t+i} \varphi_{t+j} \quad \text{and} \quad w_\infty := \sum_{j=0}^{\infty} \prod_{i=1}^{j} A_{t+i} \varphi_{t+j} \tag{5}$$

whenever the limit exists. In what follows, we say that (1) is *stable under forward iteration* if $w_n \to w_\infty$ and $w_\infty < \infty$ with probability one.

Proposition 2.1. *Let $\{A_t\}$ and $\{\varphi_t\}$ be stationary and ergodic. If, in addition,*

$$\mathbb{E} \ln \|A_{t+1}\| < 0 \quad \textit{and} \quad \mathbb{E} \ln \|\varphi_t\| < \infty, \tag{6}$$

then (1) is stable under forward iteration.

The norm on $\|A_{t+1}\|$ in (6) is the induced operator norm $\|A\| := \sup_{\|y\|=1} \|Ay\|$. For scalars this is just absolute value. The following proof draws heavily on ideas from a study of backward looking random coefficient models in Brandt (1986).

Proof. By Cauchy's root criterion (which is valid for Banach space valued sequences), to show that w_n converges absolutely, it suffices to show that

$$\limsup_{j \to \infty} \left\| \prod_{i=1}^{j} A_{t+i} \, \varphi_{t+j} \right\|^{1/j} < 1. \tag{7}$$

To prove (7), observe that

$$\begin{aligned} \left\| \prod_{i=1}^{j} A_{t+i} \, \varphi_{t+j} \right\|^{1/j} &\leqslant \prod_{i=1}^{j} \|A_{t+i}\|^{1/j} \; \|\varphi_{t+j}\|^{1/j} \\ &= \exp\left(\frac{1}{j} \sum_{i=1}^{j} \ln \|A_{t+i}\| + \frac{\ln \|\varphi_{t+j}\|}{j} \right). \end{aligned}$$

From the conditions in (6) and the law of large numbers for stationary and ergodic sequences, $\frac{1}{j} \sum_{i=1}^{j} \ln \|A_{t+i}\|$ converges almost surely to a negative constant and $\frac{\ln\|\varphi_{t+j}\|}{j}$

converges to zero.[5] It follows that

$$\limsup_{j\to\infty}\left(\frac{1}{j}\sum_{i=1}^{j}\ln\|A_{t+i}\|+\frac{\ln\|\varphi_{t+j}\|}{j}\right)<0 \quad \mathbb{P}\text{-a.s.}$$

Hence

$$\limsup_{j\to\infty}\exp\left(\frac{1}{j}\sum_{i=1}^{j}\ln\|A_{t+i}\|+\frac{\ln\|\varphi_{t+j}\|}{j}\right)<1 \quad \mathbb{P}\text{-a.s.}$$

The bound in (7) follows. □

Example 2.4. Returning to example 2.2, let $G_{t+1} := \ln(D_{t+1}/D_t)$ denote dividend growth, so that $A_{t+1} = M_{t+1}G_{t+1}$. The conditions of proposition 2.1 will hold if the stochastic discount factor and are stationary and ergodic, $\mathbb{E}\ln\mathbb{E}_t M_{t+1}G_{t+1}$ is finite, and, in addition,

$$\mathbb{E}\ln M_{t+1}+\mathbb{E}\,G_{t+1}<0. \tag{8}$$

A familiar example is Mehra and Prescott (1985), where $M_{t+1} = \beta u'(D_{t+1})/u'(D_t)$ with $u(c) = c^{1-\gamma}/(1-\gamma)$ and $\{G_t\}$ is stationary and ergodic (being, in their case, a function of a uniformly ergodic Markov process). Assuming $\gamma > 1$ and using stationarity of $\{G_t\}$, the restriction in (8) translates to

$$\mathbb{E}\,G_t > \frac{\ln\beta}{\gamma-1}.$$

For a given discount factor, stability under forward iteration requires the expected steady state growth rate of dividends is sufficiently large, or that the agent is sufficiently risk averse.

Example 2.5. Consider an overlapping generations (OLG) model with money, as studied by Lucas (1972), Sargent and Wallace (1983) and other authors. The economy is infinitely lived with two generations. Agents live for two periods, working in the first and consuming in the second. Lifetime utility is $U(\ell_t, c_t) = -u_1(\ell_t) + \beta\mathbb{E}_t[u_2(c_{t+1})]$, where ℓ_t is labor, c_{t+1} is consumption and u_1 and u_2 satisfy standard shape restrictions. Output per agent is $\ell_t Z_t$, where Z_t is a productivity shock. Equilibrium in the goods and money markets requires that $P_t\ell_t Z_t = M$, where M is a fixed supply of money. Individual optimality requires that agents solve $\max_{\ell_t, c_{t+1}}\{-u_1(\ell_t) + \beta\mathbb{E}_t u_2(c_{t+1})\}$

[5] We use the fact that $z_n/n \to 0$ a.s. as $n\to\infty$ whenever $\{z_n\}$ is stationary and ergodic with finite mean. This holds because if $s_n := \sum_{j=1}^{n} z_j$, then $z_n/n = s_n/n - [(n-1)/n](s_{n-1}/(n-1))$. The two sample means converge to the same number a.s. by stationarity and ergodicity.

subject to $P_{t+1}c_{t+1} \leqslant P_t \ell_t Z_t$ and $0 \leqslant \ell_t \leqslant L$. Taking the first order condition and using the equilibrium constraint to eliminate P_t and P_{t+1} yields

$$\ell_t u_1'(\ell_t) = \beta \mathbb{E}_t \left\{ u_2'(\ell_{t+1} Z_{t+1}) \ell_{t+1} Z_{t+1} \right\}. \tag{9}$$

Under standard assumptions, the function $\xi(\ell) := \ell u_1'(\ell)$ is one to one, so, with $\psi(x) := u_2'(x)x$ and $Y_t := \xi(\ell_t)$, we can write (9) as

$$Y_t = \beta \mathbb{E}_t \left\{ \psi[\xi^{-1}(Y_{t+1}) Z_{t+1}] \right\}. \tag{10}$$

This model is quite nonlinear but simplifies in certain cases. If preferences have the power form $u_1(x) = u_2(x) = x^{1-\gamma}/(1-\gamma)$, then (10) becomes

$$Y_t = \mathbb{E}_t \beta Z_{t+1}^{1-\gamma} Y_{t+1}$$

The conditions of proposition 2.1 require that $\{Z_t\}$ is ergodic and that $\mathbb{E} \ln\{\beta Z_{t+1}^{1-\gamma}\} < 0$, or $\mathbb{E} \ln Z_t > \ln \beta/(\gamma - 1)$ when $\gamma > 1$.

2.3. **Solutions with Finite Expectations.** The conditions of proposition 2.1 ensure that the random variable on the right hand side of (4), which, in asset pricing models, corresponds to the current valuation of a flow of payoffs across different states of the world, is well defined and finite. They do not, however, imply that the expectation in (4) is finite. In this section we consider finiteness of the expectation.[6]

To address this problem, we will find it convenient to add a small amount of additional structure. In particular, we will assume that the model in (1) can be expressed as

$$Y_t = \mathbb{E}_t \left[A(X_t, X_{t+1}) Y_{t+1} \right] + \varphi(X_t), \qquad t \in \mathbb{Z}. \tag{11}$$

In other words, A_{t+1} can be written as $A(X_t, X_{t+1})$ and φ_t can be written as $\varphi(X_t)$ for some Borel measurable maps A and φ and some *state process* $\mathcal{X} := \{X_t\}_{t \in \mathbb{Z}}$. We assume that $\mathcal{X}$ is a stationary, exogenous X-valued Markov process with stochastic kernel Q, so that $Q(x, B)$ represents the probability of transitioning from x into set B in one step. $\mathcal{X}$ is defined on some underlying probability space $(\Omega, \mathscr{F}, \mathbb{P})$.[7] It is *not* required to be ergodic at this stage.

[6] Note, however, that our conditions guaranteeing finite expectations, while general, do not guarantee existence of the random variable in (4). For a full treatment, both approaches are required.

[7] A stochastic kernel Q on $(\mathsf{X}, \mathscr{B})$ is a function from $(\mathsf{X}, \mathscr{B})$ to $[0, 1]$ such that $B \mapsto Q(x, B)$ is a probability measure on $(\mathsf{X}, \mathscr{B})$ for each $x \in \mathsf{X}$, and $x \mapsto Q(x, B)$ is $\mathscr{B}$-measurable for each $B \in \mathscr{B}$. An X-valued Markov process $\{X_t\}$ on probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is said to be Markov with stochastic kernel Q if $\mathbb{P}\{X_{t+1} \in B \mid X_t = x\} = Q(x, B)$ for all $x \in \mathsf{X}$ and $B \in \mathscr{B}$.

The set X is called the *state space*, and can be any Polish space.[8] We let π denote the common marginal distribution of X_t, so $\pi(B) = \mathbb{P}\{X_t \in B\}$ for all $B \in \mathscr{B}$ and $t \in \mathbb{Z}$. Since $\mathcal{X}$ is assumed to be stationary, π is a stationary distribution of Q.[9] As before, Y will be a separable Banach space with norm $\|\cdot\|$. In most applications we will have $\mathsf{Y} = \mathbb{R}^m$ for some $m \geqslant 1$ and $\|\cdot\|$ is equal to the ordinary Euclidean norm.

When considering equilibria, one concern is the need to accommodate unbounded solutions. To this end, we replace boundedness with the weaker condition that at least one moment is finite. This shifts our search for equilibria out of the space of bounded functions and into spaces of integrable functions. In particular, given $p \geqslant 1$, we consider as our candidate set $L_p(\mathsf{X}, \mathsf{Y}, \pi)$, the Bochner–Lebesgue space of Borel measurable functions $f \colon \mathsf{X} \to \mathsf{Y}$ such that

$$\|f\|_\pi := \left\{ \int \|f(x)\|^p \pi(\mathrm{d}x) \right\}^{1/p} < \infty.$$

When considering this space at fixed p, we are seeking candidate functions f such that the induced solution $Y_t = f(X_t)$ has finite p-th moment. Functions equal π-almost everywhere are identified, so that $\|\cdot\|_\pi$ defines a norm on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ and together they form a Banach space.

We make use of the following concepts: The operator norm of any bounded linear operator $\mathbf{L}$ from $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ to itself is the induced norm $\|\mathbf{L}\|_\pi := \sup_{\|f\|_\pi = 1} \|\mathbf{L}f\|_\pi$. As usual, $\mathbf{L}^i$ represents the i-th composition of $\mathbf{L}$ with itself and $\mathbf{L}^0$ is the identity map $\mathbf{I}$. A scalar λ is called an eigenvalue of $\mathbf{L}$ if there exists a nonzero $f \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$ such that $\mathbf{L}f = \lambda f$. The spectrum $\sigma(\mathbf{L})$ of $\mathbf{L}$ is all scalars $\lambda \in \mathbb{C}$ such that $\mathbf{L} - \lambda \mathbf{I}$ is not bijective.[10] The spectral radius $r(\mathbf{L})$ is $r(\mathbf{L}) := \max\{|\lambda| \,:\, \lambda \in \sigma(\mathbf{L})\}$.

We are interested in two operators in particular. Given Q, the stochastic kernel of $\mathcal{X}$, and any $\mathscr{B}$-measurable function $f \colon \mathsf{X} \to \mathsf{Y}$, we let $\mathbf{Q}f$ be the function from X to Y defined (when the integral exists) by

$$\mathbf{Q}f(x) = \int f(y) Q(x, \mathrm{d}y) \qquad (x \in \mathsf{X}).$$

[8] That is, X is separable and completely metrizable. These weak topological restrictions are only used to ensure measurability of random elements.

[9] That is, $\int Q(x, B)\pi(\mathrm{d}x) = \pi(B)$ for all $B \in \mathscr{B}$.

[10] The set $\sigma(\mathbf{L})$ is nonempty and compact in the complex plane, and every eigenvalue of $\mathbf{L}$ lies in $\sigma(\mathbf{L})$. Conversely, if $\mathbf{L}$ compact (i.e., the image of the unit ball under $\mathbf{L}$ is lies in a compact set), then the set of eigenvalues is at most countable, and an element of $\sigma(\mathbf{L})$ is either an eigenvalue of $\mathbf{L}$ or the zero element of $L_p(\mathsf{X}, \mathsf{Y}, \pi)$. See, e.g., Kantorovich and Akilov (1982).

$\mathbf{Q}$ is called the *Markov operator* associated with Q and its properties are discussed in Stokey and Lucas (1989) and Meyn and Tweedie (2009) for the case $\mathsf{Y} = \mathbb{R}$. Properties for the case where Y is a Banach space are similar. In particular, $\mathbf{Q}$ is a bounded linear operator on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ with $r(\mathbf{Q}) = \|\mathbf{Q}\|_\pi = 1$. Proofs are given in the appendix.

In addition, given A in (11), let $\mathbf{A}$ represent the operator

$$\mathbf{A} f(x) = \int A(x, x') f(x') Q(x, \mathrm{d}x') \qquad (x \in \mathsf{X}). \tag{12}$$

We call $\mathbf{A}$ the *valuation operator* by analogy with asset pricing models. (If f gives payoffs in different states and $A(x, x')$ is a stochastic discount factor, then $\mathbf{A} f$ provides expected discounted values of the payoff.) Conditions under which $\mathbf{A}$ is well defined will be stated below.

Regarding solutions to (11), a Y-valued stochastic process $\mathcal{Y} = \{Y_t\}_{t \in \mathbb{Z}}$ defined on on $(\Omega, \mathscr{F}, \mathbb{P})$ is called an *equilibrium* if it satisfies (11) with probability one for all $t \in \mathbb{Z}$. We call an equilibrium $\mathcal{Y}$ a *stationary equilibrium* if it is both an equilibrium and a stationary stochastic process, and a *stationary Markov equilibrium* if, in addition, there exists a Borel measurable function $f \colon \mathsf{X} \to \mathsf{Y}$ such that

$$Y_t = f(X_t) \quad \text{for all} \;\; t \in \mathbb{Z}. \tag{13}$$

As before, we have the option to iterate forwards, producing the "solution"

$$Y_t^* = \sum_{j=0}^{\infty} \mathbb{E}_t \left[\prod_{i=1}^{j} A(X_{t+i-1}, X_{t+i}) \varphi(X_{t+j}) \right], \tag{14}$$

where $\prod_{i=1}^{0} A(X_{t+i-1}, X_{t+i}) = 1$. Notice that we have swapped the order of expectation and summation relative to (4), the reason being that the results discussed below map immediately to the representation in (14). The connection between the two representations is treated in section 2.3.1.

As was the case with (4), the right hand side of (14) can be undefined or infinite unless conditions are imposed to ensure that the series of random elements on the right hand side converges. We state conditions under which convergence occurs and $\{Y_t^*\}$ defined in (14) is a stationary equilibrium. In fact $\{Y_t^*\}$ is a stationary Markov equilibrium with $Y_t^* = f^*(X_t)$ and

$$f^*(x) := \sum_{j=0}^{\infty} \mathbb{E} \left[\prod_{i=1}^{j} A(X_{t+i-1}, X_{t+i}) \varphi(X_{t+j}) \mid X_t = x \right]. \tag{15}$$

The next result is a straightforward application of the Neumann series lemma. In the statement, $r(\mathbf{A})$ is the spectral radius of the valuation operator.

Proposition 2.2. *If $\varphi \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$ and the valuation operator is a bounded linear operator on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ satisfying $r(\mathbf{A}) < 1$, then*

(a) f^* in (15) is an element of $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ and
(b) setting $Y_t^* = f^*(X_t)$ for all t defines a stationary Markov equilibrium.

Proof. Let $\mathbf{I}$ denote the identity mapping on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$. Regarding the first claim, since $r(\mathbf{A}) < 1$, Gelfand's formula implies the existence of an $i \in \mathbb{N}$ such that $\|\mathbf{A}^i\|_\pi < 1$. As $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ is a Banach space, this and the Neumann series theorem imply that $\mathbf{I} - \mathbf{A}$ is a bijection on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ and hence the inverse exists and equals $\sum_{i=0}^{\infty} \mathbf{A}^i$. (see, e.g., theorem 2.3.1 and corollary 2.3.3 of Atkinson and Han (2009)). Since $\varphi \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$, it follows that $f^* = \sum_{j=0}^{\infty} \mathbf{A}^j \varphi$ is a well-defined element of $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ satisfying

$$f^* - \mathbf{A} f^* = \varphi. \tag{16}$$

Let $\{Y_t^*\}$ be defined by $Y_t^* = f^*(X_t)$ for all t Fix $t \in \mathbb{Z}$ and observe that

$$\mathbb{E}_t\, A(X_t, X_{t+1}) Y_{t+1}^* + \varphi(X_t) = \int A(X_t, x') f^*(x') Q(X_t, \mathrm{d}x') + \varphi(X_t).$$

In view of (16) and the definition of $\mathbf{A}$, we have

$$\mathbb{E}_t\, A(X_t, X_{t+1}) Y_{t+1}^* + \varphi(X_t) = f^*(X_t) = Y_t^*.$$

Hence $\{Y_t^*\}$ is a stationary Markov equilibrium.

A straightforward calculation based on the definition (12) shows that

$$\mathbf{A}^j \varphi(x) = \mathbb{E}\left[\prod_{i=1}^{j} A(X_{t+i-1}, X_{t+i}) \varphi(X_{t+j}) \mid X_t = x\right].$$

Hence f^* satisfies (15). This completes the proof. $\square$

Notice also that, since f^* satisfies (15), it follows that $f^*(X_t)$ is equal to the right hand side of (14). Hence, with $Y_t^* = f^*(X_t)$, the equality in (14) is valid.

Remark 2.2. In the statement of proposition 2.2, $r(\mathbf{A})$ is the spectral radius of the operator from $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ to itself defined in (12) . If $A(x, x') \equiv A$ for some fixed linear operator $A \colon \mathsf{Y} \to \mathsf{Y}$, then $\mathbf{A} = \mathbf{Q} A$. If A is scalar-valued, then, since $r(\mathbf{Q}) = \|\mathbf{Q}\|_\pi = 1$ (see above), we have $r(\mathbf{A}) = |A| r(\mathbf{Q}) = |A|$.

Note the role of π here. One reason we weight the L_p space with π is that $\|\mathbf{Q}\|_\pi = 1$, which is a key part of the preceding argument.

The next result gives uniqueness of the fundamental solution across a substantial class of stationary equilibria.

Proposition 2.3. *If the conditions of proposition 2.2 hold and $\{Y_t\}$ is a stationary Markov equilibrium with finite p-th moment, then*

$$\mathbb{P}\left\{Y_t = Y_t^* \text{ for all } t \in \mathbb{Z}\right\} = 1. \tag{17}$$

Proof. Let $\{Y_t\}$ be a stationary Markov equilibrium with finite p-th moment. By definition, $Y_t = f(X_t)$ for some $f \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$. Since $\{Y_t\}$ are $\{Y_t^*\}$ both equilibria, we can iterate forward on (11) to obtain

$$Y_t - Y_t^* = \mathbb{E}_t \left\{ \prod_{i=1}^{j} A(X_{t+i-1}, X_{t+i})[f(X_{t+j}) - f^*(X_{t+j})] \right\}.$$

We can write this as $Y_t - Y_t^* = \mathbf{A}^j g(X_t)$ where $g := f^* - f$, which in turn yields

$$\mathbb{E}\|Y_t - Y_t^*\|^p = \mathbb{E}\|\mathbf{A}^j g(X_t)\|^p \leqslant \|\mathbf{A}^j\|^p \, \mathbb{E}\|g(X_t)\|^p.$$

The right hand side is equal to $\|\mathbf{A}^j\|^p \, \|f - f^*\|_\pi^p$. Since $r(\mathbf{A}) < 1$ and $f - f^* \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$, and since this bound is true for all j, we conclude that $\mathbb{E}\|Y_t - Y_t^*\| = 0$. The claim now follows from the argument given under equation (60). $\square$

2.3.1. *Connection to the Random Sums.* The conditions of proposition 2.2 imply that the representation of Y_t^* as a sum of expectations in (14) is valid. The conditions of proposition 2.1 guarantee that the random variable inside the expectation in (4) is well defined. A remaining technical issue is whether the expectation can be passed through the sum, ensuring that the two expressions are the same. The next result gives one set of conditions under which the answers to these questions are affirmative.

Proposition 2.4. *Let the conditions of propositions 2.1 and 2.2 hold. If A and φ are scalar-valued and nonnegative, then*

$$f^*(x) = \mathbb{E}_x \left\{ \sum_{j=0}^{\infty} \left[\prod_{i=1}^{j} A(X_{t+i-1}, X_{t+i}) \varphi(X_{t+j}) \right] \right\}. \tag{18}$$

Here $\mathbb{E}_x = \mathbb{E}[\cdot \mid X_t = x]$. The proposition assures us that (4) is valid, since $Y_t^* = f^*(X_t)$, which, from (18), is equal to the right hand side of (4).

Proof. By definition, we have $f^*(x) = \lim_{n\to\infty} \mathbb{E}_x w_n$. Since the conditions of proposition 2.1 are assumed, we know that w_∞ is well-defined, and that $\lim_{n\to\infty} w_n = w_\infty$ almost surely. Since w_n is the sum of n nonnegative random variables it is monotone increasing, and hence $f^*(x) = \lim_{n\to\infty} \mathbb{E}_x w_n = \mathbb{E}_x \lim_{n\to\infty} w_n = \mathbb{E}_x w_\infty$. This is precisely the claim in proposition 2.4. □

2.3.2. *Application: The Classic Asset Pricing Model.* We can use proposition 2.2 to recover and extend a well known existence and uniqueness result due to Robert Lucas Jr. (which is based around contraction arguments in a space of bounded functions). To this end, consider again the asset pricing equation (2). Letting $C_t = d(X_t)$ and applying the change of variable $Y_t = P_t\, u'(d(X_t))$, this equation can be expressed as $Y_t = \mathbb{E}_t\, A(X_t, X_{t+1}) Y_{t+1} + \varphi(X_t)$ when

$$A(x, x') := \beta \quad \text{and} \quad \varphi(x) := \beta \int u'(d(x')) d(x') Q(x, \mathrm{d}x'). \tag{19}$$

Lucas (1978) assumes that u is concave and bounded, which in turn gives $0 \leqslant u'(c(x'))c(x') \leqslant M$ for some $M \in \mathbb{N}$. Hence φ is likewise bounded. Since bounded functions are in $L_p(\mathsf{X}, \mathbb{R}, \pi)$ for every $p \geqslant 1$, the first condition of proposition 2.2 is satisfied, and it remains only to show that the valuation operator is a bounded linear operator on $L_p(\mathsf{X}, \mathbb{R}, \pi)$ satisfying $r(\mathbf{A}) < 1$. Here $A = \beta$, so this is immediate from remark 2.2. Hence proposition 2.2 implies the existence of a unique stationary Markov equilibrium, with finite p-th moment for all p. This recovers the existence and uniqueness results on equilibrium prices in Lucas (1978), proposition 3.

Notice how these results rely on boundedness of the utility function to prove existence and uniqueness of the price system. In applications utility is rarely bounded. Brogueira and Schütze (2015) use a weighted supremum norm approach to extend Lucas's results to the case of CRRA utility $u(c) = c^{1-\gamma}/(1-\gamma)$ and lognormal dividends $D_t = d(X_t) := \exp(X_t)$. The state process is as in (24). In this case φ in (19) becomes

$$\varphi(x) = \beta \exp\left\{(1-\gamma)\left(\rho x + b + \frac{(1-\gamma)\sigma^2}{2}\right)\right\}. \tag{20}$$

Since π is normal and φ is exponential, we have $\varphi \in L_p(\mathsf{X}, \mathbb{R}, \pi)$ for every $p \geqslant 1$, so the first condition of proposition 2.2 is again satisfied. Otherwise the arguments are unchanged, so a uniquely defined fundamental solution $Y_t^* = f^*(X_t)$ exists. Every moment of Y_t^* is finite. The conditions of proposition 2.1 and proposition 2.4 are also easily verified.

The conditions used here are weaker than in Brogueira and Schütze (2015). For example, Brogueira and Schütze (2015) stated a pair of joint restrictions on parameters that suffice for existence and uniqueness. As shown above, these restrictions are unnecessary. Brogueira and Schütze (2015) also assume a positively correlated state process. This assumption is likewise unnecessary.

3. Verifying Conditions on the Valuation Operator

In general, the conditions of proposition 2.1 are straightforward to check in applications, but the conditions of proposition 2.2 are less so. Even showing that the valuation operator $\mathbf{A}$ is a bounded linear operator on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ is nontrivial. In this section, we study how the conditions of proposition 2.2 can be verified. We show that they simplify in certain useful cases. Throughout this section, we will assume that the state space X is a Borel subset of $\mathbb{R}^m$ and Q has a *density representation* q.[11] The corresponding stationary distribution is represented by density π on X.

3.1. An L_2 Bound on the Valuation Operator.

To estimate the norm of the valuation operator, we define the functional Γ by

$$\Gamma(A) := \int \int \|A(x, x')\|^2 q(x, x')^2 \frac{\pi(x)}{\pi(x')} \, \mathrm{d}x \, \mathrm{d}x'. \tag{21}$$

This value lies in $\mathbb{R}_+ \cup \{+\infty\}$. The next lemma is proved in the appendix.

Lemma 3.1. *On $L_2(\mathsf{X}, \mathsf{Y}, \pi)$, the norm of $\mathbf{A}$ satisfies $\|\mathbf{A}\|_\pi \leqslant \sqrt{\Gamma(A)}$.*

One reason that this lemma is useful is that if $\Gamma(A)$ is finite, then we know that $\mathbf{A}$ is a bounded linear operator on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ when $p = 2$. Hence we can progress to checking the spectral radius condition in proposition 2.2. Second, if $\Gamma(A) < 1$, then, since $r(\mathbf{A}) \leqslant \|\mathbf{A}\|_\pi$ always holds, we have $r(\mathbf{A}) < 1$, and all the conditions of proposition 2.2 are satisfied.

Of course $\Gamma(A)$ is not itself trivial to evaluate. As we now show, it simplifies when certain conditions are met. One of these cases is discussed next.

[11] That is, there exists a measurable function $q \colon \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ such that $Q(x, B) = \int_B q(x, x') \, \mathrm{d}x'$ for all x in X and all B in $\mathscr{B}$.

3.1.1. *Time Reversibility.* A stochastic kernel Q on X with stationary distribution π is called *reversible* with respect to π if

$$\int \int f(x)g(x')Q(x, \mathrm{d}x')\pi(\mathrm{d}x) = \int \int g(x)f(x')Q(x, \mathrm{d}x')\pi(\mathrm{d}x) \tag{22}$$

for all bounded, $\mathscr{B}$-measurable f, g mapping X to $\mathbb{R}$. Examples of reversible processes are found among such processes as Levy, Gaussian processes, Cox-Ingersoll-Ross, Dirichlet and Gamma processes.[12] A sufficient condition when $Q(x, \mathrm{d}x')$ has density representation $q(x, x')\,\mathrm{d}x'$ is the so-called detailed balance condition

$$q(x, x')\pi(x) = q(x', x)\pi(x') \quad \text{for all} \quad (x, x') \in \mathsf{X} \times \mathsf{X}. \tag{23}$$

(Here π is the stationary distribution represented as a density.) One example is when $\mathsf{X} = \mathbb{R}$ and the state process is a Gaussian processes such as the linear AR(1) model

$$X_{t+1} = \rho X_t + b + \sigma W_{t+1}, \quad \{W_t\} \overset{\text{IID}}{\sim} N(0, I) \quad \text{and} \quad |\rho| < 1. \tag{24}$$

The stochastic kernel is $q(x, x') = N(\rho x + b, \sigma^2)$ and the unique stationary density is

$$\pi := N(\mu, \nu), \quad \text{where} \quad \mu := \frac{b}{1 - \rho} \text{ and } \nu := \frac{\sigma^2}{1 - \rho^2}. \tag{25}$$

This process time reversible and satisfies the detailed balance condition with respect to π (see, e.g., Khare and Zhou (2009)).

3.1.2. *Bounds Under Time Reversibility.* It turns out that time reversibility is useful in bounding the norm of the valuation operator. In the next result, $q_2(x, x') := \int q(x, z)q(z, x')\,\mathrm{d}z$ represents the two-step transition density.

Lemma 3.2. *Let q and π satisfy the detailed balance condition and let $\mathsf{Y} = \mathbb{R}$. If there exists a measurable function $a \colon \mathsf{X} \to \mathbb{R}$ such that either $A(x, x') = a(x')$ for all $x \in \mathsf{X}$ or $A(x, x') = a(x)$ for all $x' \in \mathsf{X}$, then*

$$\Gamma(A) = \int a(x)^2 q_2(x, x)\,\mathrm{d}x. \tag{26}$$

[12] See, for example, Boyd et al. (2005), Khare and Zhou (2009), Gouriéroux and Jasiak (2006) and Longla and Peligrad (2012)). Markov processes generated by symmetric copula models are also time reversible. as are those generated by the most common variants of Markov chain Monte Carlo algorithms, such as Metropolis–Hastings or the Gibbs sampler. See, Häggström (2002).

Proof. If $A(x, x') = a(x')$ for all $x \in \mathsf{X}$, then the detailed balance condition and the definition of $\Gamma(A)$ in (21) gives.

$$\begin{aligned}\Gamma(A) &= \int\int \frac{a(x')^2 q(x,x') q(x,x') \pi(x)}{\pi(x')} \,\mathrm{d}x \,\mathrm{d}x' \\ &= \int\int \frac{a(x')^2 q(x,x') q(x',x) \pi(x')}{\pi(x')} \,\mathrm{d}x \,\mathrm{d}x' \\ &= \int a(x')^2 \int q(x',x) q(x,x') \,\mathrm{d}x \,\mathrm{d}x'.\end{aligned}$$

Using the definition of q_2 verifies the claim. If, on the other hand, $A(x, x') = a(x)$ for all $x' \in \mathsf{X}$, then the detailed balance condition and the definition of $\Gamma(A)$ in (21) give

$$\Gamma(A) = \int\int \frac{a(x)^2 q(x,x') q(x,x') \pi(x)}{\pi(x')} \,\mathrm{d}x \,\mathrm{d}x' = \int\int a(x)^2 q(x,x') q(x',x) \,\mathrm{d}x \,\mathrm{d}x'.$$

Changing the order of integration (the integrand is nonnegative and jointly measurable, so we can apply Fubini's theorem) and using the definition of q_2 completes the proof. □

Corollary 3.1. *If, in addition to the conditions of lemma 3.2, $\mathcal{X}$ obeys the AR(1) specification in section 3.1.1, then*

$$\Gamma(A) = \frac{1}{1-\rho^2} \int a(z)^2 \varphi(z, m, s) \,\mathrm{d}z, \tag{27}$$

where

$$\varphi(\cdot, m, s) = N(m, s^2) \quad \textit{with} \quad m := \frac{b(1+\rho)}{1-\rho^2} \quad \textit{and} \quad s := \sigma \frac{\sqrt{1+\rho^2}}{1-\rho^2}.$$

Proof. The two step transition is $X_{t+2} = \rho^2 X_t + \rho b + b + \rho\sigma W_{t+1} + \sigma W_{t+2}$. It follows that $q_2(x, x')\,\mathrm{d}x'$ is the normal density $N(\rho^2 x + b(1+\rho), \sigma^2(1+\rho^2))$. Evaluating this density at $x' = x$ gives

$$q_2(x,x) = (2\pi\sigma^2(1+\rho^2))^{-1/2} \exp\left(-\frac{(x - \rho^2 x - b(1+\rho))^2}{2\sigma^2(1+\rho^2)}\right).$$

With some rearranging, we can write this as a scaled univariate normal density. That is, $q_2(x,x) = \varphi(x, m, s)/(1-\rho^2)$ where $\varphi(\cdot, m, s)$ is the $N(m, s^2)$ density with m and s as given in the statement of the corollary. Hence $\Gamma(A)$ in (26) reduces to (27). □

Because the normal density has thin tails, finiteness of $\Gamma(A)$ in (27) is a weak condition. The examples below illustrate this point.

3.1.3. *An Illustrative Example.* Recall the standard asset price evaluation formula in in (3). One variation is found in Abel (1990), where habit persistence leads to a stochastic discount factor of the form

$$M_{t+1} = \beta \left(\frac{D_{t+1}}{D_t}\right)^{-\gamma} \left(\frac{v_{t+1}}{v_t}\right)^{\gamma-1} \frac{H_{t+2}}{E_t H_{t+1}}.$$

Here $v_t = [c_{t-1}^r C_{t-1}^r]^\alpha$, with c_t being individual consumption and C_t being aggregate consumption, $\alpha \geqslant 0$ and $r \geqslant 0$, $H_{t+2} = 1 - \beta\alpha r (D_{t+2}/D_{t+1})^{1-\gamma}(D_{t+1}/D_t)^{-\alpha(1-\gamma)}$. The constant $\beta \in (0,1)$ is the nonstochastic time discount factor. Dividend growth is assumed to follow $\ln(D_{t+1}/D_t) = X_{t+1}$, where $X_{t+1} = \rho X_t + x_0 + \sigma W_{t+1}$ with $\{W_t\}$ IID and standard normal. In equilibrium, $c_t = C_t = D_t$. As shown in Calin et al. (2005), by using these equalities and the definitions above, and by restricting the state process to

$$X_{t+1} = \rho X_t + b + \sigma W_{t+1} \quad \text{with} \quad b := x_0 + \sigma^2(1-\gamma), \tag{28}$$

we can plug M_{t+1} back into (3) to obtain the forward looking model

$$Y_t = \frac{k_0 \exp(k_1 X_t)}{1 - k_2 \exp(k_1 X_t)} \mathbb{E}_t \left\{ [1 - k_2 \exp[k_1(X_{t+1} - \sigma^2(1-\gamma))][1 + Y_{t+1}] \right\}, \tag{29}$$

where k_0, k_1 and k_2 are constants that depend on parameters.[13]

Following Calin et al. (2005), we focus on the "external" habit formation case, where $r = 0$ and hence $k_2 = 0$. In this case, the forward looking restriction in (29) can be expressed as $Y_t = \mathbb{E}_t[A(X_t, X_{t+1})Y_{t+1}] + \varphi(X_t)$ where

$$A(X_t, X_{t+1}) = \varphi(X_t) = k_0 \exp(k_1 X_t) \tag{30}$$

As shown in lemma 3.1, the norm of the valuation operator is bounded by $\sqrt{\Gamma(A)}$. Since the state process is linear and Gaussian, we can apply corollary 3.1 with $A(x, x') = a(x) := k_0 \exp(k_1 x)$. Evaluating the integral in (27) gives

$$\begin{aligned} \Gamma(A) &= \frac{1}{1-\rho^2} \int k_0^2 \exp(2k_1 x) \varphi(x, m, s) \, \mathrm{d}x \\ &= \frac{k_0^2}{1-\rho^2} \exp\{2k_1(m + k_1 s^2)\}. \end{aligned}$$

where m and s are as given in corollary 3.1. Clearly this term is finite, so $\Gamma(A)$ is finite and hence the valuation operator is a bounded linear operator. If, in addition, the parameters are such that this term is strictly less than one, then $r(\mathbf{A}) \leqslant \|\mathbf{A}\|_\pi < 1$, and hence all the conditions of proposition 2.2 are satisfied (when $p = 2$).

[13]In particular, $k_0 := \beta \exp(b(1-\gamma) + \sigma^2(\gamma-1)^2/2)$, $k_1 := (1-\gamma)(\rho - \alpha)$, and $k_2 := \alpha r k_0$.

3.2. **Checking the Spectral Radius Condition.** When checking the conditions of proposition 2.2, a key property is the spectral radius condition $r(\mathbf{A}) < 1$ on the valuation operator. In evaluating the spectral radius, time reversibility can again be used to simplify computations. The main ideas are presented in this section. Throughout we suppose that $\mathsf{Y} = \mathbb{R}$ and $p = 2$, so that $L_p(\mathsf{X}, \mathsf{Y}, \pi) = L_2(\mathsf{X}, \mathbb{R}, \pi)$. The inner product on $L_2(\mathsf{X}, \mathbb{R}, \pi)$ is $\langle f, g \rangle_\pi := \int f(x) g(x) \pi(\mathrm{d}x)$, under which $L_2(\mathsf{X}, \mathbb{R}, \pi)$ is a separable Hilbert space.

3.2.1. *Some Diagonalization Results.* We will exploit an implication of reversibility of $q(x, \mathrm{d}x')$ with respect to π, which is that the corresponding Markov operator $\mathbf{Q}$ is self-adjoint on $L_2(\mathsf{X}, \mathbb{R}, \pi)$. That is, $\langle \mathbf{Q} f, g \rangle_\pi = \langle f, \mathbf{Q} g \rangle_\pi$ for all $f, g \in L_2(\mathsf{X}, \mathbb{R}, \pi)$. See, for example, Khare and Zhou (2009). If $\mathbf{Q}$ is also compact, then, by the spectral decomposition theorem (see, e.g., Debnath and Mikusinski (2005)), $\mathbf{Q}$ is diagonalizable, in the sense that there exists a complete orthonormal basis $\{e_j\}_{j \geqslant 0}$ of $L_2(\mathsf{X}, \mathbb{R}, \pi)$ such that

$$\mathbf{Q} g = \sum_{j=0}^{\infty} \lambda_j \langle g, e_j \rangle_\pi e_j \quad \text{for all } g \in L_2(\mathsf{X}, \mathbb{R}, \pi). \tag{31}$$

Equality is in the sense of convergence in $L_2(\mathsf{X}, \mathbb{R}, \pi)$. Each e_j is necessarily an eigenfunction of $\mathbf{Q}$ and λ_j is the corresponding eigenvalue.

In many cases we can give explicit representations for the eigenfunctions and eigenvalues. For example, in the case of the linear AR(1) state process in (24), we have the following result:

Lemma 3.3. *If π is as defined in (25) and*

$$e_j := h_j \circ \tau \quad \textit{where} \quad \tau(x) := \frac{\sqrt{1-\rho^2}}{\sigma} \left[x - \frac{b}{1-\rho} \right] \tag{32}$$

and h_j is the j-th normalized Hermite polynomial,[14] *then $\{e_i\}_{i \geqslant 0}$ forms an orthonormal basis of $L_2(\mathsf{X}, \mathbb{R}, \pi)$. Moreover, the Markov operator $\mathbf{Q}$ corresponding to (24) admits the decomposition in (31) with $\{e_i\}_{i \geqslant 0}$ as given in (32) and $\lambda_i = \rho^i$ for all $i \geqslant 0$.*

This is a simple variation on a well known result. A proof is given in the appendix.[15]

[14] We refer to normalized "probabilists" Hermite polynomials, which are orthogonal with respect to the weight function formed by the standard normal density φ. Normalized means that $\int h_j^2 \, \mathrm{d}\varphi = 1$ for all j.

[15] The Gaussian setting falls within the so-called Meixner class, which also includes the Gamma, Poisson and negative binomial distributions. All such distributions have representations as stationary

3.2.2. *Diagonalization and Valuation Operators.* The standard approach to evaluating the spectral radius of an operator such as $\mathbf{A}$ is to take the first n elements of a complete orthonormal basis $\{e_j\}$ of $L_2(\mathsf{X}, \mathsf{Y}, \pi)$, compute the $n \times n$ matrix $\mathbb{A}$, where $\mathbb{A}_{ij} := \langle \mathbf{A} e_i, e_j \rangle$, and evaluate the spectral radius of $\mathbb{A}$ using standard routines for matrices. If $\mathbf{A}$ is compact, then the spectral radius of $\mathbb{A}$ converges to $r(\mathbf{A})$ as $n \to \infty$.[16] The nontrivial aspect of this procedure is computing the n^2 inner products $\langle \mathbf{A} e_i, e_j \rangle$. However, as we now show, this procedure simplifies in many applications.

We make two assumptions: First, we suppose that the detailed balance condition (23) holds and $\mathbf{Q}$ is compact, so that $\mathbf{Q}$ is diagonalizable, and there exists a complete orthonormal basis $\{e_j\}_{j \geqslant 0}$ of $L_2(\mathsf{X}, \mathbb{R}, \pi)$ such that $\mathbf{Q} e_j = \lambda_j e_j$ for all $j \in \mathbb{N}$. Second, we assume that A in the definition of the valuation operator (see (12)) is a function of either x or x' but not both, so either

$$\mathbf{A} f(x) = \int f(x') g(x') q(x, x') \, \mathrm{d}x' \tag{33}$$

or

$$\mathbf{A} f(x) = g(x) \int f(x') q(x, x') \, \mathrm{d}x', \tag{34}$$

for some function $g \colon \mathsf{X} \to \mathbb{R}$. We call operators of the form (33) *type I* valuation operators. Operators of the form (34) we call *type II* valuation operators.

In the case of type I valuation operators, the detailed balance condition gives

$$\begin{aligned} \langle \mathbf{A} e_i, e_j \rangle_\pi &= \int \left[\int g(x') e_i(x') q(x, x') \, \mathrm{d}x' e_j(x) \right] \pi(x) \, \mathrm{d}x \\ &= \int g(x') e_i(x') \left[\int q(x, x') \pi(x) e_j(x) \, \mathrm{d}x \right] \mathrm{d}x' \\ &= \int g(x') e_i(x') \left[\int q(x', x) \pi(x') e_j(x) \, \mathrm{d}x \right] \mathrm{d}x'. \end{aligned}$$

---

distributions of kernel densities that can be written as $q(x, y) = \pi(y) \sum_{j=0}^{\infty} \lambda_j e_j(x) e_j(y)$ where $\{\lambda_j\}_{j \geqslant 0}$ is a real, square summable sequence satisfying $\lambda_0 = 1$ and $\{e_j\}_{j \geqslant 0}$ is a complete orthonormal basis for $L_2(\mathsf{X}, \mathbb{R}, \pi)$ consisting of normalized orthogonal polynomials, with $e_0 \equiv 1$. If $\mathbf{Q}$ is the Markov operator associated with q, then each e_k is an eigenfunction of $\mathbf{Q}$ and $\mathbf{Q}$ is diagonalizable with respect to $\{e_j\}$. In the Gaussian case, $\{e_j\}$ is the Chebychev-Hermite polynomials, while in the Gamma case, $\{e_j\}$ is the Laguerre polynomials. In the Poisson case, $\{e_j\}$ is the Poisson–Charlier polynomials. For more discussion see Griffiths (2009).

[16] See, for example, Atkinson (1967). One condition for compactness of $\mathbf{A}$ in the density case $Q(x, \mathrm{d}x') = q(x, x') \, \mathrm{d}x'$ is that $\int \int \kappa(x, x')^2 q(x, x')^2 \pi(x) \pi(x') \, \mathrm{d}x \, \mathrm{d}y$ is finite (Kantorovich and Akilov, 1982, p. 326).

But $\int q(x', x) e_j(x)\,\mathrm{d}x = \mathbf{Q} e_j(x') = \lambda_j e_j(x')$, so that, under this choice of basis, the double integral simplifies to

$$\langle \mathbf{A} e_i, e_j \rangle_\pi = \lambda_j \int g(x') e_i(x') \pi(x') e_j(x')\,\mathrm{d}x'.$$

With $\mathbb{A}_{ij} := \langle \mathbf{A} e_i, e_j \rangle_\pi$, the spectral radius of $\mathbf{A}$ can now be approximated by computing the spectral radius of the matrix $\mathbb{A} := (\mathbb{A}_{ij})_{1 \leqslant i,j \leqslant n}$ with n sufficiently large. Notice also that we can further reduce computation time by expressing $\mathbb{A}$ as $\mathbb{A} = MD$, where $D := \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $M_{ij} := \int g e_i e_j\,\mathrm{d}\pi$. Since M is symmetric, we can populate it by computing only $n(n+1)/2$ terms.

In the case of type II operators, we have

$$\begin{aligned}
\langle \mathbf{A} e_i, e_j \rangle_\pi &= \int \left[\int g(x) e_i(x') q(x, x')\,\mathrm{d}x' \right] e_j(x) \pi(x)\,\mathrm{d}x \\
&= \int g(x) \left[\int q(x, x') e_i(x')\,\mathrm{d}x' \right] e_j(x) \pi(x)\,\mathrm{d}x \\
&= \lambda_i \int g(x) e_i(x) e_j(x) \pi(x)\,\mathrm{d}x.
\end{aligned}$$

In the notation just defined, the corresponding matrix $\mathbb{A}$ can be written as DM.

3.3. **Application: The Habit Persistence Model.** Recall the habit persistence model discussed in section 3.1.3. As shown in section 3.1.3, when $r = 0$, the valuation operator is a bounded linear operator on $L_2(\mathsf{X}, \mathbb{R}, \pi)$, where π is the stationary density corresponding to the state process (28). All the conditions of proposition 2.2 will be satisfied if, in addition, $r(\mathbf{A}) < 1$ on $L_2(\mathsf{X}, \mathbb{R}, \pi)$.

We discussed analytical conditions in section 3.1.3 that imply $\Gamma(A) < 1$ and hence $r(\mathbf{A}) \leqslant \|\mathbf{A}\|_\pi \leqslant \sqrt{\Gamma(A)} < 1$. We can think of these conditions as one step contraction conditions, since $\mathbf{A}$ itself is a contraction map (because $\|\mathbf{A}\|_\pi < 1$). Calin et al. (2005) also provide a sufficient condition for existence and uniqueness of an equilibrium price-dividend ratio, and this condition can also be thought of as a one step contraction condition. (The condition differs somewhat because that study is based on contraction in a different metric.)

The aim of this section is to demonstrate that spectral radius methods allow us to establish existence and uniqueness for a much broader set of parameters. In particular, we show that conditions based on a one step contraction are often relatively strict. The idea is that, with the weaker condition $r(\mathbf{A}) < 1$, the operator $\mathbf{A}$ can be initially

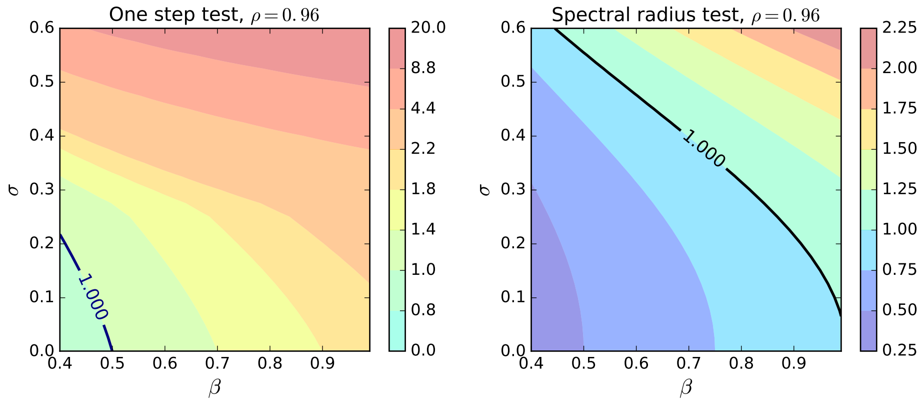

FIGURE 1. One step and spectral radius conditions, $\rho = 0.96$

expansive in some directions, provided that the contractive components eventually dominate.

To show this, we compare the results of the one step condition of Calin et al. (2005) with the condition $r(\mathbf{A}) < 1$ at a range of parameter values. Figure 1 shows one set of results. In each sub-figure, the horizontal and vertical axes show grid points for β and σ respectively. Pairs (β, σ) with test values strictly less than one (points to the south west of the 1.0 contour line) are where the respective condition holds. The left sub-figure is for the condition given in eq. (7) of Calin et al. (2005), while the right sub-figure gives the spectral radius $r(\mathbf{A})$. The spectral radius condition $r(\mathbf{A}) < 1$ is satisfied for a significantly broader range of parameters. Put differently, the sufficient condition in eq. (7) of Calin et al. (2005) fails for many parameterizations that do in fact have unique stationary Markov equilibria.[17]

The exercise is repeated figure 2 when $\rho = 0.98$ and the results are similar. In all computations we exploit the simplifications described in section 3.2.

4. Computation

In this section we briefly discuss computation, focusing on the case where $\mathsf{Y} = \mathbb{R}^n$ for some $n \in \mathbb{N}$ and the p in L_p is set to 2. The main message of this section is that projection methods can be greatly simplified in some cases with careful selection of the set of basis functions.

[17]The other parameter values are $\gamma = 2.5$, $x_0 = 0.1$ and $\alpha = 1$.

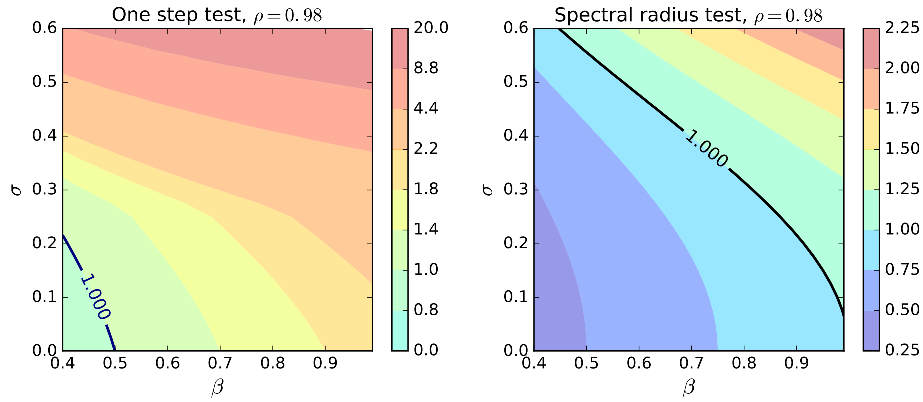

FIGURE 2. One step and spectral radius conditions, $\rho = 0.98$

4.1. **Computing Markov Equilibria in Random Coefficient Models.** For random coefficient models (see (11)), we seek to solve the equation $f = \mathbf{A}f + \varphi$ in $L_2(\mathsf{X}, \mathsf{Y}, \pi)$, where $\mathbf{A}$ is defined in (12). Consider the case where f is scalar-valued. Since $\mathbf{A}$ is linear and $r(\mathbf{A}) < 1$, we can use ordinary projection and still have guaranteed convergence. If $\{e_j\}_{j \geqslant 0}$ is any complete orthonormal basis of $L_2(\mathsf{X}, \mathsf{Y}, \pi)$, the Galerkin equations are

$$\langle f, e_j \rangle_\pi = \sum_{i=0}^{n-1} \langle f, e_i \rangle_\pi \mathbb{A}_{ij} + \langle \varphi, e_j \rangle_\pi, \quad j = 0, \ldots, n-1,$$

where $\mathbb{A}_{ij} := \langle \mathbf{A} e_i, e_j \rangle_\pi$. Solving this finite dimensional linear system at n and produces a vector c_n, with i-th element $c_{n,i}$. Setting $f_n^* := \sum_{i=0}^{n-1} c_{n,i} e_i$, we have $f_n^* \to f^*$ in $L_2(\mathsf{X}, \mathsf{Y}, \pi)$ as $n \to \infty$. See, for example, theorem 4.8 of Cheney (2001).

The nontrivial part of this exercise is computing the n^2 double integrals in the definition of of $\mathbb{A}_{ij}$. However, this often simplifies if the state process is time reversible. For example, if $A(X_t, X_{t+1}) = g(X_t)$ for some function g and q satisfies the detailed balance condition, then $\mathbf{A}$ is a type II valuation operator in the language of section 3.2. As shown in that section, if we choose as our orthonormal basis the basis that diagonalizes the Markov operator corresponding to the state process, then we need only compute $n(n+1)/2$ *single* integrals.

The habit persistence model discussed in section 3.1.3 is of this type (see, in particular, equation (30)), so we can exploit the procedures described above. Figure 3 shows a plot of the log price-dividend ratio computed using this method. The details of the model are as discussed in section 3.1.3. The price-dividend ratio is shown as a

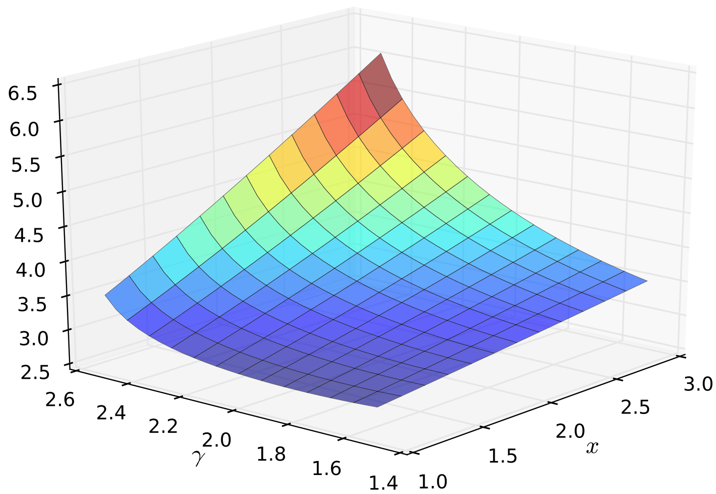

FIGURE 3. Log price-dividend ratio as a function of γ and the state

function of γ and the state x. The other parameters are $\rho = 0.96$, $\sigma = 0.1$, $\beta = 0.96$ and $\alpha = 1.0$. The number of basis elements is $n = 60$.

4.2. **Partial Linearity and Time Reversibility.** In this section we discuss the special case where $A(X_t, X_{t+1})$ is constant. (We showed in section 2.3.2 that the standard Lucas asset pricing model can be framed in this way with a suitable change of variable.) We show that the projection based computational approach simplifies further in this setting. For the sake of brevity, we restrict our discussion to the case $\mathsf{Y} = \mathbb{R}$.

Consider the partially linear forward looking equation $Y_t = \alpha \, \mathbb{E}_t [Y_{t+1}] + \varphi(X_t)$, where α, Y_t and φ are scalar valued. If $\varphi \in L_2(\mathsf{X}, \mathbb{R}, \pi)$ and $|\alpha| < 1$, then by proposition 2.2 and remark 2.2, there is a unique $f^* \in L_2(\mathsf{X}, \mathbb{R}, \pi)$ satisfying

$$f^* = \alpha \mathbf{Q} f^* + \varphi \tag{35}$$

The next result provides a simple expression for f^* when $\mathbf{Q}$ is diagonalizable.

Proposition 4.1. *If $\varphi \in L_2(\mathsf{X}, \mathbb{R}, \pi)$ and $\mathbf{Q}$ satisfies (31), then*

$$f^* = \sum_{j=0}^{\infty} (1 - \alpha\lambda_j)^{-1} \langle \varphi, e_j \rangle_\pi \, e_j. \tag{36}$$

Proof. Let f^* be as defined in (36). We claim that f^* satisfies (35) when $\mathbf{Q}$ is the diagonalizable operator in (31). To see this, recall that $\mathbf{Q}$ is continous and linear, and that e_j is an eigenvector of $\mathbf{Q}$ with eigenvalue λ_j. Hence

$$\mathbf{Q}f^* = \sum_{j=0}^{\infty} \frac{1}{1-\alpha\lambda_j} \langle \varphi, e_j \rangle_\pi \, \mathbf{Q}e_j = \sum_{j=0}^{\infty} \frac{\lambda_j}{1-\alpha\lambda_j} \langle \varphi, e_j \rangle_\pi \, e_j.$$

Since $\{e_j\}$ is a complete orthonormal basis, we then have

$$\begin{aligned} \varphi + \alpha \mathbf{Q} f^* &= \varphi + \sum_{j=0}^{\infty} \frac{\alpha\lambda_j}{1-\alpha\lambda_j} \langle \varphi, e_j \rangle_\pi \, e_j \\ &= \sum_{j=0}^{\infty} \langle \varphi, e_j \rangle_\pi e_j + \sum_{j=0}^{\infty} \frac{\alpha\lambda_j}{1-\alpha\lambda_j} \langle \varphi, e_j \rangle_\pi \, e_j = \sum_{j=0}^{\infty} \frac{1}{1-\alpha\lambda_j} \langle \varphi, e_j \rangle_\pi \, e_j = f^*. \end{aligned}$$

In particular, f^* satisfies (35). □

In the Gaussian setting of section 3.1.1, the solution (36) becomes

$$f^* = \sum_{i=0}^{\infty} (1-\alpha\rho^i)^{-1} \langle \varphi, e_i \rangle_\pi \, e_i \tag{37}$$

where $\{e_j\}$ is as defined in (32).

Example 4.1. Recall the CRRA utility asset pricing model with lognormal dividends from section 2.3.2. To compute f^* we take this expression for φ and insert it into (37). The price function p^* is obtained from f^* by reversing the change of variable $f(x) = p(x)\, u'(d(x))$. That is,

$$p^*(x) = \exp(\gamma x) \sum_{i=0}^{\infty} (1-\beta\rho^i)^{-1} \langle \varphi, e_i \rangle_\pi \, e_i(x). \tag{38}$$

5. General Nonlinear Models

In this section we set out a more general forward looking equation and study existence and uniqueness of solutions.

5.1. Problem Statement.

We analyze forward looking models where the equilibrium condition can be expressed as

$$Y_t = \mathbb{E}_t \, G(X_t, X_{t+1}, Y_{t+1}) \tag{39}$$

for all $t \in \mathbb{Z}$, the integers.[18] The sequence $\mathcal{Y} := \{Y_t\}_{t\in\mathbb{Z}}$ evolves in Banach space Y and is understood as endogenous. The function $G \colon \mathsf{X} \times \mathsf{X} \times \mathsf{Y} \to \mathsf{Y}$ is a Borel measurable

[18] In the theory that follows, nothing changes if we replace $t \in \mathbb{Z}$ with $t \in \{0, 1, \ldots\}$.

function that defines a relationship between the endogenous variables and the state process. Together G and $\mathcal{X}$ are the primitives and $\mathcal{Y}$ is the object that we wish to solve for.

Example 5.1. For agents with risk sensitive preferences (Hansen and Sargent (1995)), a typical formulation of the present value of lifetime utility associated with consumption path $\{C_t\}$ is $U_t = u(C_t) + \frac{\beta}{\sigma} \ln \mathbb{E}_t \exp(\sigma U_{t+1})$, where u is a flow utility function and β and σ are parameters. Assume as in Tallarini (2000) or Anderson (2005) that C_t is a function of some given state process. Since the expectation lies inside the log function, this equation is not in the form of (39). However, if we set $Y_t := \mathbb{E}_t \exp(\sigma U_{t+1})$, then the equation is $U_t = u(C_t) + \frac{\beta}{\sigma} \ln Y_t$, or $\exp(\sigma U_t) = \exp\{\sigma u(C_t) + \beta \ln Y_t\}$. Shifting this equation forward one period and taking expectations conditional on time t gives

$$Y_t = \mathbb{E}_t \exp\{\sigma u(C_{t+1}) + \beta \ln Y_{t+1}\}.$$

This is a version of (39). If we solve it for Y_t, we can recover U_t via $U_t = u(C_t) + \frac{\beta}{\sigma} \ln Y_t$.

5.2. **Existence and Uniqueness.** To begin, observe that if f induces a stationary Markov equilibrium—that is, if $\{Y_t\}$ defined by (13) is an equilibrium—then $f(X_t) = \mathbb{E}_t \, G(X_t, X_{t+1}, f(X_{t+1}))$ must hold with probability one. State-by-state, this restriction can be expressed as

$$f(x) = \int G(x, x', f(x')) Q(x, \mathrm{d}x') \qquad (x \in \mathsf{X}). \tag{40}$$

Any solution to the functional equation (40) induces a stationary Markov equilibrium via $Y_t = f(X_t)$. This prompts us to study fixed points of the operator $\mathbf{T}$ defined by

$$\mathbf{T} f(x) = \int G(x, x', f(x')) Q(x, \mathrm{d}x') \qquad (x \in \mathsf{X}). \tag{41}$$

Rather than embedding $\mathbf{T}$ in a space of bounded functions—which would preclude unbounded solutions—we treat it as a mapping from $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ into itself. As before, $p \geqslant 1$ and π is the distribution of each X_t. To pursue this approach, it is necessary that $\mathbf{T}$ is invariant on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$. We will say that $(G, \mathcal{X})$ is *order p regular* if $\mathbf{T}$ has this property; that is, if $\mathbf{T} f$ is in $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ whenever f is in $L_p(\mathsf{X}, \mathsf{Y}, \pi)$. We provide sufficient conditions for order p regularity in section 5.3.

The second strand of conditions we require relate to stability under iteration. Below, we say that $(G, \mathcal{X})$ is *order p contracting* if there exists a measurable function $\kappa \colon \mathsf{X} \times \mathsf{X} \to \mathbb{R}_+$ such that, for all (x, x') in $\mathsf{X} \times \mathsf{X}$ and (y_1, y_2) in $\mathsf{Y} \times \mathsf{Y}$,

$$\|G(x, x', y_1) - G(x, x', y_2)\| \leqslant \kappa(x, x') \, \|y_1 - y_2\|, \tag{42}$$

and, in addition, $\mathbf{K}$ defined by

$$\mathbf{K}f(x) = \int f(x')\kappa(x,x')Q(x,\mathrm{d}x') \qquad (x \in \mathsf{X}) \tag{43}$$

is a bounded linear operator on $L_p(\mathsf{X}, \mathbb{R}, \pi)$ with $r(\mathbf{K}) < 1$. As a further property, we will say that $(G, \mathcal{X})$ is *strongly order p contracting* if it is order p contracting and the operator $\mathbf{K}$ in (43) satisfies $\|\mathbf{K}\|_\pi < 1$.

Remark 5.1. The spectral radius condition $r(\mathbf{K}) < 1$ in the order p contraction property will typically be checked numerically at each parameterization. Examples are given below.

Remark 5.2. In some nonlinear models, a function κ that fits the bound in (42) can be found by studying the derivative of the mapping $y \mapsto G(x, x', y)$. For example, suppose that $\mathsf{Y} = \mathbb{R}^n$ and that the Jacobian $\nabla G(x, x', y)$ of this mapping exists and is continuous. Then, by the mean value theorem, for each $y_1, y_2 \in \mathsf{Y}$, there exists a $\theta \in (0, 1)$ with

$$\|G(x, x', y_1) - G(x, x', y_2)\| \leqslant \|\nabla G(x, x', \theta y_1 + (1-\theta)y_2)\| \cdot \|y_1 - y_2\|.$$

In particular, if there exists a function κ with $\|\nabla G(x, x', y)\| \leqslant \kappa(x, x')$ for all $(x, x') \in \mathsf{X} \times \mathsf{X}$ and $y \in \mathsf{Y}$, then the bound (42) holds for this function κ.

Together, regularity and either of the contraction properties implies existence and uniqueness of a fixed point of $\mathbf{T}$. This follows from the next proposition.

Proposition 5.1. *Let $(G, \mathcal{X})$ be order p regular. The following statements are true:*

(a) If $(G, \mathcal{X})$ is order p contracting, then there exists an $m \in \mathbb{N}$ such that $\mathbf{T}^m$ is a Banach contraction on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$.
(b) If G is strongly order p contracting, then $\mathbf{T}$ is a Banach contraction on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$.

Proof of proposition 5.1. Regarding the first claim, fix $p \geqslant 1$ and let the conditions of the theorem be satisfied. By proposition 5.4, we know that $\mathbf{T}$ maps $L_p(\mathsf{X}, \mathbb{R}, \pi)$ to itself. It remains only to show that $\mathbf{T}^m$ is a contraction mapping for some m. To this end, pick $f, g \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$ and fix $j \in \mathbb{N}$ and $x \in \mathsf{X}$. By the definition of $\mathbf{T}$ and the order p contraction property, we have

$$\begin{aligned}\|\mathbf{T}^j f(x) - \mathbf{T}^j g(x)\| &= \left\| \int [G(x, x', \mathbf{T}^{j-1} f(x')) - G(x, x', \mathbf{T}^{j-1} g(x'))] \, Q(x, \mathrm{d}x') \right\| \\ &\leqslant \int \|G(x, x', \mathbf{T}^{j-1} f(x')) - G(x, x', \mathbf{T}^{j-1} g(x'))\| \, Q(x, \mathrm{d}x') \\ &\leqslant \int \kappa(x, x') \, \|\mathbf{T}^{j-1} f(x') - \mathbf{T}^{j-1} g(x')\| \, Q(x, \mathrm{d}x').\end{aligned}$$

If we let $h_j(x) := \|\mathbf{T}^j f(x) - \mathbf{T}^j g(x)\|$, then the preceding inequality can be expressed as $h_j \leqslant \mathbf{K} h_{j-1}$. Since $\kappa \geqslant 0$, the operator $\mathbf{K}$ is order preserving, in the sense that $\mathbf{K} f \leqslant \mathbf{K} g$ whenever $f \leqslant g$. This gives $h_j \leqslant \mathbf{K}^2 h_{j-2}$, and, iterating further, $h_j \leqslant \mathbf{K}^j h_0$. Since the norm on $L_p(\mathsf{X}, \mathbb{R}, \pi)$ preserves order for nonnegative functions, we then have $\|h_j\|_\pi \leqslant \|\mathbf{K}^j h_0\|_\pi$. By the submultiplicative property of the operator norm, this implies $\|h_j\|_\pi \leqslant \|\mathbf{K}^j\|_\pi \|h_0\|_\pi$. In other words,

$$\|\mathbf{T}^j f - \mathbf{T}^j g\|_\pi \leqslant \|\mathbf{K}^j\|_\pi \|f - g\|_\pi.$$

From Gelfand's formula, $\|\mathbf{K}^j\|_\pi^{1/j} \to r(\mathbf{K})$, where $r(\mathbf{K})$ is the spectral radius. By assumption, $r(\mathbf{K}) < 1$. hence there is an $m \in \mathbb{N}$ such that $\|\mathbf{K}^m\|_\pi < 1$. For this m, we see that $\mathbf{T}^m$ is a Banach contraction on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$.

Regarding the second claim in the theorem, suppose that the strong contraction property holds. Pick $f, g \in L_p(S, \mathsf{Y}, \pi)$ and fix $x \in S$. Using similar arguments to the proof of the first claim, we have

$$\begin{aligned}\|\mathbf{T} f(x) - \mathbf{T} g(x)\|^p &= \left\| \int [G(x, x', f(x')) - G(x, x', g(x'))] Q(x, \mathrm{d}x') \right\|^p \\ &\leqslant \int \|G(x, x', f(x')) - G(x, x', g(x'))\|^p \, Q(x, \mathrm{d}x') \\ &\leqslant \theta^p \int \|f(x') - g(x')\|^p \, Q(x, \mathrm{d}x').\end{aligned}$$

Integrating with respect to π and applying (56) gives

$$\|\mathbf{T} f - \mathbf{T} g\|_\pi^p \leqslant \theta^p \|f - g\|_\pi^p$$

Raising both sides to the power of $1/p$ gives stated contraction property. □

Let the conditions of proposition 5.1 be satisfied. Since $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ is complete, it follows from the theorem that, under order p regularity and the order p contraction property, there exists a unique $f^* \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$ such that

$$f^*(x) = \int G(x, x', f^*(x'))Q(x, \mathrm{d}x') \qquad (x \in \mathsf{X}). \tag{44}$$

Given this f^*, let $Y_t^* = f^*(X_t)$ for all $t \in \mathbb{Z}$. We call $\{Y_t^*\}$ the *fundamental solution*.

Theorem 5.1. *If $(G, \mathcal{X})$ is order p regular and order p contracting, then*

(a) the fundamental solution $\{Y_t^*\}$ is a stationary Markov equilibrium,
(b) the vector Y_t^* has finite p-th moment, and
(c) if $\{Y_t\}$ is a stationary Markov equilibrium such that Y_t has finite p-th moment, then

$$\mathbb{P}\left\{Y_t = Y_t^* \textit{ for all } t \in \mathbb{Z}\right\} = 1. \tag{45}$$

The proof is relatively straightforward and delayed until the appendix.

We can strengthen the uniqueness result in the case where $(G, \mathcal{X})$ is strongly order p contracting. The next proposition shows uniqueness over the class of all stationary solutions with finite first moment.

Proposition 5.2. *Let $(G, \mathcal{X})$ be first order regular and strongly first order contracting, and let $\{Y_t^*\}$ be the fundamental solution. If $\{Y_t\}$ is a Y-valued stationary stochastic process satisfying the forward looking restriction (39) and having finite first moment, then*

$$\mathbb{P}\left\{Y_t = Y_t^* \textit{ for all } t \in \mathbb{Z}\right\} = 1.$$

Proof. fix $t \geqslant 0$ and let $\{Y_t\}$ be as in the statement of the proposition. Since both processes are equilibria and G is strongly order p contracting, we have

$$\|Y_t - Y_t^*\| \leqslant \mathbb{E}_t \left\|G(X_t, X_{t+1}, Y_{t+1}) - G(X_t, X_{t+1}, Y_{t+1}^*)\right\| \leqslant \theta \, \mathbb{E}_t \left\|Y_{t+1} - Y_{t+1}^*\right\|.$$

Iterating and then taking expectations gives

$$\mathbb{E}\,\|Y_t - Y_t^*\| \leqslant \theta^i \mathbb{E}\,\|Y_{t+i} - Y_{t+i}^*\| \leqslant \theta^i \mathbb{E}\,\|Y_{t+i}\| + \theta^i \mathbb{E}\,\|Y_{t+i}^*\|$$

for all $i \geqslant 0$. Since $\{Y_t\}$ and $\{Y_t^*\}$ are both stationary with finite first moments, the right hand side of this expression converges to zero, and hence $\mathbb{E}\,\|Y_t - Y_t^*\|_p = 0$. It follows that $\mathbb{P}\{Y_t = Y_t^*\} = 1$. Countable intersections of probability one sets have probability one. □

5.3. **Sufficient Conditions.** Order p regularity and the contraction properties stated above can be non-trivial to check in applications. In this section we give a variety of sufficient conditions for the regularity and contraction properties stated above, and show how they can be used in various applications.

5.3.1. *Sufficient Conditions for Regularity.* The first result is elementary but a useful first step.

Proposition 5.3. *If* $\mathbb{E}\,\|G(X_t, X_{t+1}, f(X_{t+1}))\|^p < \infty$ *for all* $f \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$*, then* $(G, \mathcal{X})$ *is order* p *regular.*

Proof. Fix $p \geqslant 1$ and let $f \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$. By Jensen's inequality for Bochner integrals, we have $\|\mathbf{T} f(x)\|^p \leqslant \int \|G(x, x', f(x'))\|^p Q(x, \mathrm{d}x')$, and hence

$$\|\mathbf{T} f\|_\pi^p \leqslant \int \int \|G(x, x', f(x'))\|^p Q(x, \mathrm{d}x') \pi(\mathrm{d}x).$$

The right hand side is finite by assumption. Thus $(G, \mathcal{X})$ is order p regular. □

Condition 5.1. There exists a measurable function $\ell \colon \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ and a constant $m \in \mathbb{R}_+$ such that $\mathbb{E}\,|\ell(X_t, X_{t+1})|^p < \infty$ and

$$\|G(x, x', y)\| \leqslant \ell(x, x') + m\|y\| \quad \text{for all} \quad (x, x') \in \mathsf{X} \times \mathsf{X} \text{ and } y \in \mathsf{Y}.$$

Proposition 5.4. *If condition 5.1 holds, then* $(G, \mathcal{X})$ *is order* p *regular.*

Proof. Now suppose that condition 5.1 holds. Fix $f \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$. We have[19]

$$\begin{aligned}
\mathbb{E}\,\|G(X_t, X_{t+1}, f(X_{t+1}))\|^p &\leqslant \mathbb{E}\,|\ell(X_t, X_{t+1}) + m\|f(X_{t+1})\||^p \\
&\leqslant 2^{p-1}\mathbb{E}\,|\ell(X_t, X_{t+1})|^p + 2^{p-1} m^p \mathbb{E}\|f(X_{t+1})\|^p \\
&= 2^{p-1}\mathbb{E}\,|\ell(X_t, X_{t+1})|^p + 2^{p-1} m^p \|f\|_\pi^p.
\end{aligned}$$

Since $\|f\|_\pi^p < \infty$ and the first term is finite by assumption, proposition 5.3 implies that $(G, \mathcal{X})$ is order p regular. □

Remark 5.3. Condition 5.1 simplifies if we have additional structure. For example, if $G(x, x', y) = \psi(x') + \beta y$ and ψ is in $L_p(\mathsf{X}, \mathsf{Y}, \pi)$, then condition 5.1 holds: take $\ell(x, x') := \|\psi(x')\|$ and observe that $\mathbb{E}\,|\ell(X_t, X_{t+1})|^p = \int \|\psi(x)\|^p \pi(\mathrm{d}x) < \infty$.

[19] We are applying the elementary inequality $|a + b|^p \leqslant 2^{p-1}(|a|^p + |b|^p)$.

Example 5.2. Consider again the Lucas asset pricing equation (2). Letting $C_t = d(X_t)$ and applying the change of variable $Y_t = P_t\, u'(d(X_t))$, this equation can be expressed as $Y_t = \mathbb{E}_t G(X_t, X_{t+1}, Y_{t+1})$ when

$$G(x, x', y) = \beta u'(d(x'))d(x') + \beta y. \tag{46}$$

By remark 5.3, to show order p regularity, it suffices to show that $\psi(x') := \beta u'(d(x'))d(x')$ is in $L_p(\mathsf{X}, \mathbb{R}, \pi)$. Lucas (1978) assumes that u is concave and bounded, which in turn gives $0 \leqslant u'(c(x'))c(x') \leqslant M$ for some $M \in \mathbb{N}$. Since bounded functions are in $L_p(\mathsf{X}, \mathbb{R}, \pi)$ for every $p \geqslant 1$, this model is order p regular for all p.

In the next sufficient condition, we assume that X is a Borel subset of $\mathbb{R}^n$ and that $Q(x, \mathrm{d}x')$ has density representation $q(x, x')\,\mathrm{d}x' = Q(x, \mathrm{d}x')$. That is, there exists a measurable function $q\colon \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ such that $Q(x, B) = \int_B q(x, x')\,\mathrm{d}x'$ for all x in X and all B in $\mathscr{B}$. The corresponding stationary distribution is represented by density π on X. In this setting for any measurable function $\kappa\colon \mathsf{X} \times \mathsf{X} \to \mathbb{R}$, we let Γ be the functional defined by (21).

Proposition 5.5. *If there exist measurable functions $\kappa\colon \mathsf{X}\times\mathsf{X} \to \mathbb{R}$ and $b\colon \mathsf{X}\times\mathsf{X} \to \mathbb{R}$ such that $\Gamma(\kappa)$ and $\mathbb{E}[\zeta(X_t, X_{t+1})^2]$ are finite and, for all $(x, x') \in \mathsf{X} \times \mathsf{X}$ and $y \in \mathsf{Y}$,*

$$\|G(x, x', y)\| \leqslant \kappa(x, x')\|y\| + \zeta(x, x'), \tag{47}$$

then $(G, \mathcal{X})$ is second order (i.e., order 2) regular.

The proof is deferred to the appendix. Techniques for computing $\Gamma(\kappa)$ were discussed in section 3.1.2.

5.3.2. *Sufficient Conditions for Contractivity.* Next we provide sufficient conditions for contractivity. The first condition requires that, stepping backwards in time, the system is contracting in every state of the world.

Lemma 5.1. *Let $p \geqslant 1$. If there exists a $\theta \in (0, 1)$ such that, for all $(x, x') \in \mathsf{X} \times \mathsf{X}$ and all $(y_1, y_2) \in \mathsf{Y} \times \mathsf{Y}$,*

$$\|G(x, x', y_1) - G(x, x', y_2)\| \leqslant \theta\|y_1 - y_2\|, \tag{48}$$

then $(G, \mathcal{X})$ is strongly order p contracting.

Proof. If the stated conditions holds, then (42) holds with $\kappa(x, x')$ equal to the constant θ. The corresponding operator $\mathbf{K}$ in (43) is given by $\mathbf{K} = \theta\mathbf{Q}$, where $\mathbf{Q}$ is the

Markov operator. Hence $\|\mathbf{K}\|_\pi = \|\theta \mathbf{Q}\|_\pi = \theta \|\mathbf{Q}\|_\pi$. But $\|\mathbf{Q}\|_\pi = 1$ (see section 3). The stated property follows. □

Example 5.3. We saw in (46) that the dynamics of the Lucas asset pricing model model can be represented by the function $G(x, x', y) = \beta u'(d(x'))d(x') + \beta y$. Clearly $|G(x, x', y_1) - G(x, x', y_2)| \leqslant \beta |y_1 - y_2|$ for all $(x, x') \in \mathsf{X} \times \mathsf{X}$ and $y_1, y_2 \in \mathsf{Y}$. In particular, the conditions of lemma 5.1 hold, and $(G, \mathcal{X})$ is strongly order p contracting.

Example 5.4. If, in the OLG model in example 2.5, preferences have the power form $h(\ell) = \ell^{1-\gamma}/(1-\gamma)$ and $v(c) = c^{1-\gamma}/(1-\gamma)$, then $G(x, x', y) = \beta y [\zeta(x')]^{1-\gamma}$ and hence $|G(x, x', y_1) - G(x, x', y_2)| \leqslant \beta [\zeta(x')]^{1-\gamma} |y_1 - y_2|$. If ζ is bounded with $\beta[\zeta(x')]^{1-\gamma} < 1$, then $(G, \mathcal{X})$ is strongly order p contracting for any p and any choice of $\mathcal{X}$.

The operator $\mathbf{T}$ that we study in this paper can be thought of as an abstract version of an Urysohn operator, fixed point theory for which typically involves Lipschitz conditions that are uniform across all states.[20] Lemma 5.1 is of this type, and has some useful applications. But for some models, this uniformity is too strict. This fact motivates us to develop fixed point results that rely on weaker conditions, where contractivity does not necessarily hold state-by-state. We turn to this problem next.

Suppose that X is a Borel subset of $\mathbb{R}^n$ and that Q has density representation $q(x, x') \, \mathrm{d}x' = Q(x, \mathrm{d}x')$. Assume the existence of a measurable function $\kappa \colon \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ such that the contraction condition (42) holds and let Γ be the functional in (21). Let $\mathbf{K}$ be the operator defined in (43). In this setting, the following statement is true:

Proposition 5.6. *If $\Gamma(\kappa)$ is finite and $r(\mathbf{K}) < 1$, then $(G, \mathcal{X})$ is second order contracting. If $\Gamma(\kappa) < 1$, then $(G, \mathcal{X})$ is strongly second order contracting.*

Proof. If $\Gamma(\kappa)$ is finite, then lemma 3.1 implies that $\mathbf{K}$ in (43) is a bounded linear operator on $L_2(\mathsf{X}, \mathbb{R}, \pi)$. It then follows from the definition that $r(\mathbf{K}) < 1$ implies $(G, \mathcal{X})$ is second order contracting. In addition, if $\Gamma(\kappa) < 1$, then the bound in lemma 3.1 implies that $\|\mathbf{K}\|_\pi < 1$. Hence $(G, \mathcal{X})$ is *strongly* second order contracting. □

5.4. **Application: Overlapping Generations with Money.** Consider again the OLG model discussed in example 2.5. As shown in (10), the dynamics are given by $Y_t = \mathbb{E}_t \, G(X_t, X_{t+1}, Y_{t+1})$ when $G(x, x', y) = \beta \psi[\xi^{-1}(y)\zeta(x')]$, $\xi(\ell) := \ell h'(\ell)$ and

[20]See, for example, Atkinson and Han (2009), p. 217, or Stokey and Lucas (1989), lemma 17.1.

$\psi(x) := v'(x)x$. Suppose for now that v, the utility function for consumption, is bounded and concave (other possibilities are explored below). We seek conditions under which $(G, \mathcal{X})$ is order p regular and order p contracting. We will focus on the case $p = 2$.

As a first step, observe that $(G, \mathcal{X})$ is second order regular, regardless of the specification of the Markov process $\mathcal{X}$. To see this, note that in our setting the function ψ is bounded.[21] Hence G is bounded, and $(G, \mathcal{X})$ is order p regular by proposition 5.3.

Whether the contraction property holds or not will depend on the functional forms and parameters. For now, suppose that preferences are linear in leisure, so that $h(\ell) = a\ell$ for some $a > 0$, and hence $G(x, x', y) = \beta\psi[y\,\zeta(x')/a]$. Suppose further that $\mathcal{X}$ is the AR(1) process in (24), and that there exists a constant δ such that

$$\left|\frac{\partial\, G(x, x', y)}{\partial y}\right| = \beta|\psi'[y\,\zeta(x')/a]|\zeta(x')/a \leqslant \delta\zeta(x') \tag{49}$$

for all $y \in \mathsf{Y} = \mathbb{R}$. Then, by remark 5.2, the bound (42) holds with $\kappa(x, x') = \delta\zeta(x')$. To check whether the corresponding operator $\mathbf{K}$ in (43) is a bounded linear operator on $L_2(\mathsf{X}, \mathsf{Y}, \pi)$ we can use corollary 3.1. With $\kappa(x, x') = \eta(x') = \delta\zeta(x')$, we can evaluate $\Gamma(\kappa)$ using (27), which gives $\Gamma(\kappa) = \frac{\delta^2}{1-\rho^2}\int \zeta(z)^2\varphi(z, m, s)\,\mathrm{d}z$, where $\varphi(\cdot, m, s)$ is the $N(m, s^2)$ density with m and s as defined in corollary 3.1. For this integral to be finite, it suffices that $\mathbb{E}\zeta(Z)^2$ is finite whenever Z is normally distributed (i.e., the productivity process has finite second moment).

If this is true, then proposition 5.6 implies that $\mathbf{K}$ is a bounded linear operator from $L_2(\mathsf{X}, \mathbb{R}, \pi)$ to itself with $\|\mathbf{K}\|_\pi \leqslant \sqrt{\Gamma(\kappa)}$. If $\Gamma(\kappa) < 1$, then $(G, \mathcal{X})$ is strongly second order contracting, and a unique stationary Markov equilibrium exists. If not, we can study the spectral radius $r(\mathbf{K})$. In the latter case, the relevent condition is $r(\mathbf{K}) < 1$, under which a unique stationary Markov equilibrium exists.

One utility function for consumption that satisfies the stated conditions is the CARA form $v(c) = 1 - \exp(-\gamma c)$, where $\gamma > 0$. In this case, straightforward calculations show that (49) holds with $\delta := \gamma\beta/a$. Figure 4 shows (β, ρ) pairs such that the condition $r(\mathbf{K}) < 1$ is satisfied (all pairs below the 1.0 contour line). In the left sub-figure, $\sigma = 0.05$. In the right sub-figure, $\sigma = 0.05$, and the set of stable (β, ρ) pairs has increased.[22] In general, a parameterization is more likely to satisfy the second

[21]Since v is bounded and concave, there exists an M and N such that, for all $x > 0$, we have $N = v(0) \leqslant v(x) + v'(x)(0 - x) \leqslant M - v'(x)x$, and hence $0 \leqslant v'(x)x \leqslant M - N$.

[22]The other parameters are $a = 1.0$, $b = 0.0$ and $\gamma = 0.5$.

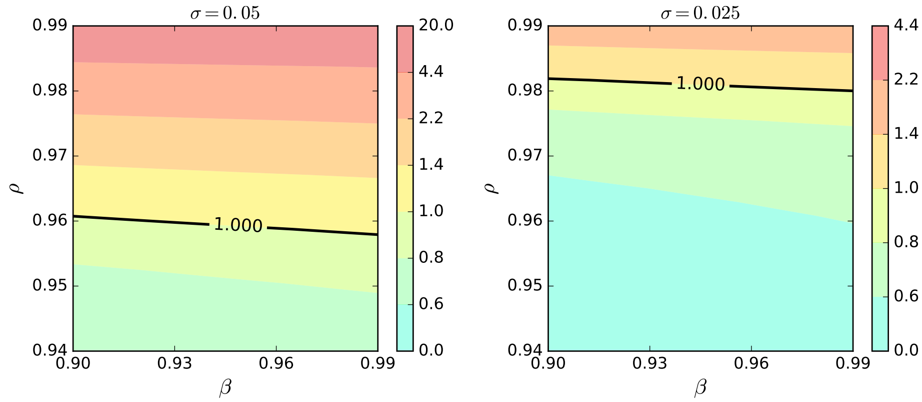

FIGURE 4. The spectral radius condition for the OLG model with money

order contraction property when the discount factor is smaller, the state process is less persistent and the volatility of the state process is lower.

5.5. **Computing Markov Equilibria in the General Setting.** Consider the general forward looking model discussed in section 5. Assume the conditions of theorem 5.1. Our next aim is to produce a consistent and globally convergent method for computing the unique fixed point f^* in proposition 5.1, where consistency means that the limit of our iterative method converges to the true fixed point f^* as the approximation architecture becomes finer and finer.

Algorithm 1 yields such a method when $\{e_j\}$ is an orthonormal basis of $L_2(\mathsf{X}, \mathsf{Y}, \pi)$. In the statement of the algorithm, Π is the distribution of (X_t, X_{t+1}) on $\mathsf{X} \times \mathsf{X}$.

Algorithm 1: Iteration composed with projection

---

fix initial $f \in L_2(\mathsf{X}, \mathsf{Y}, \pi)$;
while *a stopping condition fails* **do**
  for $i = 0, \ldots, n$ **do**
    set $c_i = \int \int \langle G(x, x', f(x')), e_i(x) \rangle \, \Pi(\mathrm{d}x, \mathrm{d}x')$;
  end
  set $f = \sum_{i=0}^{n} c_i e_i$;
end
return f;

Since $\Pi(\mathrm{d}x, \mathrm{d}x') = Q(x, \mathrm{d}x')\pi(\mathrm{d}x)$, we have

$$\begin{aligned} c_i &= \int \int \langle G(x, x', f(x')), e_i(x) \rangle \, Q(x, \mathrm{d}x')\pi(\mathrm{d}x) \\ &= \int \left\langle \int G(x, x', f(x'))Q(x, \mathrm{d}x'), e_i(x) \right\rangle \pi(\mathrm{d}x) \\ &= \langle \mathbf{T}f, e_i \rangle_\pi \, . \end{aligned}$$

Hence f in line 7 evaluates to $\sum_{i=0}^n \langle \mathbf{T}f, e_i \rangle_\pi e_i$. The implication is that stepping through the loop in algorithm 1 is equivalent to iterating with $\mathbf{P} \circ \mathbf{T}$, the composition of the orthogonal projection $\mathbf{P}$ from $L_2(\mathsf{X}, \mathsf{Y}, \pi)$ onto the span of $\{e_j\}_{j \leqslant n}$ with the operator $\mathbf{T}$.

The operator $\mathbf{P}$ has three fundamental properties that are important for us. First, it is optimal in the sense that, for each $f \in L_2(\mathsf{X}, \mathsf{Y}, \pi)$, the function $\mathbf{P}f$ equals the minimizer of $\|f - h\|_\pi$ over all $h \in \operatorname{span}\{e_0, \ldots, e_n\}$. Note that optimality with respect to $\| \cdot \|_\pi$ enforces small deviation where π puts most of its probability mass. This is desirable because it implies that the quality of approximation is maximized in regions where the state process spends most of its time—assuming some degree of ergodicity.

The second and third properties we make use of are that the orthogonal projection is linear and nonexpansive respectively, where nonexpansive means that $\|\mathbf{P}f\|_\pi \leqslant \|f\|_\pi$ for all f in $L_2(\mathsf{X}, \mathsf{Y}, \pi)$.[23] These two facts lie behind the following lemma.

Lemma 5.2. *If $\mathbf{T}$ is a contraction of modulus θ on $L_2(\mathsf{X}, \mathsf{Y}, \pi)$, then so is $\hat{\mathbf{T}} := \mathbf{P} \circ \mathbf{T}$.*

Proof. Let f and g be elements of $L_2(\mathsf{X}, \mathsf{Y}, \pi)$. Since the projection $\mathbf{P}$ is linear and nonexpansive under the norm $\| \cdot \|_\pi$, we have

$$\|\hat{\mathbf{T}}f - \hat{\mathbf{T}}\, g\|_\pi = \|\mathbf{P}\mathbf{T}f - \mathbf{P}\mathbf{T}g\|_\pi = \|\mathbf{P}(\mathbf{T}f - \mathbf{T}g)\|_\pi \leqslant \|\mathbf{T}f - \mathbf{T}g\|_\pi \leqslant \theta \|f - g\|_\pi.$$

Hence $\hat{\mathbf{T}}$ is a contraction of modulus θ as claimed. □

Now consider expanding the number of basis elements. In particular, let E_n be the linear span of $\{e_j\}_{j \leqslant n}$ and let $\mathbf{P}_n$ be the orthogonal projection onto E_n. Let f_n be the unique fixed point of $\hat{\mathbf{T}}_n := \mathbf{P}_n \circ \mathbf{T}$, existence of which follows from lemma 5.2. We can now state the following result:

Proposition 5.7. *Let $\mathbf{T}$ be a contraction of modulus θ and let f^* be the unique fixed point of $\mathbf{T}$. If $\{e_j\}_{j \geqslant 0}$ is a complete orthonormal basis of $L_2(\mathsf{X}, \mathsf{Y}, \pi)$, then $\|f_n - f^*\|_\pi \to 0$ as $n \to \infty$.*

[23] See, e.g., chapter 2 of Cheney (2001) or section 1.3 of Atkinson and Han (2009).

Proof. Fix $n \in \mathbb{N}$. By lemma 5.2, $\mathbf{T}$ and $\hat{\mathbf{T}}_1, \hat{\mathbf{T}}_2, \dots$ are all contractions of modulus θ. It then follows from lemma 2.1 of Rust (1996) that $\|f_n - f^*\|_\pi \leqslant \|\hat{\mathbf{T}}_n f^* - f^*\|_\pi / (1-\theta)$. (The norm is different here but the proof is identical.) Since $\hat{\mathbf{T}}_n = \mathbf{P}_n \circ \mathbf{T}$, this becomes

$$\|f_n - f^*\|_\pi \leqslant \frac{\|\mathbf{P}_n f^* - f^*\|_\pi}{1-\theta}.$$

The fact that $\{e_j\}_{j \geqslant 0}$ is a complete orthonormal basis implies $\|\mathbf{P}_n f^* - f^*\|_\pi \to 0$, which in turn gives $\|f_n - f^*\|_\pi \to 0$. □

6. Extensions

In this section we show how the ideas expressed above can be extended in different directions. Section 6.1 connects our analysis to the decomposition in Hansen and Scheinkman (2009), while section 6.2 considers nonstationary models.

6.1. Decompositions.

In this section we show how the decomposition in Hansen and Scheinkman (2009) can be applied to simplify computations by converting a random coefficients model into a partially linear model. The process involves a change of measure. If diagonalizability of the Markov operator is preserved, then the results in section 4.2 can be applied.

To illustrate, consider an asset price equation of the form (3). As in Hansen and Scheinkman (2009), suppose that there exists a positive function k such that $k(X_t, X_{t+1}) = M_{t+1} \frac{D_{t+1}}{D_t}$ for all $t \geqslant 0$. Under some regularity conditions, we can use the decomposition in Hansen and Scheinkman (2009) to rewrite (3) as

$$Y_t = \mathbb{E}_t \left[\lambda \frac{v(X_t)}{v(X_{t+1})} \frac{H_{t+1}}{H_t} (1 + Y_{t+1}) \right],$$

where $\lambda > 0$ is an eigenvalue of the operator $\mathbf{K}$ defined by $\mathbf{K} f(x) = \int k(x, x') f(x') Q(x, \mathrm{d}x')$, v is a π-a.e. positive eigenfunction, and $\{H_t\}$ is an $\mathscr{F}_t$-martingale. Absorbing the martingale component into the expectation gives

$$Y_t = \tilde{\mathbb{E}}_t \left[\lambda \frac{v(X_t)}{v(X_{t+1})} (1 + Y_{t+1}) \right].$$

Dividing by $v(X_t)$ leads to the expression

$$\hat{Y}_t = \lambda \tilde{\mathbb{E}}_t \left[\hat{Y}_{t+1} \right] + \lambda \, \tilde{\mathbb{E}}_t \left[\frac{1}{v(X_{t+1})} \right] \quad \text{where} \quad \hat{Y}_t := \frac{F_t}{v(X_t)}.$$

This is a special case of the partially linear models discussed in section 4.2. If the stability conditions are satisfied (see remark 2.2), then we can apply proposition 4.1, provided that time reversibility is preserved in the change of measure.

The following gives one example where the Hansen–Scheinkman decomposition preserves time reversibility. Let $\{X_t\}$ obey the AR(1) process in section 3.1.1 with $b = 0$. Consumption evolves according to $\log C_{t+1} - \log C_t = g_c + g_x X_t + g_w W_{t+1}$. The stochastic discount factor is

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} = \beta \exp[(1-\gamma)(g_c + g_x X_t + g_w W_{t+1})].$$

Substituting in $W_{t+1} = (\rho X_t - X_{t+1})/\sigma$ shows that we can write M_{t+1} as $k(X_t, X_{t+1})$. Straightforward algebra shows that

$$v(x) := \exp \left\{ \frac{(1-\gamma) g_x}{1-\rho} x \right\} \quad \text{and} \quad \lambda := \beta \exp \left\{ (1-\gamma) g_c + (1-\gamma)^2 \frac{1}{2} \left[g_w + \frac{g_x}{1-\rho} \sigma \right] \right\}$$

provide a positive eigenfunction and corresponding eigenvalue of $\mathbf{K}$. It follows that

$$\frac{H_{t+1}}{H_t} = \lambda^{-1} \frac{v(X_{t+1})}{v(X_t)} \beta \exp[(1-\gamma)(g_c + g_x X_t + g_w W_{t+1})]$$

This leads to

$$\frac{H_{t+1}}{H_t} = \exp \left(\alpha W_{t+1} - \frac{1}{2} \alpha^2 \right) \quad \text{where} \quad \alpha := (1-\gamma) \left[g_w + \frac{\sigma g_x}{1-\rho} \right].$$

The conditional expectation has the form

$$\tilde{\mathbb{E}}_t \left[f(X_{t+1}) \right] = \mathbb{E}_t \left[\exp \left(\alpha W_{t+1} - \frac{1}{2} \alpha^2 \right) f(\rho X_t + \sigma W_{t+1}) \right].$$

Some manipulations show that $\nu(w) := \exp\left(\alpha w - \frac{1}{2}\alpha^2\right) \pi_{(0,1)}(w) = N(\alpha, 1)$. The Markov operator corresponding to $\tilde{\mathbb{E}}_t$ is therefore

$$\begin{aligned}
(\tilde{\mathbf{Q}} f)(x) &= \int \exp \left(\alpha w - \frac{1}{2} \alpha^2 \right) f(\rho x + \sigma w) \pi_{(0,1)}(w) \, \mathrm{d}w \\
&= \int f(\rho x + \sigma w) \nu(w) \, \mathrm{d}w \\
&= \int f(\rho x + \alpha \sigma + \sigma w) \pi_{(0,1)}(w) \, \mathrm{d}w.
\end{aligned}$$

From the results in section 3.1.1, this operator is diagonalizable on $L_2(\mathsf{X}, \mathsf{Y}, \tilde{\pi})$ where $\tilde{\pi}$ is the stationary distribution of $\tilde{\mathbf{Q}}$.

6.2. **Nonstationary Models.** Some degree of nonstationarity in the model can be incorporated, provided that growth is not too fast. For example, consider a version of (11) where

$$Y_t = \mathbb{E}_t \left[a(X_t, X_{t+1}) Y_{t+1} \right] + \varphi_t(X_t), \qquad t \geqslant 0. \tag{50}$$

Here $\{\varphi_t\}_{t \geqslant 0}$ is a sequence of Borel measurable maps from X to Y. Apart from the nonnegative time index, other aspects of the model are the same.

To ensure convergence of the solution, we introduce the following condition:

Assumption 6.1. The function φ_0 is in $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ and there exists a positive constant ϵ such that

$$\epsilon < \frac{1}{r(\mathbf{A})} \quad \text{and} \quad \|\varphi_{t+1}\|_\pi \leqslant \epsilon \|\varphi_t\|_\pi \quad \text{for all } \ t \geqslant 0.$$

To state the main result for this section result, let

$$f_t^* := \sum_{i=0}^{\infty} \mathbf{A}^i \varphi_{t+i} \qquad (t \geqslant 0), \tag{51}$$

where convergence of the series still remains to be proved. (Equality is in the sense of norm convergence in $L_2(\mathsf{X}, \mathsf{Y}, \pi)$.)

Theorem 6.1. *If assumption 6.1 holds, then, for all $t \geqslant 0$, the series (51) converges in $L_p(\mathsf{X}, \mathsf{Y}, \pi)$, and the sequence $\{f_t^*\}$ satisfies*

$$f_t^* = \mathbf{A} f_{t+1}^* + \varphi_t, \qquad \forall\, t \geqslant 0. \tag{52}$$

Moreover, if $Y_t^* = f_t^*(X_t)$ for all t, then $\{Y_t^*\}$ satisfies (50).

Proof of theorem 6.1. Pick any $t \geqslant 0$ and $p \geqslant 1$. Since $r(\epsilon \mathbf{A}) < 1$, it follows from Gelfand's formula that we can find a $\delta < 1$ and $M_0 < \infty$ such that $\|(\epsilon \mathbf{A})^j\| \leqslant \delta^j M_0$ for all $j \geqslant 0$. If we now take $M := M_0 \|\varphi_t\|_\pi$ then

$$\|\mathbf{A}^j \varphi_{t+j}\|_\pi \leqslant \|\mathbf{A}^j\|_\pi \|\varphi_{t+j}\|_\pi \leqslant \|(\epsilon \mathbf{A})^j\|_\pi \|\varphi_t\|_\pi \leqslant \delta^j M \quad \text{for all } \ j \geqslant 0.$$

(Note that M can depend on t but does not depend on j.) Let $g_n := \sum_{i=0}^n \mathbf{A}^i \varphi_{t+i}$. By the triangle inequality in $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ and $\delta < 1$ we have

$$\|g_n\|_\pi \leqslant \sum_{i=0}^n \|\mathbf{A}^i \varphi_{t+i}\|_\pi \leqslant \sum_{i=0}^n \delta^i M \leqslant \frac{M}{1 - \delta}.$$

Hence $g_n \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$ for all n. In fact $\{g_n\}$ is Cauchy in $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ because, for any $n, j \in \mathbb{N}$,

$$\|g_n - g_{n+j}\|_\pi \leqslant \sum_{i=n+1}^{n+j} \|\mathbf{A}^i \varphi_{t+i}\|_\pi.$$

This yields

$$\lim_{n\to\infty} \sup_{j \geqslant n} \|g_n - g_{n+j}\|_\pi \leqslant \lim_{n\to\infty} \sup_{j \geqslant n} \delta^n \frac{1-\delta^j}{1-\delta} M = \lim_{n\to\infty} \delta^n \frac{1}{1-\delta} M = 0.$$

Since $\{g_n\}$ is Cauchy, its limit exists in $L_p(\mathsf{X}, \mathsf{Y}, \pi)$. In particular, the series in (51) converges for any $t \geqslant 0$. Let f_t^* denote this limit. In other words, $\{f_t^*\}$ is the sequence defined in (51).

For f_{t+1}^* we have $f_{t+1}^* = \sum_{i=0}^\infty \mathbf{A}^i \varphi_{t+i+1}$, implying that

$$\varphi_t + \mathbf{A} f_{t+1}^* = \varphi_t + \mathbf{A}\varphi_{t+1} + \mathbf{A}^2 \varphi_{t+2} + \cdots = \sum_{i=0}^\infty \mathbf{A}^i \varphi_{t+i}.$$

In other words, $\varphi_t + \mathbf{A} f_{t+1}^* = f_t^*$. This confirms (52).

Regarding the final claim of theorem 6.1, let $Y_t^* := f_t^*(X_t)$ for all t. Pick any $t \geqslant 0$. Since $\{f_t^*\}$ satisfies (52), we have

$$Y_t^* = f^*(X_t) = \varphi_t(X_t) + (\mathbf{A} f_{t+1}^*)(X_t). \tag{53}$$

The Markov property of $\mathcal{X}$ and the definition of $\mathbf{A}$ imply that

$$(\mathbf{A} f_{t+1}^*)(X_t) = \int f_{t+1}^*(x') a(X_t, x') Q(X_t, \mathrm{d}x') = \mathbb{E}_t[a(X_t, X_{t+1}) f_{t+1}^*(X_{t+1})].$$

Using our definition of $\{Y_t^*\}$ again, we see that

$$(\mathbf{A} f_{t+1}^*)(X_t) = \mathbb{E}_t[a(X_t, X_{t+1}) Y_{t+1}^*].$$

Combining this with (53) confirms (50). □

Example 6.1. Consider the model (2) again, but suppose as in Tallarini (2000) that log dividends follow a linear growth trend. That is, $\ln D_t = \theta t + X_t$, where $\{X_t\}$ is a stationary state process. With the change of variable $U_t := D_t^{1-\gamma} Y_t$, we can write (2) as $U_t = \beta \mathbb{E}_t[U_{t+1} + D_{t+1}^{1-\gamma}]$. This expands out to

$$U_t = \mathbb{E}_t[\beta U_{t+1}] + \beta\, \mathbb{E}_t \exp[(1-\gamma)(\theta(t+1) + X_{t+1})], \tag{54}$$

which is a version of (50) when $\mathsf{Y} = \mathbb{R}$,

$$\varphi_t(x) := \beta \int \exp\left\{(1-\gamma)(\theta(t+1) + x')\right\} Q(x, \mathrm{d}x') \tag{55}$$

and $a(x, x') \equiv \beta$. If π is a Gaussian distribution, one can use (56) and the expression for the mean of a lognormal distribution to show that $\|\varphi_{t+1}\|_\pi = \exp\{\theta(1-\gamma)\} \, \|\varphi_t\|_\pi$. Hence assumption 6.1 holds whenever $\exp\{\theta(1-\gamma)\} < 1/\beta$.

7. Appendix

Let Q be a stochastic kernel on $(\mathsf{X}, \mathscr{B})$ and let π be stationary for Q. Let $\mathbf{Q}$ be the Markov operator. We prove some properties of Q and $\mathbf{Q}$. One is that, if $g \colon \mathsf{X} \to \mathsf{Y}$ is in $L_p(\mathsf{X}, \mathsf{Y}, \pi)$, then

$$\int \mathbf{Q} g \, \mathrm{d}\pi = \int g \, \mathrm{d}\pi. \tag{56}$$

First we prove (56) in the scalar case. To this end, fix any $g \in L_p(\mathsf{X}, \mathbb{R}, \pi)$. We can write $\int \left[\int g(x') Q(x, \mathrm{d}x')\right] \pi(\mathrm{d}x) = \int g(x') \mu(\mathrm{d}x')$ where μ is the measure defined by $\mu(B) = \int Q(x, B) \pi(\mathrm{d}x)$. Since π is stationary for Q we have $\mu = \pi$. Hence (56) holds for scalar valued functions.

Now let Y be a separable Banach space. Let Y^* be the dual space and let $y^* \in \mathsf{Y}^*$. Recall that if μ is a finite measure on X and $f \colon \mathsf{X} \to \mathsf{Y}$ is Bochner integrable, then $\langle \int f \, \mathrm{d}\mu, y^* \rangle = \int \langle f, y^* \rangle \, \mathrm{d}\mu$. It follows that, for any $g \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$,

$$\begin{aligned}
\left\langle \int g \, \mathrm{d}\pi, y^* \right\rangle &= \int \langle g, y^* \rangle \, \mathrm{d}\pi \\
&= \int \int \langle g(x'), y^* \rangle Q(x, \mathrm{d}x') \pi(\mathrm{d}x) \\
&= \left\langle \int \int g(x') Q(x, \mathrm{d}x') \pi(\mathrm{d}x), y^* \right\rangle,
\end{aligned}$$

where the second equality uses the previously established validity of (56) in the scalar case. Since y^* is an arbitrary element of Y^*, we conclude (by the Hahn–Banach theorem) that $\int g \, \mathrm{d}\pi = \int \int g(x') Q(x, \mathrm{d}x') \pi(\mathrm{d}x)$, as was to be shown.

Lemma 7.1. *$\mathbf{Q}$ is a bounded linear self-map on $L_p(\mathsf{X}, \mathsf{Y}, \pi)$ with $r(\mathbf{Q}) = \|\mathbf{Q}\|_\pi = 1$.*

Proof of lemma 7.1. Linearity of $\mathbf{Q}$ is obvious and boundedness is weaker then the claim that $\|\mathbf{Q}\|_\pi = 1$. Hence it suffices to show that $r(\mathbf{Q}) = \|\mathbf{Q}\|_\pi = 1$.

To this end, fix $f \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$ and observe that

$$\begin{aligned}\int \|\mathbf{Q}f\|^p \,\mathrm{d}\pi &= \int \left\| \int f(x')Q(x, \mathrm{d}x') \right\|^p \pi(\mathrm{d}x) \\ &\leqslant \int \int \|f(x')\|^p Q(x, \mathrm{d}x')\pi(\mathrm{d}x) = \int \|f\|^p \,\mathrm{d}\pi,\end{aligned}$$

where we used Jensen's inequality and (56). Raising both sides to the power of $1/p$ gives $\|\mathbf{Q}f\|_\pi \leqslant \|f\|_\pi$, implying that $\|\mathbf{Q}\|_\pi \leqslant 1$. As for an linear operator, we have $r(\mathbf{Q}) \leqslant \|\mathbf{Q}\|$, and hence the proof will be complete if we can show that $r(\mathbf{Q}) \geqslant 1$, which in turn will be established if 1 is an eigenvalue of $\mathbf{Q}$. It is easy to check that if c is a fixed element of Y and $f(x) \equiv c \neq 0$, then $\mathbf{Q}f = f$. Hence 1 is an eigenvalue of $\mathbf{Q}$. $\square$

Proof of lemma 3.1. Let f be in $L_2(\mathsf{X}, \mathsf{Y}, \pi)$. By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\int \|\mathbf{A}f(x)\|^2 \pi(x)\,\mathrm{d}x &= \int \left\| \int A(x, x')f(x')q(x, x')\,\mathrm{d}x' \right\|^2 \pi(x)\,\mathrm{d}x \\ &\leqslant \int \left[\int \|A(x, x')f(x')\| q(x, x')\,\mathrm{d}x' \right]^2 \pi(x)\,\mathrm{d}x \\ &\leqslant \int \left[\int \|A(x, x')\| \, \|f(x')\| q(x, x')\,\mathrm{d}x' \right]^2 \pi(x)\,\mathrm{d}x \\ &= \int \left[\int \|f(x')\| \frac{\|A(x, x')\| q(x, x')}{\pi(x')} \pi(x')\,\mathrm{d}x' \right]^2 \pi(x)\,\mathrm{d}x \\ &\leqslant \int \int \|f(x')\|^2 \pi(x')\,\mathrm{d}x' \int \frac{[\|A(x, x')\| q(x, x')]^2}{\pi(x')^2} \pi(x')\,\mathrm{d}x' \pi(x)\,\mathrm{d}x \\ &= \|f\|_\pi^2 \int \int [\|A(x, x')\| q(x, x')]^2 \frac{\pi(x)}{\pi(x')}\,\mathrm{d}x'\,\mathrm{d}x.\end{aligned}$$

Hence $\|\mathbf{A}f\|_\pi \leqslant \sqrt{\Gamma(A)}\|f\|_\pi$. $\square$

The following discussion provides some preliminaries to the proof of lemma 3.3. Consider the normalized linear state dynamics

$$X_{t+1} = \rho X_t + (1 - \rho^2)^{1/2}\, W_{t+1} \quad \text{where} \quad \{W_t\} \overset{\text{IID}}{\sim} \pi_{(0,1)} := N(0, I) \tag{57}$$

and $-1 < \rho < 1$. Here $\pi_{(0,1)} = N(0, I)$ is the standard normal distribution on $\mathbb{R}$. The corresponding stochastic kernel is $q(x, y) = N(\rho x, (1 - \rho^2)^{1/2})$. It is known to have stationary density $\pi_{(0,1)}$ and orthogonal expansion

$$q(x, y) = \pi_{(0,1)}(y) \sum_{j=0}^{\infty} \rho^j h_j(x) h_j(y), \tag{58}$$

where ρ is the autocorrelation coefficient in (57) and $\{h_j\}$ is the normalized Hermite polynomials. (See, for example, O'Donnell (2014). The representation (58) is called the Mehler expansion of q.) The corresponding Markov operator $\mathbf{Q}_{(0,1)}$ is diagonalizable with

$$\mathbf{Q}_{(0,1)}\, g = \sum_{j=0}^{\infty} \rho^j \langle g, h_j \rangle_\pi h_j \quad \text{for all } g \in L_2(\mathsf{X}, \mathbb{R}, \pi). \tag{59}$$

The operator $\mathbf{Q}_{(0,1)}$ is called the Gaussian noise operator or the Mehler transform.

Proof of lemma 3.3. Writing $\pi_{(0,1)}$ for the standard normal distribution, we can express the relationship between π and $\pi_{(0,1)}$ as $\pi_{(0,1)} = \pi \circ \tau^{-1}$ on $\mathscr{B}$. That is, $\pi_{(0,1)}$ is the image measure of π under τ. It follows from the usual rules for integration over image measures (Dudley, 2002, theorem 4.1.11) that

(a) $f \in L_2(\mathsf{X}, \mathbb{R}, \pi_{(0,1)}) \implies f \circ \tau \in L_2(\mathsf{X}, \mathbb{R}, \pi)$ and $\int f \, \mathrm{d}\pi_{(0,1)} = \int f \circ \tau \, \mathrm{d}\pi$.
(b) $f \in L_2(\mathsf{X}, \mathbb{R}, \pi) \implies f \circ \tau^{-1} \in L_2(\mathsf{X}, \mathbb{R}, \pi_{(0,1)})$ and $\int f \, \mathrm{d}\pi = \int f \circ \tau^{-1} \, \mathrm{d}\pi_{(0,1)}$.

Let $\mathbf{M}$ be the operator from $L_2(\mathsf{X}, \mathbb{R}, \pi_{(0,1)})$ to $L_2(\mathsf{X}, \mathbb{R}, \pi)$ defined by $\mathbf{M} f = f \circ \tau$. By repeated use of (i) and (ii) above, it is straightforward to confirm that $\mathbf{M}$ is a linear bijection that preserves inner products. In other words, $\mathbf{M}$ is a Hilbert space isomorphism. Since $\{h_i\}$ is a complete orthonormal basis of $L_2(\mathsf{X}, \mathbb{R}, \pi_{(0,1)})$, this implies that $\{\mathbf{M} h_i\} = \{e_i\}$ is a complete orthonormal basis of $L_2(\mathsf{X}, \mathbb{R}, \pi)$.

To verify the rest of lemma 3.3 it suffices to show that, for any $i \geqslant 0$, we have $\mathbf{Q} e_i = \rho^i e_i$. Since $\{e_j\}$ is a complete orthonormal basis, to prove this equality it suffices to show that $\langle \mathbf{Q} e_i, e_j \rangle_\pi = \rho^i \mathbb{1}\{i = j\}$ for any fixed j.

To this end, observe that $\mathbf{Q}$ and $\mathbf{Q}_{(0,1)}$ are topologically conjugate under $\mathbf{M}$, in the sense that $\mathbf{Q} = \mathbf{M} \mathbf{Q}_{(0,1)} \mathbf{M}^{-1}$ on $L_2(\mathsf{X}, \mathbb{R}, \pi)$. Indeed, for any $f \in L_2(\mathsf{X}, \mathbb{R}, \pi)$ and any $x \in \mathbb{R}$, we have

$$\begin{aligned}
(\mathbf{M} \mathbf{Q}_{(0,1)} \mathbf{M}^{-1}) f(x) &= \int (f \circ \tau^{-1})(\rho \tau(x) + (1 - \rho^2)^{1/2} w) \pi_{(0,1)}(w) \, \mathrm{d}w \\
&= \int f(\rho x + b + \sigma w) \pi_{(0,1)}(w) \, \mathrm{d}w \\
&= (\mathbf{Q} f)(x).
\end{aligned}$$

Returning to the claim that $\langle \mathbf{Q} e_i, e_j \rangle_\pi = \rho^i \mathbb{1}\{i = j\}$, we have

$$\langle \mathbf{Q} e_i, e_j \rangle_\pi = \left\langle \mathbf{M}\mathbf{Q}_{(0,1)}\mathbf{M}^{-1} e_i, e_j \right\rangle_\pi = \left\langle \mathbf{M}\mathbf{Q}_{(0,1)} h_i, e_j \right\rangle_\pi = \rho^j \langle e_i, e_j \rangle_\pi ,$$

where the last equality is by (59) and $\mathbf{M} h_i = e_i$. The proof is now done. □

Proof of theorem 5.1. Let $\{Y_t^*\}$ be as defined. Regarding the first claim, fix $t \in \mathbb{Z}$ and observe that

$$\mathbb{E}_t \, G(X_t, X_{t+1}, Y_{t+1}^*) = \int G(X_t, x', f^*(x')) Q(X_t, \mathrm{d}x') = \mathbf{T} f^*(X_t).$$

Since f^* is a fixed point, the right hand side equals $f^*(X_t) = Y_t^*$. Hence $\{Y_t^*\}$ is a stationary Markov equilibrium. Moreover, $f^* \in L_p(\mathsf{X}, \mathsf{Y}, \pi)$ and $\mathbb{E}\|Y_t^*\|^p = \|f^*\|_\pi^p$, so the second claim also holds.

Regarding the third claim, let $\{Y_t\}$ be any other stationary Markov equilibrium with finite p-th moment. Fix $t \in \mathbb{Z}$ and $\epsilon > 0$. We will show that

$$[\mathbb{E}\|Y_t - Y_t^*\|^p]^{1/p} < \epsilon. \tag{60}$$

Since ϵ was chosen arbitrarily, (60) implies that $\mathbb{P}\{Y_t = Y_t^*\} = 1$. Moreover, countable intersections of probability one sets are probability one, which proves the third claim in the theorem. Hence we need only show that (60) holds.

To this end, we can use Jensen's inequality and the function κ from the order p contraction property to write

$$\begin{aligned} \|Y_t - Y_t^*\| &\leqslant \mathbb{E}_t \left\| G(X_t, X_{t+1}, Y_{t+1}) - G(X_t, X_{t+1}, Y_{t+1}^*) \right\| \\ &\leqslant \kappa(X_t, X_{t+1}) \, \mathbb{E}_t \left\| Y_{t+1} - Y_{t+1}^* \right\| . \end{aligned}$$

Continuing to iterate in this way gives

$$\|Y_t - Y_t^*\| \leqslant \mathbb{E}_t \prod_{j=0}^{\ell-1} \kappa(X_{t+j}, X_{t+1+j}) \|Y_{t+\ell} - Y_{t+\ell}^*\|.$$

Because $\{Y_t\}$ is a stationary Markov equilibrium, there is a fixed function $g \colon \mathsf{X} \to \mathsf{Y}$ such that $Y_{t+\ell} = g(X_{t+\ell})$. This means that

$$\begin{aligned} \|Y_t - Y_t^*\| &\leqslant \mathbb{E}_t \prod_{j=0}^{\ell-1} \kappa(X_{t+j}, X_{t+1+j}) \|g(X_{t+\ell}) - f^*(X_{t+\ell})\| \\ &= \mathbf{K}^\ell h(X_{t+\ell}) \quad \text{where } h(x) := \|g(x) - f^*(x)\|. \end{aligned}$$

Raising to the power of p and taking expectations gives

$$\mathbb{E}\|Y_t - Y_t^*\|^p \leqslant \|\mathbf{K}^\ell h\|_\pi^p \leqslant \|\mathbf{K}^\ell\|_\pi^p \, \|h\|_\pi^p = \|\mathbf{K}^\ell\|_\pi^p \, \|f^* - g\|_\pi^p. \tag{61}$$

Since $(G, \mathcal{X})$ is order p contracting, the operator $\mathbf{K}$ defined in (43) has spectral radius less than 1. Gelfand's formula then implies the existence of an $m \in \mathbb{N}$ such that $\|\mathbf{K}^m\|_\pi < 1$. Choosing $n \in \mathbb{N}$ such that $\|\mathbf{K}^m\|_\pi^n < \epsilon / \|f^* - g\|_\pi$ and setting $\ell = mn$ in (61), we have

$$(\mathbb{E}\|Y_t - Y_t^*\|^p)^{1/p} \leqslant \|\mathbf{K}^{nm}\|_\pi \, \|f^* - g\|_\pi \leqslant \|\mathbf{K}^m\|_\pi^n \, \|f^* - g\|_\pi < \epsilon.$$

This confirms (60), and the proof is now done. □

Proof of proposition 5.5. Fix $f \in L_2(\mathsf{X}, \mathsf{Y}, \pi)$ and $x \in \mathsf{X}$. By Jensen's inequality and the conditions of proposition 5.5 we have

$$\begin{aligned}
\|\mathbf{T} f(x)\|^2 &= \left\| \int G(x, x', f(x')) q(x, x') \, \mathrm{d}x' \right\|^2 \\
&\leqslant \left[\int \|G(x, x', f(x'))\| q(x, x') \, \mathrm{d}x' \right]^2 \\
&\leqslant \left\{ \int \kappa(x, x') \|f(x')\| q(x, x') \, \mathrm{d}x' + \int \zeta(x, x') q(x, x') \, \mathrm{d}x' \right\}^2 \\
&\leqslant 2 \left\{ \int \kappa(x, x') \|f(x')\| q(x, x') \, \mathrm{d}x' \right\}^2 + 2 \left\{ \int \zeta(x, x') q(x, x') \, \mathrm{d}x' \right\}^2 \\
&\leqslant 2 \left\{ \int \kappa(x, x') \|f(x')\| q(x, x') \, \mathrm{d}x' \right\}^2 + 2 \int \zeta(x, x')^2 q(x, x') \, \mathrm{d}x'.
\end{aligned}$$

From the Cauchy-Schwarz inequality,

$$\begin{aligned}
\left\{ \int \kappa(x, x') \|f(x')\| q(x, x') \, \mathrm{d}x' \right\}^2 &= \left[\int \|f(x')\| \frac{\kappa(x, x') q(x, x')}{\pi(x')} \pi(x') \, \mathrm{d}x' \right]^2 \\
&\leqslant \left[\int \|f(x')\|^2 \pi(x') \, \mathrm{d}x' \int \frac{[\kappa(x, x') q(x, x')]^2}{\pi(x')^2} \pi(x') \, \mathrm{d}x' \right]^2 \\
&= \|f\|_\pi^2 \int \frac{[\kappa(x, x') q(x, x')]^2}{\pi(x')} \, \mathrm{d}x'.
\end{aligned}$$

Hence

$$\|\mathbf{T} f(x)\|^2 \leqslant 2 \|f\|_\pi^2 \int \frac{[\kappa(x, x') q(x, x')]^2}{\pi(x')} \, \mathrm{d}x' + 2 \int \zeta(x, x')^2 q(x, x') \, \mathrm{d}x'.$$

Integrating with respect to $\pi(x) \, \mathrm{d}x$ now gives

$$\|\mathbf{T} f\|_\pi^2 \leqslant 2 \|f\|_\pi \int \int [\kappa(x, x') q(x, x')]^2 \frac{\pi(x)}{\pi(x')} \, \mathrm{d}x' \, \mathrm{d}x + 2 \int \int \zeta(x, x')^2 q(x, x') \pi(x) \, \mathrm{d}x \, \mathrm{d}x'.$$

That is, $\|\mathbf{T} f\|_\pi^2 \leqslant 2 \|f\|_\pi \Gamma(\kappa) + 2 \mathbb{E} \zeta(X_t, X_{t+1})^2$. The right hand side is finite, so $\mathbf{T} f$ is in $L_2(\mathsf{X}, \mathsf{Y}, \pi)$. Hence $(G, \mathcal{X})$ is second order regular. □

References

Abel, A. B. (1990): "Asset prices under habit formation and catching up with the Joneses," *The American Economic Review*, 38–42.

Al-Sadoon, M. M. (2016): "The Linear Systems Approach to Linear Rational Expectations Models," Tech. Rep. 875, Barcelona Graduate School of Economics.

Aliprantis, C. D. and K. C. Border (2007): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer Science & Business Media.

Alvarez, F. and N. L. Stokey (1998): "Dynamic programming with homogeneous functions," *Journal of economic theory*, 82, 167–189.

Anderson, E. W. (2005): "The dynamics of risk-sensitive allocations," *Journal of Economic theory*, 125, 93–150.

Atkinson, K. and W. Han (2009): *Theoretical Numerical Analysis: A Functional Analysis Framework*, vol. 39, Springer Science & Business Media.

Atkinson, K. E. (1967): "The numerical solution of the eigenvalue problem for compact integral operators," *Trans. Amer. Math. Soc*, 129, 458–465.

Benhabib, J., A. Bisin, and S. Zhu (2011): "The distribution of wealth and fiscal policy in economies with finitely lived agents," *Econometrica*, 79, 123–157.

Blanchard, O. J. and C. M. Kahn (1980): "The solution of linear difference models under rational expectations," *Econometrica*, 1305–1311.

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

Boyd, S., P. Diaconis, P. Parrilo, and L. Xiao (2005): "Symmetry analysis of reversible Markov chains," *Internet Mathematics*, 2, 31–71.

Brandt, A. (1986): "The stochastic equation Yn+ 1= AnYn+ Bn with stationary coefficients," *Advances in Applied Probability*, 211–220.

Brogueira, J. and F. Schütze (2015): "Existence and uniqueness of equilibrium in Lucas' Asset Pricing model when utility is unbounded," .

Cagan, P. (1956): "The Dynamics of Hyper-Inflation, in M. Friedman (ed.), Studies of the Quantity Theory of Money, Chicago University Press," .

Calin, O. L., Y. Chen, T. F. Cosimano, and A. A. Himonas (2005): "Solving asset pricing models when the price–dividend function is analytic," *Econometrica*, 73, 961–982.

Cheney, W. (2001): *Analysis for Applied Mathematics*, vol. 208, Springer Science & Business Media.

Christiano, L. J. (2002): "Solving dynamic equilibrium models by a method of undetermined coefficients," *Computational Economics*, 20, 21–55.

Debnath, L. and P. Mikusinski (2005): *Introduction to Hilbert Spaces with Applications*, Academic Press.

Dudley, R. M. (2002): *Real Analysis and Probability*, vol. 74, Cambridge University Press.

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

Farmer, R. E., D. F. Waggoner, and T. Zha (2009): "Understanding Markov-switching rational expectations models," *Journal of Economic Theory*, 144, 1849–1867.

Gouriéroux, C. and J. Jasiak (2006): "Autoregressive gamma processes," *Journal of Forecasting*, 25, 129–152.

Griffiths, B. (2009): "Stochastic processes with orthogonal polynomial eigenfunctions," *Journal of computational and applied mathematics*, 233, 739–744.

Häggström, O. (2002): *Finite Markov chains and algorithmic applications*, vol. 52, Cambridge University Press.

Hansen, L. P. and T. J. Sargent (1980): "Formulating and estimating dynamic linear rational expectations models," *Journal of Economic Dynamics and control*, 2, 7–46.

——— (1995): "Discounted linear exponential quadratic gaussian control," *IEEE Transactions on Automatic control*, 40, 968–971.

Hansen, L. P. and J. A. Scheinkman (2009): "Long-Term Risk: An Operator Approach," *Econometrica*, 77, 177–234.

Kantorovich, L. and G. Akilov (1982): *Functional Analysis*, Pergamon Press, Oxford, second ed.

Kesten, H. (1973): "Random difference equations and renewal theory for products of random matrices," *Acta Mathematica*, 131, 207–248.

Khare, K. and H. Zhou (2009): "Rates of convergence of some multivariate Markov chains with polynomial eigenfunctions," *The Annals of Applied Probability*, 19, 737–777.

Klein, P. (2000): "Using the generalized Schur form to solve a multivariate linear rational expectations model," *Journal of Economic Dynamics and Control*, 24, 1405–1423.

Labadie, P. (1986): "Comparative dynamics and risk premia in an overlapping generations model," *The Review of Economic Studies*, 53, 139–152.

LONGLA, M. AND M. PELIGRAD (2012): "Some aspects of modeling dependence in copula-based Markov chains," *Journal of Multivariate Analysis*, 111, 234–240.

LUCAS, R. E. (1972): "Expectations and the Neutrality of Money," *Journal of economic theory*, 4, 103–124.

——— (1976): "Econometric policy evaluation: A critique," in *Carnegie-Rochester conference series on public policy*, North-Holland, vol. 1, 19–46.

——— (1978): "Asset prices in an exchange economy," *Econometrica*, 1429–1445.

LUCAS, R. E. AND N. L. STOKEY (1987): "Money and Interest in a Cash-in-Advance Economy," *Econometrica: Journal of the Econometric Society*, 491–513.

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

MATKOWSKI, J. AND A. S. NOWAK (2011): "On discounted dynamic programming with unbounded returns," *Economic Theory*, 46, 455–474.

MEHRA, R. AND E. C. PRESCOTT (1985): "The equity premium: A puzzle," *Journal of Monetary Economics*, 15, 145–161.

MEYN, S. AND R. L. TWEEDIE (2009): "Markov chains and stochastic stability," .

MUTH, J. F. (1961): "Rational expectations and the theory of price movements," *Econometrica: Journal of the Econometric Society*, 315–335.

O'DONNELL, R. (2014): *Analysis of boolean functions*, Cambridge University Press.

RINCÓN-ZAPATERO, J. P. AND C. RODRÍGUEZ-PALMERO (2003): "Existence and uniqueness of solutions to the Bellman equation in the unbounded case," *Econometrica*, 71, 1519–1555.

RUST, J. (1996): "Numerical dynamic programming in economics," *Handbook of computational economics*, 1, 619–729.

SARGENT, T. J. AND M. WALLACE (1983): "A model of commodity money," *Journal of Monetary Economics*, 12, 163–187.

SIMS, C. A. (2002): "Solving linear rational expectations models," *Computational economics*, 20, 1–20.

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

TALLARINI, T. D. (2000): "Risk-sensitive real business cycles," *Journal of Monetary Economics*, 45, 507–532.

TAUCHEN, G. AND R. HUSSEY (1991): "Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models," *Econometrica*, 371–396.

TAYLOR, J. B. (1977): "Conditions for unique solutions in stochastic macroeconomic models with rational expectations," *Econometrica: Journal of the Econometric*

Society, 1377–1385.

Uhlig, H. (1999): "A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon and S. A, Oxford University Press.