Discussion Paper Series

RIEB

Kobe University

DP2011-32

Existence, Stability and Computation of Stationary Distributions: An Extension of the Hopenhayn-Prescott Theorem

Takashi KAMIHIGASHI
John STACHURSKI

December 2, 2011

Research Institute for Economics and Business Administration
Kobe University
2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Existence, Stability and Computation of Stationary Distributions: An Extension of the Hopenhayn-Prescott Theorem*

Takashi Kamihigashi[a] and John Stachurski[b]

[a] *Research Institute for Economics and Business Administration, Kobe University*

[b] *Research School of Economics, Australian National University*

December 2, 2011

Abstract

This paper strengthens the Hopenhayn and Prescott stability theorem for monotone economies. We extend the theorem to a larger class of applications, and develop new perspectives on the nature and causes of stability and instability. In addition, we show that models satisfying the Hopenhayn-Prescott theorem are ergodic, in the sense that sample averages of time series converge with probability one to their corresponding expectations under the stationary distribution, independent of initial conditions.

JEL Classification: C62, C63
Keywords: Stability, simulation, stationary equilibria

---

*We thank Yiyong Cai, Andrew McLennan, Kazuo Nishimura, Kenji Sato, and Cuong Le Van for many helpful comments. Financial support from the Japan Society for the Promotion of Science and Australian Research Council Discovery Grant DP120100321 is gratefully acknowledged. *Email addresses:* tkamihig@rieb.kobe-u.ac.jp, john.stachurski@anu.edu.au

1 Introduction

Hopenhayn and Prescott's stability theorem [28, theorem 2] is a standard tool for analysis of dynamics and stationary equilibria. For example, Huggett [30] used the theorem to study asset distributions in incomplete-market economies with infinitely-lived agents.[1] The theorem was also applied to variants of Huggett's model with features such as habit formation, endogenous labor supply, capital accumulation and international trade [22, 31, 53, 42, 55]. It was used to study the classical one-sector optimal growth model by Hopenhayn and Prescott [28], a stochastic endogenous growth model by de Hek [21], and a small open economy by Chatterjee and Shukayev [15]. It has been used in a wide range of OLG models with features such as credit rationing [1, 54], human capital [52, 40, 13, 17, 12], international trade [56, 18], nonconcave production [47], and occupational choice [40, 4, 5]. Other well-known applications include variants of Hopenhayn and Rogerson's [29] model of job turnover [11, 59], as well as variants of Hopenhayn's [27] model of entry and exit [16, 58].

The contributions of this paper are twofold. First, we show that the conditions of Hopenhayn and Prescott's stability theorem have an additional implication beyond existence, uniqueness and stability of stationary equilibria: They also imply ergodicity, in the sense that sample averages of time series converge with probability one to their corresponding expectations under the stationary distribution, independent of initial conditions. This is particularly important from a computational perspective, implying as it does that stationary cross-sectional probabilities can be calculated by time averages, where the time series is started from any point in the state space. In particular, we need no prior information about the stationary distribution to compute stationary outcomes via simulation.

Our second contribution is to weaken the conditions of the theorem, and show how it can be applied to a broader range of models. We do this by introducing a mixing condition called order reversing that is weaker than the mixing condition used by Hopenhayn and Prescott. We also relax the restriction that the state space be compact and order bounded. In this setting, we obtain general conditions for monotone, order reversing processes to attain global stability. The conditions are also necessary, and hence we are able to characterize global stability for monotone, order reversing processes. Finally, we show that the ergodicity results discussed

[1] See [32] for a recent extension of Huggett's analysis.

above are valid under these conditions.

One reason that the Hopenhayn-Prescott theorem has not previously been extended is that their proof of the existence of a stochastic steady state uses the Knaster-Tarski fixed point theorem, and for non-compact state spaces the Knaster-Tarski theorem cannot be applied, since a chain in the space of distributions need not have a supremum or an infimum. Our fixed point argument is new, combining order-theoretic and topological results to obtain existence of the stochastic steady state.

Freeing Hopenhayn and Prescott's existence result from the Knaster-Tarski theorem has the obvious benefit of permitting more general state spaces. It also relaxes a tension that is present in the original formulation of the Hopenhayn-Prescott stability theorem. In the original formulation, a compact state space is needed to apply the Knaster-Tarski theorem, which yields existence of a stationary distribution. On the other hand, the restriction to compact state spaces requires that shocks have relatively small supports, which in turn implies less mixing. Since mixing is associated with uniqueness and stability of stationary distributions, reduced mixing means that these properties are less likely to hold.[2]

As the conditions in this paper yield existence without the compact state space, they permit the study of models with larger shocks and more mixing. This makes it possible to prove existence, uniqueness and stability for models where even existence cannot be established using earlier techniques. It also allows researchers to address new questions, such as whether or not large shocks are destabilizing.

Some of these ideas are illustrated in our applications at the end of the paper. The applications include an infinite horizon optimal growth model, an overlapping generations model and a nonlinear autoregression. In all three applications, we illustrate situations where the conditions of our extended Hopenhayn-Prescott theorem are satisfied, while those of the theorem in its original formulation do not hold. As far as we are aware, no current theory from the literature on Markov processes can be used to obtain existence of a stationary distribution in these cases.

Our applications also shed some light on the extent to which large shocks are destabilizing. Our results suggest that, provided that the fundamentals of the model act against divergence, large shocks are not destabilizing. On the contrary, large shocks generate mixing, and mixing promotes stability.

Concerning related literature, the sequence of results leading to the Hopenhayn-Prescott stability theorem began with the seminal contribution of Razin and Yahav

[2]For a discussion of the relationship between mixing and stability see Stokey *et al.* [62, p. 380].

[57]. Razin and Yahav introduced a condition now called the monotone mixing condition (MMC), and showed that the MMC implies global stability for monotone and suitably continuous Markov processes evolving on a closed interval of $\mathbb{R}$. Stokey, Lucas and Prescott [62] generalized this result to multiple dimensions.

Meyn and Tweedie [44] and Herández-Lerma and Lasserre [25] give excellent overviews of the classical theory of discrete time Markov processes. For monotone Markov models, important contributions to the theory aside from those mentioned above were provided by Dubins and Freedman [23], Bhattacharya and Lee [7], Bhattacharya and Majumdar [8] and Bhattacharya *et al.* [9]. These authors studied stability in the monotone setting via a "splitting condition," defined in terms of an ordering on the state space. This condition is stricter than order reversing. At the same time, the literature on splitting contains many important results not treated in this paper.[3]

The rest of the paper is structured as follows: Section 2 reviews some basic definitions concerning Markov processes, and introduces the concept of order reversing. Section 3 states the main results (theorems 3.1–3.4), and compares them to the original formulation of the Hopenhayn-Prescott stability theorem. Section 4 provides sufficient conditions for order reversing, and other results useful for checking the conditions of theorems 3.1–3.4. Section 5 gives applications and section 6 concludes.

2 Preliminaries

At each time $t = 0, 1, \ldots,$ the state of the economy is described by a point X_t in topological space S. The space S is equipped with its Borel sets $\mathscr{B}_S$ and a closed partial order $\leq$. An order interval of S is a set of the form $[a, b] := \{x \in S : a \leq x \leq b\}$. A function $f \colon S \to \mathbb{R}$ is called *increasing* if $f(x) \leq f(y)$ whenever $x \leq y$. A subset B of S is called *order bounded* if there exists an order interval $[a, b] \subset S$ with $B \subset [a, b]$. In addition, B is called *increasing* if its indicator function $\mathbb{1}_B$ is increasing, and *decreasing* if $\mathbb{1}_B$ is decreasing.

To simplify terminology, we often use the word "distribution" to mean "probability measure on $(S, \mathscr{B}_S)$". The set of all probability measures on $(S, \mathscr{B}_S)$ will be denote by $\mathscr{P}_S$. We let cbS denote the continuous bounded functions from S to $\mathbb{R}$, and ibS denote the set of increasing bounded measurable functions from S to $\mathbb{R}$. We

[3]A recent technical note by Kamihigashi and Stachurski [36] also studied monotone Markov processes. The result in that note is used in the proof of our main stability result.

use inner product notation to represent integration, so that

$$\langle \mu, h \rangle := \int h(x)\mu(dx) \quad \text{for } \mu \in \mathscr{P}_S \text{ and } h \in ibS \cup cbS.$$

We adopt the standard definitions of convergence in distribution and stochastic domination: Given sequence $\{\mu_n\}_{n=0}^{\infty}$ in $\mathscr{P}_S$, we say that μ_n *converges to* μ and write $\mu_n \to \mu_0$ if $\langle \mu_n, h \rangle \to \langle \mu_0, h \rangle$ for all $h \in cbS$. We say that μ_2 *stochastically dominates* μ_1 and write $\mu_1 \preceq \mu_2$ if $\langle \mu_1, h \rangle \leq \langle \mu_2, h \rangle$ for all $h \in ibS$.

Following Hopenhayn and Prescott [28], we assume that S is a normally ordered polish space.[4] Hopenhayn and Prescott assume in addition that S is compact, with least element a and greatest element b.[5] Since we wish to include more general state spaces such as $\mathbb{R}^n$, we make the weaker assumption that a subset of S is compact if and only if it is closed and order bounded. This is obviously the case in Hopenhayn and Prescott's setting, where all subsets of S are order bounded, and any closed subset is compact. It also holds for $S = \mathbb{R}^n$ with its standard partial order, since order boundedness is then equivalent to boundedness. In addition, it holds in common state spaces such as $\mathbb{R}^n_+$ or $\mathbb{R}^n_{++}$, or in any set of the form $I_1 \times \cdots \times I_n \subset \mathbb{R}^n$, where each I_i is an open, closed, half-open or half-closed interval in $\mathbb{R}$.[6]

Throughout the paper, we suppose that the model under consideration is time-homogeneous and Markovian. The dynamics of such a model can be summarized by a stochastic kernel Q, where $Q(x, B)$ represents the probability that the state moves from $x \in S$ to $B \in \mathscr{B}_S$ in one unit of time. As usual, we require that $Q(x, \cdot) \in \mathscr{P}_S$ for each $x \in S$, and that $Q(\cdot, B)$ is measurable for each $B \in \mathscr{B}_S$.

Here and below, $(\Omega, \mathscr{F}, \mathbb{P})$ denotes a fixed probability space on which all random variables are defined, and $\mathbb{E}$ is the corresponding expectations operator. Given $\mu \in \mathscr{P}_S$ and stochastic kernel Q, an S-valued stochastic process $\{X_t\}$ is called *Markov-*(Q, μ) if X_0 has distribution μ and $Q(x, \cdot)$ is the conditional distribution of X_{t+1} given $X_t = x$.[7] If μ is the distribution $\delta_x \in \mathscr{P}_S$ concentrated on $x \in S$, we

[4] A polish space is a separable and completely metrizable topological space. The space $(S, \leq)$ is normally ordered if, given any closed increasing set I and closed decreasing set D with $I \cap D = \emptyset$, there exists an f in $ibS \cap cbS$ such that $f(x) = 0$ for all $x \in D$ and $f(x) = 1$ for all $x \in I$.

[5] A point a is called a *least element* of S if $a \in S$ and $a \leq x$ for all $x \in S$. A point b is called a *greatest element* of S if $b \in S$ and $x \leq b$ for all $x \in S$.

[6] A simple example that does not satisfy our assumptions is $S = (0, 1) \cup (2, 3)$. In this case the order interval $[0.5, 2.5]$ is closed and order bounded but not compact.

[7] More formally, $\mathbb{P}[X_{t+1} \in B \,|\, \mathscr{F}_t] = Q(X_t, B)$ almost surely for all $B \in \mathscr{B}_S$, where $\mathscr{F}_t$ is the σ-algebra generated by the history $X_0, \ldots, X_t$.

call $\{X_t\}$ *Markov-*(Q, x). We call $\{X_t\}$ *Markov-*Q if $\{X_t\}$ is Markov-(Q, μ) for some $\mu \in \mathscr{P}_S$.

Example 2.1. Many economic models result in processes for the state variables represented by nonlinear, vector-valued stochastic difference equations. As a generic example, consider the S-valued process

$$X_{t+1} = F(X_t, \xi_{t+1}), \qquad \{\xi_t\} \overset{\text{IID}}{\sim} \phi, \tag{1}$$

where $\{\xi_t\}$ takes values in $Z \subset \mathbb{R}^m$, the function $F\colon S \times Z \to S$ is measurable, and ϕ is a probability measure on the Borel sets of Z. Let Q_F be the kernel

$$Q_F(x, B) := \mathbb{P}\{F(x, \xi_t) \in B\} = \phi\{z \in Z : F(x, z) \in B\}. \tag{2}$$

Then $\{X_t\}$ in (1) is Markov-Q_F.[8]

For each $t \in \mathbb{N}$, let Q^t be the t-th order kernel, defined by

$$Q^1 := Q, \quad Q^t(x, B) := \int Q^{t-1}(y, B) Q(x, dy) \qquad (x \in S,\ B \in \mathscr{B}_S).$$

The value $Q^t(x, B)$ represents the probability of transitioning from x to B in t steps.

A sequence $\{\mu_n\} \subset \mathscr{P}_S$ is called *tight* if, for all $\epsilon > 0$, there exists a compact $K \subset S$ such that $\mu_n(K) \geq 1 - \epsilon$ for all n. A stochastic kernel Q is called *bounded in probability* if $\{Q^t(x, \cdot)\}_{t \geq 0}$ is tight for all $x \in S$. Intuitively, Q is bounded in probability if, for any initial condition, the entire sequence of distributions is almost supported on a single compact set, and hence probability mass does not diverge as $n \to \infty$.

For each Q we define two operators, sometimes called the left and right Markov operators. The left Markov operator maps $\mu \in \mathscr{P}_S$ into $\mu Q \in \mathscr{P}_S$, where

$$(\mu Q)(B) := \int Q(x, B) \mu(dx) \qquad (B \in \mathscr{B}_S). \tag{3}$$

The right Markov operator maps bounded measurable function $h\colon S \to \mathbb{R}$ into bounded measurable function Qh, where

$$(Qh)(x) := \int h(y) Q(x, dy) \qquad (x \in S).$$

[8] Although the process (1) is only first order, models including higher order lags of the state and shock process can be rewritten in the form of (1) by redefining the state variables.

The interpretation of the left Markov operator $\mu \mapsto \mu Q$ is that it shifts the distribution for the state forward by one time period. In particular, if $\{X_t\}$ is Markov-(Q, μ), then μQ^t is the distribution of X_t. The interpretation of the right Markov operator $h \mapsto Qh$ is that $(Q^t h)(x)$ is the expectation of $h(X_t)$ given $X_0 = x$. If Q_F is the kernel in (2), then $(Q_F h)(x) = \int h[F(x,z)]\phi(dz)$. It is well known (see, e.g., Stokey *et al.* [62, p. 219]) that the left and right Markov operators are adjoint, in the sense that, for any such h and any $\mu \in \mathscr{P}_S$, we have $\langle \mu, Qh \rangle = \langle \mu Q, h \rangle$. Also, given any $x \in S$, $B \in \mathscr{B}_S$ and $t \in \mathbb{N}$, the t-th order kernel and the left and right Markov operators are related by $Q^t(x, B) = (\delta_x Q^t)(B) = (Q^t \mathbb{1}_B)(x)$. Here $\mathbb{1}_B$ is the indicator function of B.

If $\mu^* \in \mathscr{P}_S$ and $\mu^* Q = \mu^*$, then μ^* is called *stationary* (or *invariant*) for Q. If Q has a unique stationary distribution μ^* in $\mathscr{P}_S$, and, in addition, $\mu Q^t \to \mu^*$ as $t \to \infty$ for all $\mu \in \mathscr{P}_S$, then Q is called *globally stable*. In this case, μ^* is naturally interpreted as the long-run equilibrium of the economic system. If μ^* is stationary, then any Markov-(Q, μ^*) process $\{X_t\}$ is strict-sense stationary with $X_t \sim \mu^*$ for all t.

If $\mu \in \mathscr{P}_S$ and $\mu Q \preceq \mu$, then μ is called *excessive*. If $\mu \preceq \mu Q$, then μ is called *deficient*. If Q satisfies $\mu Q \preceq \mu' Q$ whenever $\mu \preceq \mu'$, then Q is called *increasing*.[9] It is in fact sufficient to check that $Q(x, \cdot) \preceq Q(x', \cdot)$ whenever $x \leq x'$. A third equivalent condition is that $Qh \in ibS$ whenever $h \in ibS$.[10] If, on the other hand, $Qh \in cbS$ whenever $h \in cbS$, then Q is called *Feller*. Finally, if Q and Q' are two kernels with $\mu Q \preceq \mu Q'$ for all $\mu \in \mathscr{P}_S$, then we say that Q' *dominates* Q and write $Q \preceq Q'$. An equivalent condition is that $Qh \leq Q'h$ for all $h \in ibS$.[11]

Remark 2.1. Let Q be an increasing stochastic kernel. If A is an increasing set, then $x \mapsto Q(x, A)$ is increasing. If A is a decreasing set, then $x \mapsto Q(x, A)$ is decreasing.

[9]Many examples of models with increasing kernels were given in the introduction. Other examples not discussed there include various infinite horizon optimal growth models with features such as irreversible investment, renewable resources, distortions, and capital-dependent utility. Increasing kernels are also found in stochastic OLG models besides those mentioned previously, such as models with limited commitment, and in a variety of stochastic games. See, for example, [2, 3, 26, 6, 49, 51, 19, 45]. For an empirical test of the increasing property, see [38].

[10]For example, suppose that $Qh \in ibS$ whenever $h \in ibS$. Pick any $\mu \in \mathscr{P}_S$. Fixing $h \in ibS$, we have $Qh \leq Q'h$. Integrating with respect to μ gives $\langle \mu, Qh \rangle \leq \langle \mu, Q'h \rangle$, or, equivalently, $\langle \mu Q, h \rangle \leq \langle \mu Q', h \rangle$. Since h was an arbitrary element of ibS, we have shown that $\mu Q \preceq \mu Q$. The proof of the converse is also straightforward.

[11]If g and h are two real-valued functions on S, then $g \leq h$ means that $g(x) \leq h(x)$ for all $x \in S$.

Remark 2.2. If S has a least element a, then δ_a is deficient for any kernel Q, because $\delta_a \preceq \mu$ for every $\mu \in \mathscr{P}_S$, and hence $\delta_a \preceq \delta_a Q$. Similarly, if S has a greatest element b, then δ_b is excessive for Q.

Example 2.2. Let F and Q_F be as in example 2.1. If $x \mapsto F(x,z)$ is increasing, then Q_F is increasing.[12] If $x \mapsto F(x,z)$ is continuous, then Q_F is Feller.

Example 2.3. Let F and Q_F be as in example 2.1. Consider a second process

$$X_{t+1} = G(X_t, \xi_{t+1}), \qquad \{\xi_t\} \overset{\text{IID}}{\sim} \phi,$$

where $G\colon S \times Z \to S$ is measurable. Let Q_G be the corresponding stochastic kernel. If $G(x,z) \leq F(x,z)$ for all $(x,z) \in S \times Z$, then $Q_G \preceq Q_F$.[13]

2.1 Order Reversing

Our first step is to introduce a new order-theoretic mixing condition. We will say that a stochastic kernel Q is *order reversing* if, given any x and x' in S with $x' \leq x$, and independent Markov-Q processes $\{X_t\}$ and $\{X'_t\}$ starting at x and x' respectively, there exists a $t \in \mathbb{N}$ with $\mathbb{P}\{X_t \leq X'_t\} > 0$. In other words, the initial ordering is reversed at some point in time with positive probability.

It is helpful to provide a second, more succinct, definition. To do so, let

$$\mathbb{G} := \text{graph}(\leq) := \{(y, y') \in S \times S : y \leq y'\},$$

so that $y \leq y'$ iff $(y, y') \in \mathbb{G}$. Also, let Q be a stochastic kernel on S, and consider the product kernel $Q \times Q$ on $S \times S$ defined by

$$(Q \times Q)((x, x'), A \times B) = Q(x, A) Q(x', B) \tag{4}$$

for $(x, x') \in S \times S$ and $A, B \in \mathscr{B}_S$.[14] The product kernel represents the stochastic

[12] The statement that $F(\cdot, z)$ is increasing means that $x, x' \in S$ with $x \leq x'$ and $z \in Z$ implies $F(x,z) \leq F(x',z)$. Since $Q_F h(x) = \int h[F(x,z)]\phi(dz)$, to prove that Q_F is increasing, it suffices to show that if $x \leq x'$ and $h \in ibS$, then $\int h[F(x,z)]\phi(dz) \leq \int h[F(x',z)]\phi(dz)$. As $h \in ibS$ and $F(\cdot, z)$ is increasing for each z, this follows from monotonicity of the integral.

[13] To see this, observe that if $h \in ibS$, then $Q_G h \leq Q_F h$, since, for any $x \in S$, we have $(Q_G h)(x) = \int h[G(x,z)]\phi(dz) \leq \int h[F(x,z)]\phi(dz) = (Q_F h)(x)$.

[14] Sets of the form $A \times B$ with $A, B \in \mathscr{B}_S$ provide a semi-ring in the product σ-algebra $\mathscr{B}_S \otimes \mathscr{B}_S$ that also generates $\mathscr{B}_S \otimes \mathscr{B}_S$. Defining the probability measure $Q((x,x'), \cdot)$ on this semi-ring uniquely defines $Q((x,x'), \cdot)$ on all of $\mathscr{B}_S \otimes \mathscr{B}_S$. See, e.g., [24, theorem 3.2.7].

kernel of the joint process $\{(X_t, X_t')\}$ when $\{X_t\}$ and $\{X_t'\}$ are independent Markov-Q processes. Using this notation, Q is order reversing if and only if

$$\forall\, x, x' \in S \text{ with } x' \leq x, \;\; \exists\, t \in \mathbb{N} \text{ such that } (Q \times Q)^t((x, x'), \mathbb{G}) > 0.$$

This second definition emphasizes the fact that order reversing is a property of the kernel Q alone (taking S and $\leq$ as given).

Remark 2.3. In verifying order reversing, it is clearly sufficient to check the existence of a t with $(Q \times Q)^t((x, x'), \mathbb{G}) > 0$ for *arbitrary* pair $x, x' \in S$. Often this is just as easy, and much of the following discussion proceeds accordingly.

Example 2.4. Suppose we are studying a dynamic model of household wealth. Informally, the model is order reversing if, for two households receiving idiosyncratic shocks from the same distribution, the wealth of the first household is less than that of the second at some point in time with non-zero probability, regardless of initial wealth for each of the two households.

Example 2.5. Let S be a compact metric space with least element a and greatest element b, and let Q be an increasing kernel on S. In this setting, Q is said to satisfy the MMC [57, 62, 28] whenever

$$\exists\, \bar{x} \in S \text{ and } k \in \mathbb{N} \text{ such that } Q^k(a, [\bar{x}, b]) > 0 \text{ and } Q^k(b, [a, \bar{x}]) > 0. \tag{5}$$

Under these conditions, Q is order reversing: If we start independent Markov-Q process $\{X_t^a\}$ and $\{X_t^b\}$ at a and b respectively, then (5) implies the order reversal $X_k^b \leq X_k^a$ occurs at time k with positive probability. Since Q is increasing, closer initial conditions only make this event more likely.[15]

Example 2.6. Consider the stochastic kernel $Q(x, B) = \mathbb{P}\{\rho x + \xi_t \in B\}$ on $S = \mathbb{R}$ associated with the linear AR(1) model

$$X_{t+1} = \rho X_t + \xi_{t+1}, \qquad \{\xi_t\} \overset{\text{IID}}{\sim} N(0, 1). \tag{6}$$

This kernel fails to satisfy the MMC. On the other hand, it is order reversing. To see this, fix $(x, x') \in \mathbb{R}^2$, and take two independent Markov-Q processes

$$X_{t+1} = \rho X_t + \xi_{t+1} \text{ with } X_0 = x, \qquad X'_{t+1} = \rho X'_t + \xi'_{t+1} \text{ with } X'_0 = x',$$

[15] To be precise, let $\bar{x}$ and k be as in (5). Fix $x, x' \in S$ and let $\{X_t\}$ and $\{X_t'\}$ be independent, Markov-(Q, x) and Markov-(Q, x') respectively. By independence and $\{X_k \leq \bar{x} \leq X_k'\} \subset \{X_k \leq X_k'\}$, we have $\mathbb{P}\{X_k \leq \bar{x}\}\mathbb{P}\{\bar{x} \leq X_k'\} = \mathbb{P}\{X_k \leq \bar{x} \leq X_k'\} \leq \mathbb{P}\{X_k \leq X_k'\}$. But $\mathbb{P}\{\bar{x} \leq X_k'\} = Q^k(x, [a, \bar{x}])$ and $\mathbb{P}\{X_k \leq \bar{x}\} = Q^k(x, [\bar{x}, b])$ are strictly positive by (5) and remark 2.1. Hence Q is order reversing.

where $\{\xi_t\}$ and $\{\xi'_t\}$ are IID, standard normal, and independent of each other. We can see that $\mathbb{P}\{X_t \leq X'_t\} > 0$ is satisfied with $t = 1$, because

$$\mathbb{P}\{X_1 \leq X'_1\} = \mathbb{P}\{\rho x + \xi_1 \leq \rho x' + \xi'_1\} = \mathbb{P}\{\xi_1 - \xi'_1 \leq \rho(x' - x)\}.$$

Since $\xi_1 - \xi'_1$ is Gaussian, this probability is strictly positive.

3 Main Results

We can now state our main results, which concern stability and ergodicity of increasing, order reversing stochastic kernels.

3.1 Global Stability

Our first result extends Hopenhayn and Prescott's stability theorem to a broader class of models. It also characterizes the set of increasing order reversing kernels that are globally stable. The proof is in section 7.

Theorem 3.1. *Let Q be a stochastic kernel that is both increasing and order reversing. Then Q is globally stable if and only if*

1. *Q is bounded in probability, and*
2. *Q has either a deficient or an excessive distribution.*

Remark 3.1. In terms of sufficient conditions for global stability, the order reversing assumption cannot be omitted, even for *existence* of a stationary distribution. In particular, there exist increasing kernels that are bounded in probability and possess an excessive or deficient distribution, but have no stationary distribution.[16]

To see that the conditions of theorem 3.1 are weaker than those of the original Hopenhayn-Prescott stability theorem, suppose as they do that S is a compact metric space with least element a and greatest element b, and Q is an increasing kernel satisfying the MMC. The conditions of theorem 3.1 then hold. First, Q is increasing

[16]An example is the kernel Q associated with the deterministic process on $S = \mathbb{R}_+$ defined by $X_{t+1} = 1/2 + \sum_{n=0}^{\infty} \mathbb{1}\{n \leq X_t < n+1\}(n + (X_t - n)/2)$. It is easy to check that $X_{t+1} > X_t$ with probability one, and hence X_{t+1} and X_t can never have the same distribution. On the other hand, Q is increasing, bounded in probability (because each interval $[n, n+1)$ is absorbing) and has the deficient distribution δ_0 (cf., remark 2.2).

by assumption. Second, Q is order reversing, as shown in example 2.5. Third, Q is bounded in probability, since S is compact and hence $\{Q^t(x, \cdot)\}$ is always tight. Fourth, Q has a deficient distribution because S has a least element (see remark 2.2).

To see that the conditions of theorem 3.1 are strictly weaker than those of Hopenhayn and Prescott, consider the AR(1) model (6) with $\rho \in [0, 1)$. Here the Gaussian shocks force us to choose the state space $S = \mathbb{R}$, which is not compact, and the Hopenhayn-Prescott theorem in its original formulation cannot be applied. On the other hand, all the conditions of theorem 3.1 are satisfied.[17] (Of course the AR(1) model is a trivial example. Nontrivial applications are presented in section 5.)

Regarding the proof of theorem 3.1, boundedness in probability and existence of an excessive or deficient distribution generalize Hopenhayn and Prescott's assumption that S is compact and has a least and greatest element. As Hopenhayn and Prescott show, if S is compact and has a least and greatest element, then the Knaster-Tarski fixed point theorem implies that every increasing stochastic kernel has a stationary distribution. Adding the MMC then yields uniqueness and global stability. In our setting, the same fixed point argument cannot be applied. As remark 3.1 shows, our mixing condition plays an essential role in the proof of existence, and the proof is fundamentally different to the Knaster-Tarski fixed point argument.

We make two final remarks. First, one of the most attractive features of the MMC is that it is straightforward to check in applications when it holds. In section 4.3, we provide conditions for order reversing that are also straightforward to verify when they hold. Second, there is no continuity requirement in theorem 3.1. However, in many applications the kernel Q will have the Feller property (see example 2.2). If Q is Feller, then condition 2 can be omitted. Since this result is likely to be useful, we state it as a second theorem.

Theorem 3.2. *Let Q be increasing, order reversing, and Feller. Then Q is globally stable if and only if Q is bounded in probability.*

3.2 Ergodicity

As stated in the introduction, we show in this paper that the conditions of the Hopenhayn-Prescott stability theorem have additional implications. We now turn

[17] That the model is order reversing was shown in example 2.5. Monotonicity follows from example 2.2. Boundedness in probability is shown in example 4.1 below. For existence of a μ with $\mu \preceq \mu Q$, we can take $\mu = N(0, (1-\rho^2)^{-1})$.

to this point. To begin, let Q be a stochastic kernel on S with stationary distribution μ^*, and let h be a measurable function from S into $\mathbb{R}$. We say that the pair (Q, h) satisfies the strong law of large numbers if, for any $x \in S$ and any Markov-(Q, x) process $\{X_t\}$, we have

$$\frac{1}{n} \sum_{t=1}^{n} h(X_t) \to \int h \, d\mu^* \quad \text{as } n \to \infty \tag{7}$$

with probability one (i.e., $\mathbb{P}$-almost surely).[18]

Theorem 3.3. *Let Q be increasing, order reversing and bounded in probability. If Q has a stationary distribution and $h \in ibS$, then (Q, h) satisfies the strong law of large numbers.*

Remark 3.2. One of the most important implications of theorem 3.3 is that ergodicity holds under the conditions of the original Hopenhayn-Prescott stability theorem, since all of the conditions of theorem 3.3 are satisfied in that case.

Remark 3.3. In view of theorems 3.1 and 3.2, when Q is increasing, order reversing and bounded in probability, existence of a stationary distribution will be guaranteed whenever Q has an excessive or deficient distribution, or when Q is Feller. Theorem 3.3 then applies.

Remark 3.4. Since finite intersections of probability one sets have probability one, if $h_1, \ldots, h_k$ all satisfy (7), then $h := \alpha_1 h_1 + \cdots + \alpha_k h_k$ satisfies (7) for any scalars $\alpha_1, \ldots, \alpha_k$. As a result, the implications of theorem 3.3 extend to any h in the linear span of ibS. For example, they extend to any bounded measurable decreasing function, and to any function of bounded variation when S is an interval of $\mathbb{R}$. As we prove below, they also extend to any continuous bounded function under an additional assumption on the state space.

The most significant aspect of theorem 3.3 is that the probability one convergence is valid for any initial condition $x \in S$. This allows us to compute stationary probabilities and expectations by simulating time series from arbitrary initial conditions. Let us now consider the particular problem of computing the stationary distribution itself. To this end, note that theorem 3.3 concerns the behavior of the *empirical*

[18]For a general discussion of the strong law of large numbers in the classical Markov setting see [44, chapter 17] or [25, chapter 2]. For a law of large numbers under the assumptions of compact state space and Feller stochastic kernel see [62, chapter 14].

distribution μ_n^X corresponding to the process $\{X_t\}$, which is defined by

$$\int h \, d\mu_n^X := \frac{1}{n} \sum_{t=1}^{n} h(X_t) \quad \text{for measurable } h \colon S \to \mathbb{R}.$$

Ideally we would like the conditions of theorem 3.3 to imply that μ_n^X converges to μ^* with probability one, which is to say that the event $\int h \, d\mu_n^X \to \int h \, d\mu^*$ for all $h \in cbS$ has probability one. This is not implied by theorem 3.3. In fact we cannot even say that the event $\int h \, d\mu_n^X \to \int h \, d\mu^*$ for all $h \in ibS$ has probability one, because, for each choice of $h \in ibS$, the probability one set on which $\int h \, d\mu_n^X \to \int h \, d\mu^*$ holds depends on h, and the function class ibS is uncountable.[19] Nevertheless, with an additional restriction on S, we are able to obtain the desired probability one convergence of μ_n^X to μ^*. The new restriction strengthens our separability assumption for S, and is satisfied for many common state spaces (e.g., when S is a cone in $\mathbb{R}^m$ with the usual partial order):

Assumption 3.1. There exists a countable subset A of S such that, given any $x \in S$ and neighborhood U of x, there are $a, a' \in A$ such that $a, a' \in U$ and $a \leq x \leq a'$.

Theorem 3.4. *If the conditions of theorem 3.3 hold and assumption 3.1 is satisfied, then, for any $x \in S$ and any Markov-(Q, x) process $\{X_t\}$, the empirical distribution μ_n^X satisfies $\mu_n^X \to \mu^*$ as $n \to \infty$ with probability one.*[20]

4 Verifying the Conditions

Theorem 3.1 requires that Q is increasing, order reversing, bounded in probability, and possesses an excessive or deficient distribution. A sufficient condition for Q to be increasing was given in example 2.2. In this section, we present a number of sufficient conditions for the remaining properties. Throughout the following discussion, we use the simple AR(1) model for illustrative purposes. Nontrivial applications are deferred to section 5.

[19] In fact *ibS* is not only uncountable, but typically non-separable. This is the essence of the problem.

[20] That is to say, $\mathbb{P}\{\omega \in \Omega : \langle \mu_n^{X(\omega)}, h \rangle \to \langle \mu^*, h \rangle, \ \forall \, h \in cbS\} = 1$.

4.1 Boundedness in Probability

Boundedness in probability is a standard condition in the Markov process literature. In this section, we briefly review some standard techniques for checking boundedness in probability, and introduce a new one based on order-theoretic ideas.

Let Q be a stochastic kernel on $S = \mathbb{R}$, and let $\{X_t\}$ be Markov-(Q, x). Then Q is bounded in probability when $\sup_t \mathbb{E}\,|X_t| < \infty$ for any initial x. The same statement is valid if we replace $|X_t|$ with X_t^2. Intuitively, boundedness of these moments means that the process does not diverge.

To go beyond the case of $S = \mathbb{R}$, recall that $V \colon S \to \mathbb{R}_+$ is called *coercive* if the sublevel sets $L_a := \{x \in S : V(x) \leq a\}$ are precompact for all $a > 0$.[21] It is known that Q is bounded in probability if

$$\forall\, x \in S, \; \exists \text{ a coercive function } V \text{ with } \sup_t \int V(y) Q^t(x, dy) < \infty. \tag{8}$$

This condition is in fact equivalent to boundedness in probability, and *necessary* for global stability [44, lemma D.5.3]. Thus every proof of global stability requires verification of (8), either directly or indirectly. One way to verify (8) directly is via a "drift" condition. For example, let Q_F be the kernel (2). Then (8) will be satisfied if there exist positive constants α and β with $\alpha < 1$ and

$$\mathbb{E}\, V[F(x, \xi_t)] \leq \alpha V(x) + \beta, \qquad \forall\, x \in S. \tag{9}$$

Example 4.1. Consider the AR(1) process (6) with $S = \mathbb{R}$. Here (9) is satisfied for $V(x) := |x|$ whenever $|\rho| < 1$. Indeed, by the triangle inequality, $\mathbb{E}\,|\rho x + \xi_t| \leq |\rho| \cdot |x| + \mathbb{E}\,|\xi_t|$. This corresponds to (9) with $\alpha := |\rho|$ and $\beta := \mathbb{E}\,|\xi_t|$.

More examples of how to verify (8) are given in section 5.[22] Before finishing this section, we introduce a new result that can be used to check boundedness in probability, and also relates to our techniques for checking existence of deficient and excessive distributions discussed in section 4.2 below.

[21] A subset A of S is called *precompact* if every sequence in A has a subsequence converging to an element of S. For $S = \mathbb{R}^n$, a natural example of a coercive function is $V(x) = \|x\|$. V is coercive on S because L_a is the compact set $\bar{B}(0, a) = \{x \in \mathbb{R}^n : x \leq a\}$. However, if $S = \mathbb{R}^n_{++}$, then $V(x) = \|x\|$ is not coercive, because $L_a = \mathbb{R}^n_{++} \cap \bar{B}(0, a)$, which is not precompact. On the other hand, $V(x) = 1/\|x\| + \|x\|$ is coercive on $\mathbb{R}^n_{++}$. In essence, V is coercive on state space S if $V(x_n) \to \infty$ whenever x_n "diverges" towards the "edges" of S. (Here and elsewhere, $\|\cdot\|$ refers to the euclidean norm, but all of the statements in this footnote remain valid if we use any other norm on $\mathbb{R}^n$.)

[22] Further examples can be found in [61, 48, 35, 37, 44].

Proposition 4.1. *Let Q_ℓ, Q, Q_u be stochastic kernels on S. If $Q_\ell \preceq Q \preceq Q_u$ and both Q_ℓ and Q_u are bounded in probability, then Q is bounded in probability.*

4.2 Existence of Excessive and Deficient Distributions

Condition 2 of theorem 3.1 requires existence of either an excessive or a deficient distribution. If S has a least element or a greatest element then the condition always holds (see remark 2.2). However, there are many settings where S has neither ($S = \mathbb{R}^n$ and $S = \mathbb{R}^n_{++}$ are obvious examples), and the existence is harder to verify. In this case, one can work more carefully with the definition of the model to construct excessive and deficient distributions. One example is Zhang [64], who constructs such distributions for the stochastic optimal growth model. However, it is useful to have a more systematic method that is relatively straightforward to check in different applications. To this end we provide the following result:

Proposition 4.2. *Let Q be a stochastic kernel on S. If there exists another kernel Q_u such that Q_u is Feller, bounded in probability and $Q \preceq Q_u$, then Q has an excessive distribution. Likewise, if Q_ℓ is Feller, bounded in probability and $Q_\ell \preceq Q$, then Q has a deficient distribution.*

Examples of how to use this result are provided in the applications. In addition, we note that propositions 4.1 and 4.2 can be combined with theorem 3.1 to obtain the following stability result:

Theorem 4.1. *Suppose that Q is increasing and order reversing. If there exist kernels $Q_\ell \preceq Q \preceq Q_u$ such that Q_ℓ and Q_u are bounded in probability and at least one of them is Feller, then Q is globally stable.*

4.3 Order Reversing

In this section we give sufficient conditions for order reversing. To state them, we introduce two new definitions: We call kernel Q on S *upward reaching* if, given any x and c in S, there exists a $t \in \mathbb{N}$ such that $Q^t(x, \{y \in S : c \leq y\}) > 0$. We call Q *downward reaching* if, given any x and c in S, there exists a $t \in \mathbb{N}$ such that $Q^t(x, \{y \in S : y \leq c\}) > 0$.

Example 4.2. The AR(1) process in (6) is both upward and downward reaching. For example, fix x, c in $S = \mathbb{R}$, and take $t = 1$. We have

$$Q(x, \{y \in S : y \leq c\}) = \mathbb{P}\{\rho x + \xi_1 \leq c\} = \mathbb{P}\{\xi_1 \leq c - \rho x\},$$

which is positive because $\xi_t \sim N(0, 1)$. Hence Q is downward reaching.

We can now present the main result of this section.

Proposition 4.3. *Suppose that Q is bounded in probability. If Q is either upward or downward reaching, then Q is order reversing.*

Corollary 4.1. *The statements in theorem 3.1 and theorem 3.2 remain valid if order reversing is replaced by either upward or downward reaching.*

Using proposition 4.3, we can also provide more specialized results for the model in example 2.1. To simplify the exposition, we assume without loss of generality that Z is the support of ϕ.[23] Suppose that, in addition to prior assumptions, the state space S is a Borel subset of $\mathbb{R}^n$. Given vectors x and y in S, we write $x < y$ if $x_i < y_i$ for all i. Finally, as an additional point of notation, observe that each finite path of shock realizations $\{z_i\}_{i=1}^t \subset Z$ and initial condition $X_0 = x \in S$ determines a path $\{x_i\}_{i=0}^t$ for the state variable up until time t via (1). Let $F^t(x, z_1, \ldots, z_t)$ denote the value of x_t determined in this way.[24]

Proposition 4.4. *Suppose that $x \mapsto F(x, z)$ is increasing for each $z \in Z$, F is continuous on $S \times Z$, and Q_F is bounded in probability. Then Q_F is globally stable if any one of the following three conditions holds:*

1. $\forall\, x, c \in S, \;\; \exists\, \{z_i\}_{i=1}^k \subset Z$ *such that* $F^k(x, z_1, \ldots, z_k) < c$.
2. $\forall\, x, c \in S, \;\; \exists\, \{z_i\}_{i=1}^k \subset Z$ *such that* $F^k(x, z_1, \ldots, z_k) > c$.
3. $\forall\, x, x' \in S, \;\; \exists\, \{z_i\}_{i=1}^k$ *and* $\{z_i'\}_{i=1}^k$ *with* $F^k(x, z_1, \ldots, z_k) < F^k(x', z_1', \ldots, z_k')$.

Example 4.3. Consider the AR(1) model in (6), where $F(x, z) = \rho x + z$. All of conditions 1–3 in proposition 4.4 hold. For example, taking condition 1, fix $x, c \in \mathbb{R}$. We need to choose a shock sequence that drives the process below c when it starts at x. This can be done in one step, by choosing z_1 such that $\rho x + z_1 < c$.

[23] That is, $\phi(Z) = 1$, and $\phi(G) > 0$ whenever $G \subset Z$ is nonempty and open. Z can always be re-defined so that this assumption is valid.

[24] Formally, $F^1 := F$ and $F^{t+1}(x, z_1, \ldots, z_{t+1}) := F(F^t(x, z_1, \ldots, z_t), z_{t+1})$ for all $t \in \mathbb{N}$.

5 Applications

We now turn to more substantial applications of the results described above.

5.1 Optimal Growth

The infinite-horizon stochastic optimal growth model forms the foundations of many dynamic models, spanning fields such as growth, development, international trade, monetary policy, fiscal policy, commodity pricing and environmental economics. The existence of globally stable, non-trivial stochastic equilibria is of fundamental importance when comparing predictions with data. While global stability of the elementary one-sector concave model with IID shocks and inelastic labor is well-known, conditions for stability of many of the variations used in applications remain to be investigated.

Consider an elementary model, with consumption chosen in order to maximize $\mathbb{E}\sum_{t=0}^{\infty}\beta^t u(c_t)$ subject to $0 \leq k_{t+1} + c_t \leq \xi_t f(k_t)$. All variables are nonnegative and $\{\xi_t\} \overset{\text{IID}}{\sim} \phi$. For now, we assume that u is bounded with $u' > 0$, $u'' < 0$, and $u'(0) = \infty$; while $f(0) = 0$, $f' > 0$, $f'' < 0$, $f'(0) = \infty$ and $f'(\infty) = 0$. To study the dynamics of the optimal process, we take $y_t = \xi_t f(k_t)$ as the state variable, and consider the income process $y_{t+1} = \xi_t f(y_t - \sigma(y_t))$, where $\sigma(\cdot)$ is the optimal consumption policy. Let Q be the corresponding stochastic kernel. For the state space we take $S = \mathbb{R}_{++}$. Zero is deliberately excluded from S so that any stationary distribution on S is automatically non-trivial.

Brock and Mirman [10] were the first to prove global stability of Q, in the case where ξ_t has support $[a, b]$ for some $0 < a < b$.[25] The case where ξ_t has unbounded support was treated in [61, 48, 35, 64]. For example, Stachurski [61] replaced the assumption of bounded support with the small tail conditions $\mathbb{E}\,\xi_t < \infty$ and $\mathbb{E}\,(1/\xi_t) < \infty$, in addition to the assumption that ξ_t has a density that is continuous and everywhere positive. Zhang [64] showed that global stability continues to hold when ξ_t has no density, provided that $\mathbb{P}\{\xi_t \leq z\} > 0$ for all $z \in S$, or $\mathbb{P}\{\xi_t \geq z\} > 0$ for all $z \in S$.

All of these global stability results, proved using a range of specialized arguments, can be obtained as special cases of either one of theorem 3.1 or theorem 3.2. Indeed, it is well-known that under the stated assumptions Q is increasing, Feller

[25] See [28, 51, 35, 14] for additional discussion of the case where ξ_t has bounded support.

and bounded in probability (cf., e.g., [62, p. 393] and [48, proposition 4.4]). Hence it suffices to show that Q is either order reversing, upward reaching or downward reaching (theorem 3.2 and corollary 4.1). For the bounded shock case, Hopenhayn and Prescott [28] proved that the MMC holds, which implies order reversing (see example 2.5). For the unbounded case, consider the conditions of Zhang [64] mentioned above. Suppose in particular that $\mathbb{P}\{\xi_t \leq z\} > 0$ for all $z \in S$. In this case, for all fixed y_0 and $\bar{y}$ in S, we have

$$\mathbb{P}\{y_1 \leq \bar{y}\} = \mathbb{P}\{\xi_1 f(y_0 - \sigma(y_0)) \leq \bar{y}\} = \mathbb{P}\{\xi_1 \leq \bar{y}/f(y_0 - \sigma(y_0))\} > 0. \tag{10}$$

Hence Q is downward reaching, and therefore globally stable.[26]

These initial stability results can be extended in many ways. For example, take again the conditions of Zhang [64], but without the assumption that f is concave. Models with nonconcave reproduction arise in areas such as renewable resource exploitation, when f is biologically determined.[27] Without concavity of f, optimal consumption is not necessarily continuous, and Q is no longer Feller. Moreover, without additional assumptions, the MMC does not apply, Q is not irreducible, the splitting condition fails, the model is not an expected contraction and the standard Harris recurrence conditions are not satisfied.[28] Indeed, to the best of our knowledge, global stability of Q—or even existence of a stationary distribution—cannot be established using any result in the existing literature.[29] On the other hand, theorem 3.1 can easily be applied. Q is still increasing and bounded in probability [48]. Existence of an excessive distribution is not difficult to establish.[30] Moreover, the downward reaching proof in (10) goes through unchanged. Hence theorem 3.1 applies, and Q is globally stable.

[26]Since $S = \mathbb{R}_{++}$ and optimal consumption is interior, we have $f(y_0 - \sigma(y_0)) > 0$.

[27]For motivation and further discussion, see, for example, [20, 41, 46].

[28]For a discussion of irreducibility and Harris recurrence, see, e.g., [44]. On the splitting condition, see, e.g., [7, 8]. For expected contractions, see, e.g., [60, p. 1952].

[29]The closest result is [48, theorem 3.1]. That result requires that ξ_t has a density that is continuous and strictly positive everywhere on S.

[30]To do so we can use proposition 4.2. Since $f'(\infty) = 0$, we can choose positive constants α, β with $\alpha \mathbb{E}\,\xi_t < 1$ and $f(x) \leq \alpha x + \beta$ [48, proposition 4.3]. Now take $G(x,z) := z(\alpha x + \beta)$, so that $F(x,z) := z f(x - \sigma(x)) \leq z f(x) \leq G(x,z)$. Letting Q_F and Q_G be the corresponding kernels, the last inequality implies $Q_F \preceq Q_G$. In view of proposition 4.2, it remains only to show that Q_G is Feller and bounded in probability. Since $G(\cdot, z)$ is continuous, Q_G is Feller. Using $\alpha \mathbb{E}\,\xi_t < 1$, condition (9) can be established for $V(x) = x + 1/x$, which is coercive on $S = \mathbb{R}_{++}$. Boundedness in probability then follows.

Remark 5.1. It is interesting to note that in order to prove stability we used order reversing, and to prove order reversing we relied on nonzero probability of arbitrarily bad productivity shocks. These shocks are stabilizing rather than destabilizing because the Inada conditions prevent divergence, and the large shocks promote mixing.

Since the conditions of theorems 3.3 and 3.4 apply, we can compute stationary moments or the stationary distribution itself at each set of parameters, by simulating from an arbitrary initial condition. Figure 1 shows a collection of stationary distributions for $\log y_t$, each one corresponding to a different value of the discount factor β. Here the production function is nonconcave, as might be the case in a model of biological resource exploitation.[31] For this model, a sudden shift in the optimal harvest policy occurs around $\beta = 0.965$. As a result, a very small difference in the patience of the agent can lead to a large difference in the steady state population of the stock.

5.2 An OLG Model of Wealth Distribution

Next we consider an OLG model of wealth distribution. The model can be viewed as a stochastic version of the small open economy of Matsuyama [43], but we introduce persistence in inequality by assuming that an old agent provides financial support to her child. This is a common assumption in the literature on wealth distribution (e.g., [4, 5, 13, 17, 39, 40, 52, 54, 56]).

Agents live for two periods, consuming only when old. Households consist of one old agent and one child. There is a unit mass of such households indexed by $i \in [0, 1]$. In each period t, the old agent of household i provides financial support b_t^i to her child. The child has the option to become an entrepreneur, investing one unit of the consumption good in a "project," and receiving stochastic output $\theta + \eta_{t+1}^i$ in period $t + 1$. Let $k_{t+1}^i \in \{0, 1\}$ be young agent i's investment in the project. If the remainder $b_t^i - k_{t+1}^i$ is positive, then she invests this quantity at the world risk-free rate R. If it is negative then she borrows $k_{t+1}^i - b_t^i$ at the same risk-free rate.

[31] The utility function is $u(x) = 1 - \exp(-\theta x^\gamma)$ and production is $f(x) = x^\alpha \ell(x)$, where ℓ is the logistic function $\ell(x) = a + (b - a)/(1 + \exp(-c(x - d)))$. The parameters are $a = 1$, $b = 2$, $c = 20$, $d = 1$, $\theta = 0.5$, $\gamma = 0.9$ and $\alpha = 0.5$. The discount factor β ranges from 0.945 to 0.99. The shock is lognormal $(-0.1, 0.2)$. The optimal policy is calculated by fitted value function iteration. To compute each stationary distribution, we simulated a time series of length 10^6 from the process $y_{t+1} = \xi_t f(y_t - \sigma(y_t))$, where $\sigma(\cdot)$ is the computed policy.

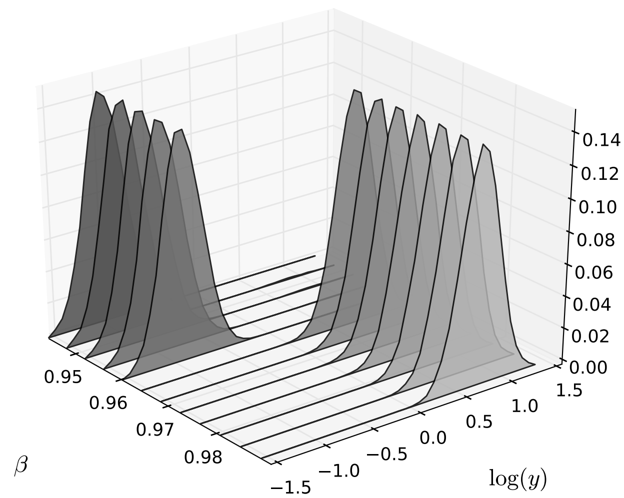

Figure 1: Stationary distributions

Independent of her investment choice, she receives an endowment of e^i_{t+1} units of the consumption good when old. Suppressing the i superscript to simplify notation, her wealth at the beginning of period $t+1$ is therefore

$$w_{t+1} = (\theta + \eta_{t+1})k_{t+1} - R(k_{t+1} - b_t) + e_{t+1}. \tag{11}$$

We assume that

$$e_{t+1} = \rho e_t + \epsilon_{t+1}, \quad 0 < \rho < 1. \tag{12}$$

The idiosyncratic shocks $\{\eta_t\}$ and $\{\epsilon_t\}$ are taken to be IID and nonnegative, and ϵ_t satisfies $\mathbb{P}\{\epsilon_t > z\} > 0$ for any $z \geq 0$. (For example, ϵ_t might be lognormal.) We also assume that $R < \theta$, which implies that becoming an entrepreneur is always profitable, even *ex-post*, and every agent would choose to do so absent additional constraint. Due to a credit market imperfection, however, each agent may borrow only up to a fraction $\lambda \in (0,1)$ of $\theta + \rho e_t$, the minimum possible value of her old-age income. That is,

$$R(k_{t+1} - b_t) \leq \lambda(\theta + \rho e_t). \tag{13}$$

As becoming an entrepreneur is always profitable, young agents do so whenever feasible, implying

$$k_{t+1} = \kappa(b_t, e_t) := \mathbb{1}\{R(1 - b_t) \leq \lambda(\theta + \rho e_t)\}. \tag{14}$$

(Here $\mathbb{1}\{\cdot\}$ is an indicator function.) Let c_{t+1} denote consumption at $t + 1$. It is common in the literature on wealth distribution to assume that each agent derives utility from her own consumption and financial support to her child. Following this approach, we assume that young agents maximize $\mathbb{E}_t[c_{t+1}^{1-\gamma} b_{t+1}^{\gamma}]$ subject to (11), (13), and the budget constraint $c_{t+1} + b_{t+1} = w_{t+1}$. Regarding the parameter γ we assume that $\gamma R < 1$. Maximization of $c_{t+1}^{1-\gamma} b_{t+1}^{\gamma}$ subject to the budget constraint implies that $b_{t+1} = \gamma w_{t+1}$. Combining this equality, (11) and (12), we obtain

$$b_{t+1} = \gamma[(\theta + \eta_{t+1} - R)\kappa(b_t, e_t) + Rb_t + \rho e_t + \epsilon_{t+1}]. \tag{15}$$

Together, (12) and (15) define a Markov process with state vector $X_t := (b_t, e_t)$ taking values in state space $S := [0, \infty) \times [0, \infty)$. Let Q denote the corresponding stochastic kernel.[32]

Recalling that $R < \theta$, $\rho \in (0, 1)$ and $\eta_{t+1} \geq 0$, and observing that $\kappa(b_t, e_t)$ is increasing in (b_t, e_t), we can see from (12) and (15) that (b_{t+1}, e_{t+1}) is increasing in (b_t, e_t) when the values of the shocks are held fixed. Hence Q is increasing (cf., example 2.2). On the other hand, (15) is discontinuous in (b_t, e_t), so Q is not Feller.

As far as we are aware, no existing Markov process theory can be used to show that Q is globally stable unless additional conditions are imposed. In contrast, global stability can be obtained in a straightforward way from theorem 3.1. To begin, let $m_\eta := \mathbb{E}\,\eta_t$ and $m_\epsilon := \mathbb{E}\,\epsilon_t$. To see that Q is bounded in probability, we can take expectations of (12) and iterate backwards to obtain

$$\mathbb{E}\,e_t \leq m_\epsilon/(1 - \rho) + \rho^t e_0 \leq m_\epsilon/(1 - \rho) + e_0 =: \bar{e} \tag{16}$$

for all t. In addition, it follows from (15) and (16) that

$$\mathbb{E}\,b_{t+1} \leq \gamma[\theta + m_\eta - R + R\mathbb{E}\,b_t + \bar{e}].$$

Using $\gamma R < 1$ and iterating backwards, we obtain the bound

$$\mathbb{E}\,b_t \leq \gamma[\theta + m_\eta - R + \bar{e}]/(1 - \gamma R) + b_0 \tag{17}$$

[32]We do not exclude $(0, 0)$ from the state space since it is not an absorbing state.

for all t. Together, (16) and (17) imply that Q is bounded in probability.[33] Since $\mathbb{P}\{\epsilon_t > z\} > 0$ for any $z \geq 0$, and since both b_t and e_t can be made arbitrarily large by choosing ϵ_t sufficiently large (see (12) and (15)), it follows that Q is upward reaching, and thus order reversing by proposition 4.3. In view of these results and theorem 3.1, Q will be globally stable whenever it has a deficient or excessive distribution. Since $(0,0)$ is a least element for S, remark 2.2 implies that Q has a deficient distribution, and we conclude that Q is globally stable.

Since the conditions of theorems 3.3 and 3.4 apply, we can compute the stationary distribution at each set of parameters by simulating from an arbitrary initial condition. Figure 2 shows smoothed histograms representing the marginal stationary distribution of wealth at two different values of λ, while figure 3 is a two-dimensional projection of the bivariate stationary distribution, represented as a histogram.[34] The shift in figure 2 shows how the distribution of wealth in the stationary equilibrium can be extremely sensitive to the value of the borrowing constraint parameter λ.

5.3 Nonlinear Autoregression

Consider a general additive shock nonlinear autoregressive model of the form

$$X_{t+1} = f(X_t) + \xi_{t+1}, \quad \{\xi_t\} \overset{\text{IID}}{\sim} \phi, \quad \mathbb{E}\,\|\xi_t\| < \infty, \tag{18}$$

where $S = \mathbb{R}^n$, and $f\colon \mathbb{R}^n \to \mathbb{R}^n$. Let Q be the stochastic kernel, defined by $Q(x, B) = \mathbb{P}\{f(x) + \xi_t \in B\}$. We assume that (i) f is increasing, (ii) $\mathbb{P}\{\xi_t \leq z\}$ is non-zero for all $z \in \mathbb{R}^n$, and (iii) there exists an $\alpha \in [0, 1)$ and $L \geq 0$ such that $\|f(x)\| \leq \alpha\|x\| + L$ for all $x \in \mathbb{R}^n$. The last assumption is a growth condition on f. Global stability cannot hold without some restriction along these lines. The second assumption is used below to prove that Q is downward reaching. It can be replaced by: $\mathbb{P}\{\xi_t \geq z\}$ is non-zero for all $z \in \mathbb{R}^n$. If this condition holds then Q will be upward reaching.

We know of no previous results that can be used to prove global stability here without additional assumptions. On the other hand, a straightforward proof of

[33] The function $V(b, e) = b + e$ is coercive on S, and equations (16) and (17) imply that $\sup_t \mathbb{E}\,[V(b_t, e_t)] \leq \sup_t \mathbb{E}\,[b_t] + \sup_t \mathbb{E}\,[e_t] < \infty$, which gives (8).

[34] In figure 2, the values of λ are 0.57 and 0.58. In figure 3, λ is set to 0.58. The other parameters are $\gamma = 0.2$, $R = 1.05$, $\theta = 1.1$ and $\rho = 0.9$. The shock ϵ is lognormal with parameters $\mu = -3$ and $\sigma = 0.1$. The shock η is beta with shape parameters 3,10. The simulated time series are of length 10^6 for each stationary distribution. The distributions in figure 2 are calculated with Gaussian nonparametric kernel density estimates.

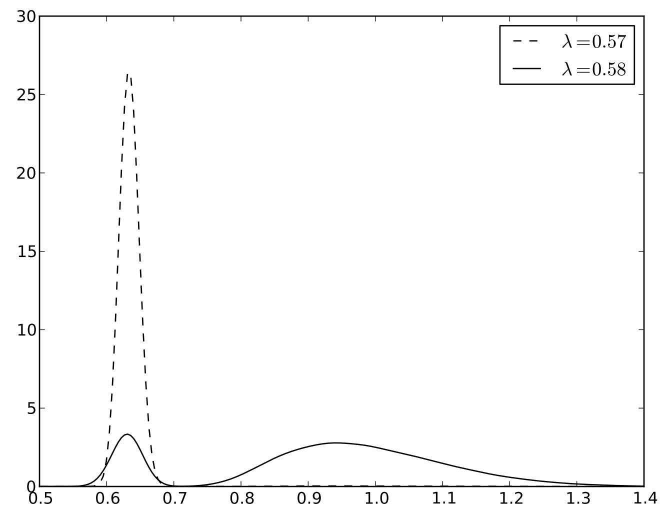

Figure 2: Stationary distribution of wealth

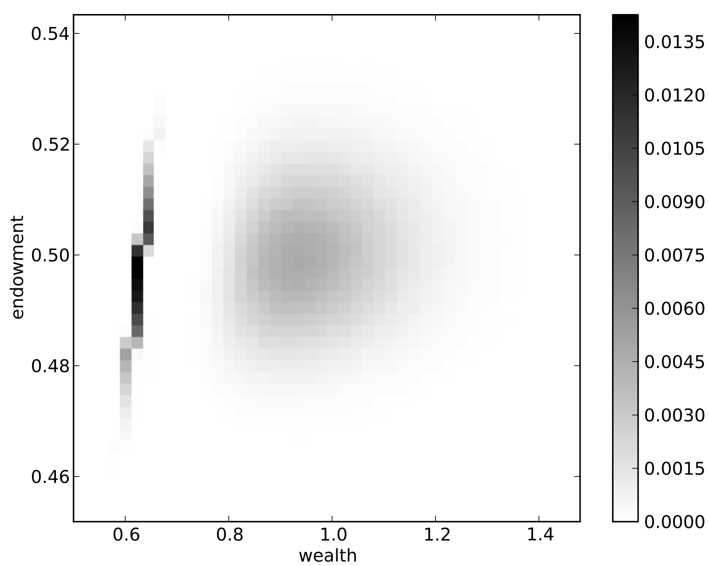

Figure 3: Bivariate stationary distribution with $\lambda = 0.58$

global stability can be constructed via theorem 3.1. To begin, note that $F(x,z) := f(x) + z$ is increasing in x for each z because f is increasing, and hence, by example 2.2, Q is increasing. Second, Q is bounded in probability, as can be shown by taking $V(x) := \|x\|$ in (9).[35] Third, Q is downward reaching by an argument similar to that of example 4.2: If we fix $x, c \in S = \mathbb{R}^n$ and let E be the event $\{\xi_1 \leq c - f(x)\}$, then, by assumption, $\mathbb{P}(E) > 0$. Moreover, if E occurs, then $f(x) + \xi_1 \leq c$. Hence Q is downward reaching. In view of proposition 4.3, Q is also order reversing.

To complete the proof of global stability via theorem 3.1, it only remains to show existence of a deficient or excessive distribution. For this purpose, we use proposition 4.2. To apply the proposition, we aim to find a kernel Q_u such that Q_u is Feller, bounded in probability, and satisfies $Q \preceq Q_u$. To construct Q_u, consider the process

$$X_{t+1} = g(X_t) + \xi_{t+1}, \qquad g(x) := \alpha\|x\|\mathbf{1} + L\mathbf{1}. \tag{19}$$

Here $\mathbf{1}$ is the unit vector in $\mathbb{R}^n$, and α and L are as given in the assumptions following (18). We let Q_u be the stochastic kernel corresponding to (19). Q_u is easily shown to be Feller and bounded in probability.[36] Moreover, $f(x) + z \leq g(x) + z$ for all x and z, because if $f_i(x)$ is the i-th component of $f(x)$, then $f_i(x) \leq |f_i(x)| \leq \|f(x)\| \leq \alpha\|x\| + L$, and hence $f(x) \leq \alpha\|x\|\mathbf{1} + L\mathbf{1} = g(x)$. It now follows from example 2.3 that $Q \preceq Q_u$. We conclude that the conditions of proposition 4.2 are satisfied, and the proof of global stability is done.

6 Conclusion

The Hopenhayn-Prescott stability theorem has become an important tool for assessing the dynamics of stochastic economic models. This paper strengthens and extends their theorem. We strengthen their results by establishing ergodicity (theorems 3.3 and 3.4). In addition, we extend the theorem to cover a significantly broader class of models (theorem 3.1). In particular, the version of the Hopenhayn-Prescott theorem presented in this paper opens up to study the dynamics of models with large shocks and high degrees of mixing. This makes it possible to develop new perspectives on the stability problem, and extend Hopenhayn and Prescott's stability results to models where relatively little other structure beyond mixing is imposed. We provide applications in section 5 where global stability can be established

[35] By the triangle inequality, $\mathbb{E}\,\|f(x) + \xi_1\| \leq \|f(x)\| + \mathbb{E}\,\|\xi_1\| \leq \alpha\|x\| + L + \mathbb{E}\,\|\xi_1\|$.

[36] Since g is continuous, Q is Feller. In addition, (9) is valid for the coercive function $V(x) := \|x\|_\infty := \max_{i=1}^n |x_i|$, and hence boundedness in probability also holds.

even though neither the original Hopenhayn-Prescott theorem nor any other existing Markov process theory yields stability (or even existence of a stationary distribution). Our applications cover an infinite horizon model, an overlapping generations model and a nonlinear autoregression.

Other significant new results contained in the paper are propositions 4.1–4.4 and theorem 4.1. These results provide additional stability conditions and aid in verification of order reversing and other conditions of our main stability theorems. Their usefulness is illustrated in the applications discussed in section 5.

7 Proofs

Before proving theorem 3.1, we need some additional results and notation. To begin, let Φ be any stochastic kernel on polish space D, let $x \in D$ and let D-valued stochastic process $\{X_t\}$ be Markov-(Φ, x). The joint distribution of $\{X_t\}$ over the sequence space D^∞ will be denoted by $\mathbf{P}_x^\Phi$. For example, $\mathbf{P}_x^\Phi\{X_t \in B\} = \Phi^t(x, B)$ for any $B \subset D$, and $\mathbf{P}_x^\Phi \cup_{t=0}^\infty \{X_t \in B\}$ is the probability that the process ever enters B. The symbol $\mathbf{E}_x^\Phi$ represents the expectations operator corresponding to $\mathbf{P}_x^\Phi$.

For given kernel Φ, we say that Borel set $B \subset D$ is

- *strongly accessible* if $\mathbf{P}_x^\Phi \cup_{t=0}^\infty \{X_t \in B\} = 1$ for all $x \in D$, and
- *C-accessible* if, for all compact $K \subset D$, there exists an $n \in \mathbb{N}$ with $\inf_{x \in K} \Phi^n(x, B) > 0$.

The following lemma is fundamental to our results, although the proofs are delayed to maintain continuity.

Lemma 7.1. *Let B be a Borel subset of D. If Φ is bounded in probability and B is C-accessible, then B is strongly accessible.*

Let Q be a given kernel on S, and let $Q \times Q$ be the product kernel (4). For given pair $(x, x') \in S \times S$, let $\{X_t\}$ and $\{X'_t\}$ be Markov-(Q, x) and Markov-(Q, x') respectively, and also independent of each other. As discussed in section 2.1, the bivariate process $\{(X_t, X'_t)\}$, which takes values in $S \times S$, is Markov-$(Q \times Q, (x, x'))$. Its joint distribution over the sequence space $(S \times S)^\infty$ is denoted by $\mathbf{P}_{x,x'}^{Q \times Q}$. In this notation, Q is order reversing if

$$\forall\, x, x' \in S \text{ with } x \geq x', \;\; \exists\, k \geq 0 \text{ such that } \mathbf{P}_{x,x'}^{Q \times Q}\{X_k \leq X'_k\} > 0.$$

Following Kamihigashi and Stachurski [36], Q is called *order mixing* if $\mathbf{P}_{x,x'}^{Q \times Q} \cup_{t=0}^\infty \{X_t \leq X'_t\} = 1$ for all $x, x' \in S$. Put differently, Q is order mixing if $\mathbb{G} := \{(y, y') \in S \times S : y \leq y\}$ is strongly accessible for the product kernel $Q \times Q$.

Lemma 7.2. *If Q is bounded in probability, then so is the product kernel $Q \times Q$.*

Lemma 7.3. *If Q is increasing and bounded in probability, then $\{\mu Q^t\}$ is tight for all $\mu \in \mathscr{P}_S$.*

Lemma 7.4. *If Q is increasing and order reversing, then $\mathbb{G}$ is C-accessible for $Q \times Q$.*

Proofs are given at the end of this section.

Let us now turn to the proof of theorem 3.1. The proof proceeds as follows: First we show that under the conditions of the theorem, Q is order mixing. Using order mixing, we then go on to prove existence of a stationary distribution, and global stability.

Lemma 7.5. *If Q is increasing, bounded in probability and order reversing, then Q is order mixing.*

Proof. To show that Q is order mixing we need to prove that $\mathbb{G}$ is strongly accessible for $Q \times Q$ under the conditions of theorem 3.1. Since Q is bounded in probability, $Q \times Q$ is also bounded in probability (lemma 7.2), and hence, by lemma 7.1, it suffices to show that $\mathbb{G}$ is C-accessible for $Q \times Q$. This follows from lemma 7.4. $\square$

We now prove global stability, making use of order mixing. In the sequel, we define *icbS* to be the bounded, increasing and continuous functions from S to $\mathbb{R}$ (i.e., $icbS = ibS \cap cbS$). We will make use of the following results, which are proved at the end of this section.

Lemma 7.6. *Let $\mu, \mu', \mu_n \in \mathscr{P}_S$.*

1. *$\mu \preceq \mu'$ iff $\langle \mu, h \rangle \leq \langle \mu', h \rangle$ for all $h \in icbS$,*
2. *$\mu = \mu'$ iff $\langle \mu, h \rangle = \langle \mu', h \rangle$ for all $h \in icbS$, and*
3. *$\mu_n \to \mu$ iff $\{\mu_n\}$ is tight and $\langle \mu_n, h \rangle \to \langle \mu, h \rangle$ for all $h \in icbS$.*

Proof of theorem 3.1. We begin by showing that if Q is globally stable, then conditions 1–2 of the theorem hold. Regarding condition 1, fix $x \in S$. Global stability implies that $\{\mu Q^t\}$ is convergent for each $\mu \in \mathscr{P}_S$, and hence $\{Q^t(x, \cdot)\} = \{\delta_x Q^t\}$ is convergent. Since convergent sequences are tight [24, proposition 9.3.4] and $x \in S$ was arbitrary, we conclude that Q is bounded in probability, and condition 1 is satisfied. Condition 2 is trivial, because global stability implies existence of a stationary distribution, and every stationary distribution is both deficient and excessive.

Next we show that if Q is increasing, order reversing and conditions 1–2 of theorem 3.1 hold, then Q has at least one stationary distribution. By lemma 7.5, Q is order mixing, and hence, by Kamihigashi and Stachurski [36, theorem 3.1], for any ν and ν' in $\mathscr{P}_S$ we have

$$\lim_{t \to \infty} |\langle \nu Q^t, h \rangle - \langle \nu' Q^t, h \rangle| = 0, \qquad \forall\, h \in ibS. \tag{20}$$

Now let $\{\mu Q^t\}$ be a tight and $\preceq$-monotone sequence, existence of which is guaranteed by conditions 1–2 of theorem 3.1 and lemma 7.3. We suppose without loss of generality that

$\{\mu Q^t\}$ is $\preceq$-increasing, since the other case changes nothing in what follows except the direction of the inequalities.

By Prohorov's theorem [24, theorem 11.5.4], tightness implies existence of a subsequence of $\{\mu Q^t\}$ converging to some $\psi^* \in \mathscr{P}_S$. Since $\{\mu Q^t\}$ is $\preceq$-increasing, it follows that, for any given $h \in icbS$, the entire sequence $\langle \mu Q^t, h \rangle$ converges up to $\langle \psi^*, h \rangle$. Because $\{\mu Q^t\}$ is tight, part 3 of lemma 7.6 implies that $\mu Q^t \to \psi^*$.

In addition to $\mu Q^t \to \psi^*$, we also have $\mu Q^t \preceq \psi^*$ for all $t \geq 0$, because for any $h \in icbS$ and $t \geq 0$ we have

$$\langle \mu Q^t, h \rangle \leq \sup_{t \geq 0} \langle \mu Q^t, h \rangle = \lim_{t \to \infty} \langle \mu Q^t, h \rangle = \langle \psi^*, h \rangle.$$

The inequality $\mu Q^t \preceq \psi^*$ now follows from part 1 of lemma 7.6.

Next, we claim that $\psi^* \preceq \psi^* Q$. To see this, pick any $h \in icbS$. Since $\mu Q^t \preceq \psi^*$ for all t, and since $Qh \in ibS$,

$$\langle \mu Q^t, Qh \rangle \leq \langle \psi^*, Qh \rangle = \langle \psi^* Q, h \rangle.$$

Using this inequality and the fact that $h \in cbS$, we obtain

$$\langle \psi^*, h \rangle = \lim_{t \to \infty} \langle \mu Q^{t+1}, h \rangle = \lim_{t \to \infty} \langle \mu Q^t, Qh \rangle \leq \langle \psi^* Q, h \rangle.$$

Hence $\langle \psi^*, h \rangle \leq \langle \psi^* Q, h \rangle$ for all $h \in icbS$, and $\psi^* \preceq \psi^* Q$ as claimed. Iterating on this inequality we obtain $\psi^* \preceq \psi^* Q^t$ for all t.

To summarize our results so far, we have $\mu Q^t \preceq \psi^* \preceq \psi^* Q \preceq \psi^* Q^t$ for all $t \geq 0$, and hence

$$\langle \mu Q^t, h \rangle \leq \langle \psi^*, h \rangle \leq \langle \psi^* Q, h \rangle \leq \langle \psi^* Q^t, h \rangle \quad \text{for all} \quad h \in icbS.$$

Applying (20), we obtain $\langle \psi^*, h \rangle = \langle \psi^* Q, h \rangle$ for all $h \in icbS$. By lemma 7.6, this implies that $\psi^* = \psi^* Q$. In other words, ψ^* is stationary for Q.

It remains to show that Q is globally stable. Fixing $\mu \in \mathscr{P}_S$ and applying (20) again, we have

$$\langle \mu Q^t, h \rangle \to \langle \psi^*, h \rangle, \quad \forall\, h \in ibS. \tag{21}$$

Since $icbS \subset ibS$ and $\{\mu Q^t\}$ is tight (cf., lemma 7.3), this implies that $\mu Q^t \to \psi^*$ (lemma 7.6, part 3). Finally, uniqueness is also immediate, because if μ is also stationary, then by (21) we have $\langle \mu, h \rangle = \langle \psi^*, h \rangle$ for all $h \in icbS$. By lemma 7.6, we then have $\mu = \psi^*$. □

Proof of theorem 3.2. Under the conditions of the theorem, Q is order mixing, as proved in lemma 7.5. In addition, boundedness in probability and the Feller property guarantee the existence of a stationary distribution by the Krylov-Bogolubov theorem [44, proposition 12.1.3 and lemma D.5.3]. Given existence of a stationary distribution ψ^*, the proof that Q is globally stable is now identical to the proof of the same claim given for theorem 3.1 (see the discussion surrounding equation (21)). □

Next we turn to the proof of theorem 3.3. In the proof, S^∞ is the space of S-valued sequences, and $\mathscr{B}_S^\infty$ is the product σ-algebra. If $\{X_t\}$ is Markov-(Q, x) then $\mathbf{P}_x^Q$ denotes its distribution on $(S^\infty, \mathscr{B}_S^\infty)$, and $\mathbf{E}_x^Q$ is the corresponding expectation. Without loss of generality, we can set $(\Omega, \mathscr{F}) = (S^\infty, \mathscr{B}_S^\infty)$, and define X_t to be the t-th projection of S^∞ onto S, so that $X_t(\omega) = X_t(x_0, x_1, \ldots) = x_t$ [44, chapter 3]. The shift operator θ is defined as usual by $\theta(x_0, x_1, \ldots) = (x_1, x_2, \ldots)$. For any random variable H we set $\theta H(\omega) = H(\theta\omega)$. We say that H is *shift-invariant* if $\theta H = H$. We make use of the strong Markov property, which states [44, p. 66] that if τ is a stopping time and $\mathscr{F}_\tau$ is the induced σ-algebra, then $\mathbf{E}_x^Q[\theta^\tau H \,|\, \mathscr{F}_\tau] = \mathbf{E}_{X_\tau}^Q H$ on the set $\{\tau < \infty\}$. Finally, we say that a bounded measurable function f is *harmonic* if $Qf = f$.

Lemma 7.7. *Let Q be any stochastic kernel, let H be a bounded random variable, and let $f(x) := \mathbf{E}_x^Q H$. If H is shift invariant, then f is harmonic. If, in addition, τ is a stopping time satisfying $\mathbf{P}_x^Q\{\tau < \infty\} = 1$, then $\mathbf{E}_x^Q f(X_\tau) = f(x)$ for all $x \in S$.*

Proof. The first assertion follows from the second by choosing $\tau \equiv 1$. Regarding the second assertion, fix $x \in S$. The strong Markov property and the law of iterated expectations give $\mathbf{E}_x^Q\, f(X_\tau) = \mathbf{E}_x^Q\, \mathbf{E}_{X_\tau}^Q H = \mathbf{E}_x^Q\, \mathbf{E}_x^Q[\theta^\tau H \,|\, \mathscr{F}_\tau] = \mathbf{E}_x^Q\, \theta^\tau H = \mathbf{E}_x^Q\, H = f(x)$. $\square$

Lemma 7.8. *Let Q be any stochastic kernel and let $f\colon S \to \mathbb{R}_+$. If f is harmonic, then the set $\{x \in S : f(x) = 0\}$ is absorbing.*[37]

Proof. Let $F := \{x \in S : f(x) = 0\}$. The claim is that $Q(x, F) = 1$ for all $x \in F$. To see this, fix $x \in F$ and observe that $0 = f(x) = Qf(x) = \mathbf{E}_x^Q f(X_1) = \mathbf{E}_x^Q f(X_1)\mathbb{1}\{f(X_1) > 0\}$. It follows that $\mathbf{P}_x^Q \mathbb{1}\{f(X_1) > 0\} = 0$, or $Q(x, F) = \mathbf{P}_x^Q \mathbb{1}\{f(X_1) = 0\} = 1$. $\square$

The following lemma is also relatively straightforward, although the proof is delayed until the end of this section:

Lemma 7.9. *Let Q be order mixing and let D be a Borel set. If D is nonempty, decreasing and absorbing, then D is strongly accessible.*

We endow the set S^∞ with the pointwise order inherited from $(S, \leq)$. In particular, we say that $\{x_t\} \leq \{x'_t\}$ if $x_t \leq x'_t$ in S for all t. We let ibS^∞ be the set of bounded measurable increasing functions from S^∞ to $\mathbb{R}$, and for probability measures ν and ν' on $(S^\infty, \mathscr{B}_S^\infty)$ we write $\nu \preceq \nu'$ if $\int h d\nu \leq \int h d\nu'$ for all $h \in ibS^\infty$.

The next two results are proved in [34, proposition 2, theorem 2] and [25, corollary 2.5.2] respectively.

Theorem 7.1. *If Q is an increasing stochastic kernel on S and $x, y \in S$ with $x \leq y$, then $\mathbf{P}_x^Q \preceq \mathbf{P}_y^Q$.*

[37] A set $C \in \mathscr{B}_S$ is called *absorbing* for Q if $Q(x, C) = 1$ for all $x \in C$.

Theorem 7.2. *If Q is a stochastic kernel on S with unique stationary distribution μ^* and $h\colon S \to \mathbb{R}$ satisfies $\int |h| \, d\mu^* < \infty$, then there exists a set $R \in \mathscr{B}_S$ such that*

1. *$\mu^*(R) = 1$, and*
2. *$n^{-1} \sum_{t=1}^n h(X_t) \to \int h \, d\mu^*$ almost surely whenever $x \in R$ and $\{X_t\}$ is Markov-(Q, x).*

Theorem 7.2 is important from a theoretical perspective, but of limited practical value, since the set R typically depends on h and is difficult to identify. However, if the conditions of theorem 3.3 are satisfied, then R is all of S, as we now show.

Proof of theorem 3.3. Let Q be an increasing, order reversing kernel that is bounded in probability, and has stationary distribution μ^*. Note that Q is order mixing by lemma 7.5, and globally stable since the conditions of theorem 3.1 all hold. We claim that (Q, h) satisfies the strong law of large numbers whenever $h \in ibS$. To see this, fix $h \in ibS$. Without loss of generality we can assume that $\langle \mu^*, h \rangle = 0$, so the claim is that, for all $x \in S$, we have $n^{-1} \sum_{t=1}^n h(X_t) \to 0$ almost surely whenever $\{X_t\}$ is Markov-(Q, x). To prove this claim, we define

$$G := \mathbb{1} \left\{ \limsup_{n \to \infty} \frac{1}{n} \sum_{t=1}^n h(X_t) > 0 \right\} \quad \text{and} \quad \alpha(x) := \mathbf{E}_x^Q \, G.$$

The random variable G is bounded and shift-invariant, so lemma 7.7 applies to α. In addition, the function α is increasing on S. To see this, observe that G can be written as $G = g(\{X_t\})$ where $g(\{x_t\}) = \mathbb{1}\{\limsup n^{-1} \sum_{t=1}^n h(x_t) > 0\}$. Since $h \in ibS$, we have $g \in ibS^\infty$. Taking $x \leq x'$ and applying theorem 7.1, we obtain

$$\alpha(x) = \mathbf{E}_x^Q G = \int g \, d\mathbf{P}_x^Q \leq \int g \, d\mathbf{P}_{x'}^Q = \alpha(x').$$

Now let $D := \{x \in S : \alpha(x) = 0\}$. Since α is increasing, the set D is decreasing. By theorem 7.2, there exists a set R with $\pi^*(R) = 1$ and $R \subset D$. In particular, D is nonempty. By lemmas 7.7 and 7.8, the set D is absorbing. Giving that Q is order mixing, lemma 7.9 now implies that D is strongly accessible.

Pick any $x \in S$. Let $\{X_t\}$ be Markov-(Q, x), and let $\tau := \inf\{t \geq 0 : X_t \in D\}$. Because D is strongly accessible, τ is finite with probability one. As a result, lemma 7.7 applies, and we have $\alpha(x) = \mathbf{E}_x^Q \alpha(X_\tau)$. This expectation is zero, since $X_\tau \in D$ by definition. We conclude that $\alpha(x) = 0$ for any given $x \in S$. In other words,

$$\mathbf{P}_x^Q \left\{ \limsup_{n \to \infty} \frac{1}{n} \sum_{t=1}^n h(X_t) > 0 \right\} = 0 \qquad \text{for all } x \in S.$$

A symmetric argument shows that

$$\mathbf{P}_x^Q \left\{ \liminf_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} h(X_t) < 0 \right\} = 0 \qquad \text{for all } x \in S.$$

Hence (Q, h) satisfies the strong law of large numbers as claimed. $\square$

Now we turn to the proof of theorem 3.4. In the proof, we let $ic(S, [0,1])$ be the functions in $icbS$ taking values in $[0,1]$. As usual, $\mu_n \to \mu$ means that $\langle \mu_n, f\rangle \to \langle \mu, f\rangle$ for all $f \in cbS$. Also, we require the following definition: Letting $\mathscr{G}$ and $\mathscr{H}$ be sets of bounded measurable functions, we say that $\mathscr{H}$ is *monotonically approximated by* $\mathscr{G}$ if, for all $h \in \mathscr{H}$, there exist sequences $\{g_n^1\}$ and $\{g_n^2\}$ in $\mathscr{G}$ with $g_n^1 \uparrow h$ and $g_n^2 \downarrow h$ pointwise. The proofs of the next two lemmas are given at the end of this section.

Lemma 7.10. *If $\mathscr{H}$ is monotonically approximated by $\mathscr{G}$, then $\mathscr{G}$ is convergence determining for $\mathscr{H}$, in the sense that if $\{\nu_n\}$ and ν are elements of $\mathscr{P}_S$, and $\langle \nu_n, g\rangle \to \langle \nu, g\rangle$ for all $g \in \mathscr{G}$, then $\langle \nu_n, h\rangle \to \langle \nu, h\rangle$ for all $h \in \mathscr{H}$.*

Lemma 7.11. *If the conditions of theorem 3.4 hold, then there exists a countable class $\mathscr{G}$ such that (Q, g) satisfies the strong law of large numbers for every $g \in \mathscr{G}$, and, moreover, $ic(S, [0,1])$ is monotonically approximated by $\mathscr{G}$.*

Proof of theorem 3.4. Fix $x \in S$. Let $\{X_t\}$ be Markov-(Q, x), and let μ_n be the associated empirical distribution. In particular, we let $\mu_n^\omega \in \mathscr{P}_S$ be defined by $\mu_n^\omega(B) := n^{-1} \sum_{t=1}^{n} \mathbb{1}\{X_t(\omega) \in B\}$. As a first step of the proof, we claim that $\{\mu_n^\omega\}$ is tight with probability one. To see this, fix $\epsilon > 0$, and let K be a compact subset of S with $\mu^*(K) \geq 1 - \epsilon$. Since compact subsets of S are order bounded, we have $a, b \in S$ with $K \subset [a, b]$. Let $I := \{y \in S : y \geq a\}$ and $J := \{y \in S : y \not\leq b\}$. Note that both of these sets are increasing, and $[a, b] = I \setminus J$. Appealing to theorem 3.3, we let $F_a \subset \Omega$ with $\mathbb{P}(F_a) = 1$ and $\mu_n^\omega(I) \to \mu^*(I)$ whenever $\omega \in F_a$; and $F_b \subset \Omega$ with $\mathbb{P}(F_b) = 1$ and $\mu_n^\omega(J) \to \mu^*(J)$ whenever $\omega \in F_b$. Picking arbitrary $\omega \in F := F_a \cap F_b$, we have

$$\mu_n^\omega([a, b]) = \mu_n^\omega(I) - \mu_n^\omega(J) \to \mu^*(I) - \mu^*(J) = \mu^*([a, b]) \geq \mu^*(K) \geq 1 - \epsilon.$$

Since closed order intervals are compact by assumption, it follows that $\{\mu_n^\omega\}$ is tight for all ω in the probability one set F.

As the second step of the proof, we claim there exists a probability one set F' such that, for any given $\omega \in F'$, we have $\langle \mu_n^\omega, f\rangle \to \langle \mu^*, f\rangle$ for all $f \in icbS$. To see that this is so, let $\mathscr{G}$ be as in lemma 7.11. Since $\mathscr{G}$ is countable and the law of large numbers holds for every element of $\mathscr{G}$, there exists a probability one set $F' \subset \Omega$ such that, for each $\omega \in F'$, we have $\langle \mu_n^\omega, g\rangle \to \langle \mu^*, g\rangle$ for all $g \in \mathscr{G}$. Fix $\omega \in F'$. Since $ic(S, [0,1])$ is monotonically approximated by $\mathscr{G}$, lemma 7.10 implies that $\langle \mu_n^\omega, f\rangle \to \langle \mu^*, f\rangle$ for all $f \in ic(S, [0,1])$. It immediately follows that $\langle \mu_n^\omega, f\rangle \to \langle \mu^*, f\rangle$ for all $f \in icbS$.[38]

Now let F'' be the probability one set $F \cap F'$. For any $\omega \in F''$, the sequence of distributions $\{\mu_n^\omega\}$ is tight, and satisfies $\langle \mu_n^\omega, f\rangle \to \langle \mu^*, f\rangle$ for all $f \in icbS$. In view of lemma 7.6, we then have $\langle \mu_n^\omega, f\rangle \to \langle \mu^*, f\rangle$ for all $f \in cbS$. This concludes the proof of theorem 3.4. □

Proof of proposition 4.1. Pick any $x \in S$ and fix $\epsilon > 0$. Let K be a compact set such that $Q_\ell^t(x, K) \geq 1 - \epsilon$ for all t. Compact sets are order bounded by assumption, so there exists an $x_\ell \in S$ with $K \subset I := \{y \in S : y \geq x_\ell\}$. As a result, $Q_\ell^t(x, I) \geq 1 - \epsilon$ for all t. Since $Q_\ell \preceq Q$, and since I is an increasing set, we have $Q_\ell^t \mathbb{1}_I \leq Q^t \mathbb{1}_I$, and hence

$$1 - \epsilon \leq Q_\ell^t(x, I) = Q_\ell^t \mathbb{1}_I(x) \leq Q^t \mathbb{1}_I(x) = Q^t(x, I) = Q^t(x, \{y \in S : y \geq x_\ell\}).$$

Applying similar reasoning to $Q \preceq Q_u$,

$$\exists\, x_u \in S \quad \text{such that} \quad 1 - \epsilon \leq Q^t(x, \{y \in S : y \leq x_u\}) \quad \text{for all} \quad t \geq 0.$$

Combining these bounds, we obtain $Q^t(x, [x_\ell, x_u]) \geq 1 - 2\epsilon$ for all t. Evidently $[x_\ell, x_u]$ is order bounded. Moreover, since the partial order $\leq$ is assumed to be closed, $[x_\ell, x_u]$ is also compact. We have now shown that $\{Q^t(x, \cdot)\}$ is tight, and, as $x \in S$ was arbitrary, Q is bounded in probability. □

Proof of proposition 4.2. Suppose that Q_ℓ is Feller and bounded in probability with $Q_\ell \preceq Q$. By the Krylov-Bogolubov theorem [44, proposition 12.1.3 and lemma D.5.3], Q_ℓ has at least one stationary distribution μ. For this μ we have $\mu = \mu Q_\ell \preceq \mu Q$. In other words, μ is deficient for Q. A similar argument shows that if Q_u is Feller and bounded in probability with $Q \preceq Q_u$ then Q has an excessive distribution. □

Proof of proposition 4.3. Let Q be bounded in probability. Suppose first that Q is upward reaching. Pick any $(x, x') \in S \times S$. Let $\{X_t\}$ and $\{X_t'\}$ be independent, Markov-(Q, x) and Markov-(Q, x') respectively. We need to prove existence of a $k \in \mathbb{N}$ such that $\mathbb{P}\{X_k \leq X_k'\} > 0$.

Since Q is bounded in probability, there exists a compact $C \subset S$ with $\mathbb{P}\{X_t \in C\} > 0$ for all $t \geq 0$. Since compact sets are assumed to be order bounded, we can take an order interval $[a, b]$ of S with $C \subset [a, b]$. For this a, b we have $\mathbb{P}\{a \leq X_t \leq b\} > 0$ for all $t \geq 0$. As Q is upward reaching, there is a $k \in \mathbb{N}$ such that $\mathbb{P}\{b \leq X_k'\} > 0$. Using independence, we now have

$$\mathbb{P}\{X_k \leq X_k'\} \geq \mathbb{P}\{X_k \leq b \leq X_k'\} = \mathbb{P}\{X_k \leq b\}\mathbb{P}\{b \leq X_k'\} > 0,$$

as was to be shown. The proof for the downward reaching case is similar. □

[38]If $f \in icbS$, then there exists a $g \in ic(S, [0,1])$ and constants a, b such that $f = a + bg$.

Proof of corollary 4.1. We discuss only theorem 3.1, since the discussion for theorem 3.2 is almost identical. To begin, observe that if Q is globally stable then Q is bounded in probability and has an excessive distribution, as shown in the proof of theorem 3.1. Conversely, suppose that Q is increasing, upward reaching, bounded in probability and has either a deficient or an excessive distribution. By proposition 4.3, Q is order reversing. Theorem 3.1 now implies that Q is globally stable. The proof of the downward reaching case is similar. □

Proof of proposition 4.4. Let $\{\xi_t\}$ and $\{\xi'_t\}$ be IID draws from ϕ and independent of each other. Consider first condition 3. We claim that Q_F is order reversing. Fix $x, x' \in S$. Let $\{z_t\}_{t=1}^k$ and $\{z'_t\}_{t=1}^k$ be as in the statement of the proposition. Define the constant

$$\gamma := \mathbb{P}\{F^k(x, \xi_1, \ldots, \xi_k) < F^k(x', \xi'_1, \ldots, \xi'_k)\}.$$

We need only show that $\gamma > 0$. By hypothesis, $F^k(x, z_1, \ldots, z_k) < F^k(x', z'_1, \ldots, z'_k)$. By continuity of F, there exist open neighborhoods N_t of z_t and N'_t of z'_t such that

$$\tilde{z}_t \in N_t \text{ and } \tilde{z}'_t \in N'_t \text{ for } t \in \{1, \ldots, k\} \implies F^k(x, \tilde{z}_1, \ldots, \tilde{z}_k) < F^k(x', \tilde{z}'_1, \ldots, \tilde{z}'_k).$$

This leads to the estimate

$$\gamma \geq \mathbb{P} \cap_{t=1}^n \{\xi_t \in N_t \text{ and } \xi'_t \in N'_t\} = \prod_{t=1}^n \phi(N_t)\phi(N'_t).$$

Since Z is the support of ϕ, this last term is positive, and $\gamma > 0$. Thus, Q_F is order reversing. Since Q_F is also Feller, increasing and bounded in probability, theorem 3.2 implies that Q_F is globally stable.

The proof of the corollary will be complete if conditions 1–2 of the corollary imply that Q_F is upward and downward reaching respectively (see proposition 4.3). The arguments are very similar to the proof just completed and hence we omit them. □

Finally, we complete the proof of all remaining lemmas stated in this section.

Proof of lemma 7.1. Let B be a C-accessible subset of D. To prove the lemma, it suffices to show that $\mathbf{P}_x^\Phi \cup_t \{X_t \in B\} = 1$ whenever $\{\Phi^t(x, \cdot)\}$ is tight. To this end, fix $x \in D$, and assume that $\{\Phi^t(x, \cdot)\}$ is tight. Let $\tau := \inf\{t \geq 0 : X_t \in B\}$. Evidently we have $\cup_{t=0}^\infty \{X_t \in B\} = \{\tau < \infty\}$. Thus, we need to show that $\mathbf{P}_x^\Phi\{\tau < \infty\} = 1$.

Fix $\epsilon > 0$. Since $\{\Phi^t(x, \cdot)\}$ is tight, there exists a compact set C such that

$$\inf_t \mathbf{P}_x^\Phi\{X_t \in C\} = \inf_t \Phi^t(x, C) \geq 1 - \epsilon.$$

Since B is C-accessible, there exists an $n \in \mathbb{N}$ and $\delta > 0$ such that $\inf_{y \in C} \Phi^n(y, B) \geq \delta$. For $t \in \mathbb{N}$, define $p_t := \mathbf{P}_x^\Phi\{\tau \leq tn\}$. We wish to obtain a relationship between p_t and p_{t+1}. To this end, note that

$$
\begin{aligned}
\mathbb{1}\{\tau \le (t+1)n\} &= \mathbb{1}\{\tau \le tn\} + \mathbb{1}\{\tau > tn\}\mathbb{1}\{\tau \le (t+1)n\} \\
&\ge \mathbb{1}\{\tau \le tn\} + \mathbb{1}\{\tau > tn\}\mathbb{1}\{X_{(t+1)n} \in B\} \\
&\ge \mathbb{1}\{\tau \le tn\} + \mathbb{1}\{\tau > tn\}\mathbb{1}\{X_{tn} \in C\}\mathbb{1}\{X_{(t+1)n} \in B\}.
\end{aligned}
$$

Taking expectations yields

$$p_{t+1} \ge p_t + \mathbf{E}_x^{\Phi}\mathbb{1}\{\tau > tn\}\mathbb{1}\{X_{tn} \in C\}\mathbb{1}\{X_{(t+1)n} \in B\}.$$

We estimate the last expectation as follows:

$$
\begin{aligned}
&\mathbf{E}_x^{\Phi}\mathbb{1}\{\tau > tn\}\mathbb{1}\{X_{tn} \in C\}\mathbb{1}\{X_{(t+1)n} \in B\} \\
&\qquad = \mathbf{E}_x^{\Phi}[\mathbb{1}\{\tau > tn\}\mathbb{1}\{X_{tn} \in C\}\mathbf{E}_x^{\Phi}[\mathbb{1}\{X_{(t+1)n} \in B\}|\mathscr{F}_{tn}]] \\
&\qquad = \mathbf{E}_x^{\Phi}[\mathbb{1}\{\tau > tn\}\mathbb{1}\{X_{tn} \in C\}\Phi^n(X_{tn}, B)] \\
&\qquad \ge \mathbf{E}_x^{\Phi}\mathbb{1}\{\tau > tn\}\mathbb{1}\{X_{tn} \in C\}\delta \\
&\qquad = \mathbf{E}_x^{\Phi}(1 - \mathbb{1}\{\tau \le tn\})\mathbb{1}\{X_{tn} \in C\}\delta \\
&\qquad = \mathbf{E}_x^{\Phi}\mathbb{1}\{X_{tn} \in C\}\delta - \mathbf{E}_x^{\Phi}\mathbb{1}\{\tau \le tn\}\mathbb{1}\{X_{tn} \in C\}\delta \\
&\qquad \ge (1-\epsilon)\delta - \mathbf{E}_x^{\Phi}\mathbb{1}\{\tau \le tn\}\delta \\
&\qquad = (1-\epsilon)\delta - p_t\delta.
\end{aligned}
$$

$$\therefore \quad p_{t+1} \ge p_t + (1-\epsilon)\delta - p_t\delta = (1-\delta)p_t + (1-\epsilon)\delta.$$

The unique, globally stable fixed point of $q_{t+1} = (1-\delta)q_t + (1-\epsilon)\delta$ is $1-\epsilon$, so

$$1 - \epsilon \le \lim_{t\to\infty} p_t = \mathbf{P}_x^{\Phi}\{\tau < \infty\} \le 1$$

Since ϵ was arbitrary, we obtain $\mathbf{P}_x^{\Phi}\{\tau < \infty\} = 1$. □

Proof of lemma 7.2. Fix $x, x' \in S$ and $\epsilon > 0$. Since Q is bounded in probability, we can choose compact sets C and C' such that

$$Q^t(x, C) \ge (1-\epsilon)^{1/2} \quad \text{and} \quad Q^t(x', C') \ge (1-\epsilon)^{1/2} \quad \text{for all } t.$$

$$\therefore \quad (Q \times Q)^t((x, x'), C \times C') = Q^t(x, C)Q^t(x', C') \ge 1 - \epsilon \quad \text{for all } t.$$

Since $C \times C'$ is compact in the product space, $Q \times Q$ is bounded in probability. □

Proof of lemma 7.3. Fix $\mu \in \mathscr{P}_S$ and $\epsilon > 0$. Since individual elements of $\mathscr{P}_S$ are tight (Dudley, 2002, theorem 11.5.1), we can choose a compact set $C_\mu \subset S$ with $\mu(C_\mu) \ge 1 - \epsilon$. By assumption, we can take an order interval $[a, b]$ of S with $C_\mu \subset [a, b]$. For this a, b, we have

$$\mu([a,b]^c) = \mu(S \setminus [a,b]) \le \epsilon. \tag{22}$$

By hypothesis, $\{Q^t(x,\cdot)\}$ is tight for all $x \in S$, so we choose compact subsets C_a and C_b of S with $Q^t(a, C_a) \geq 1-\epsilon$ and $Q^t(b, C_b) \geq 1-\epsilon$ for all t. Since $C_a \cup C_b$ is also compact, we can take an order interval $[\alpha, \beta]$ of S with $C_a \cup C_b \subset [\alpha, \beta] \subset S$. We then have $Q^t(a, [\alpha, \beta]) \geq 1-\epsilon$ and $Q^t(b, [\alpha, \beta]) \geq 1-\epsilon$ for all t. Letting $I_\alpha := \{x \in S : x \geq \alpha\}$ and $D_\beta := \{x \in S : x \leq \beta\}$, this leads to

$$Q^t(a, I_\alpha) \geq 1-\epsilon \quad \text{and} \quad Q^t(b, D_\beta) \geq 1-\epsilon \quad \text{for all } t. \tag{23}$$

In view of remark 2.1 and (23), we have

$$a \leq x \implies Q^t(x, I_\alpha) \geq Q^t(a, I_\alpha) \geq 1-\epsilon,$$

and, by a similar argument,

$$x \leq b \implies Q^t(x, D_\beta) \geq Q^t(b, D_\beta) \geq 1-\epsilon.$$

Since $[\alpha, \beta] := \{x \in S : \alpha \leq x \leq \beta\} = I_\alpha \cap D_\beta$, we have

$$Q^t(x, [\alpha, \beta]^c) = Q^t(x, D_\beta^c \cup I_\alpha^c) \leq 2 - Q^t(x, D_\beta) - Q^t(x, I_\alpha).$$

This leads to the estimate

$$a \leq x \leq b \implies Q^t(x, [\alpha, \beta]^c) \leq 2\epsilon. \tag{24}$$

Combining (22) and (24), we now have

$$\begin{aligned}
\mu Q^t([\alpha, \beta]^c) &= \int Q^t(x, [\alpha, \beta]^c)\mu(dx) \\
&= \int_{[a,b]} Q^t(x, [\alpha, \beta]^c)\mu(dx) + \int_{[a,b]^c} Q^t(x, [\alpha, \beta]^c)\mu(dx) \\
&\leq \int_{[a,b]} 2\epsilon\, \mu(dx) + \mu([\alpha, \beta]^c) \;\leq\; 3\epsilon.
\end{aligned}$$

Since $[\alpha, \beta]$ is compact and t is arbitrary, we conclude that $\{\mu Q^t\}$ is tight. □

Proof of lemma 7.4. Let C be any compact subset of $S \times S$. We need to prove existence of an $n \in \mathbb{N}$ and $\delta > 0$ such that $(Q \times Q)^n((x, x'), \mathbb{G}) \geq \delta$ whenever $(x, x') \in C$. To do so, we introduce the function

$$\psi_n(x, x') := (Q \times Q)^n((x, x'), \mathbb{G}) = \mathbf{P}^{Q \times Q}_{x,x'}\{X_n \leq X'_n\}.$$

where (X_n, X'_n) is Markov-$(Q \times Q, (x, x'))$. Intuitively, since Q is increasing, the event $\{X_n \leq X'_n\}$ becomes less likely as x rises and x' falls, and hence $\psi_n(x, x')$ is decreasing in x and increasing in x' for each n. A routine argument confirms this is the case.

Since $C \subset S \times S$ is compact, we can take an order interval $[a, b]$ of S with $C \subset [a, b] \times [a, b]$.[39] Moreover, since Q is order reversing, we can take $n \in \mathbb{N}$ such that $\delta := \psi_n(b, a) > 0$. Observe that

$$(x, x') \in C \implies (x, x') \in [a, b] \times [a, b] \implies x \leq b \text{ and } x' \geq a.$$

$$\therefore \quad (x, x') \in C \implies (Q \times Q)^n((x, x'), \mathbb{G}) = \psi_n(x, x') \geq \psi_n(b, a) = \delta.$$

In other words, $\mathbb{G}$ is C-accessible for $Q \times Q$. □

Proof of lemma 7.6. The statement $\mu \preceq \mu'$ iff $\langle \mu, h \rangle \leq \langle \mu', h \rangle$ for all $h \in icbS$ holds for every normally ordered space, as shown by Whitt [63, theorem 2.6]. Moreover, since $\preceq$ is a partial order on $\mathscr{P}_S$ [33, theorem 2], and hence antisymmetric, it follows that $\mu = \mu'$ iff $\langle \mu, h \rangle = \langle \mu', h \rangle$ for all $h \in icbS$. Regarding the third assertion of the lemma, observer first that if $\mu_n \to \mu$, then since S is polish the sequence $\{\mu_n\}$ is tight [24, theorem 11.5.3]. The statement $\langle \mu_n, h \rangle \to \langle \mu, h \rangle$ whenever $h \in icbS$ is obvious. To prove the converse, suppose that $\{\mu_n\}$ is tight and $\langle \mu_n, h \rangle \to \langle \mu, h \rangle$ for all $h \in icbS$. Take any subsequence $\{\mu_n\}_{n \in \mathbb{N}_1}$ of $\{\mu_n\}$. By tightness and Prohorov's theorem [24, theorem 11.5.4], this subsequence has a subsubsequence converging to some $\nu \in \mathscr{P}_S$:

$$\exists\, \mathbb{N}_2 \subset \mathbb{N}_1 \quad \text{such that} \quad \lim_{n \in \mathbb{N}_2} \langle \mu_n, h \rangle = \langle \nu, h \rangle \quad \text{for all} \quad h \in cbS.$$

Since $\langle \mu_n, h \rangle \to \langle \mu, h \rangle$ for all $h \in icbS$, we now have $\lim_{n \in \mathbb{N}_2} \langle \mu_n, h \rangle = \langle \nu, h \rangle = \langle \mu, h \rangle$ for all $h \in icbS$, and hence $\nu = \mu$. We have now shown that every subsequence of $\{\mu_n\}$ has a subsubsequence converging to μ, and hence the entire sequence also converges to μ. □

Proof of lemma 7.9. Pick any $x \in S$. Since D is nonempty we can choose some $x^d \in D$. We let $\{X_t\}$ and $\{X_t^d\}$ be independent and Markov-(Q, x) and (Q, x^d) respectively, defined on some common probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Define $\tau := \inf\{t \geq 0 : X_t \leq X_t^d\}$. Since Q is order mixing, $\mathbb{P}\{\tau < \infty\} = 1$. Since D is absorbing and $x^d \in D$, the set $\{X_t^d \in D\}$ also has probability one. Therefore $\mathbb{P}\{X_t^d \in D\} \cap \{\tau < \infty\} = 1$. Moreover, since D is decreasing and, by definition, $X_\tau \leq X_\tau^d$ on $\{\tau < \infty\}$, we have

$$\{X_t^d \in D\} \cap \{\tau < \infty\} \subset \{X_t \in D\} \cap \{\tau < \infty\} \subset \cup_{t=0}^{\infty} \{X_t \in D\}.$$

It now follows that $\mathbb{P} \cup_{t=0}^{\infty} \{X_t \in D\} = 1$. As x was arbitrary, we conclude that D is strongly accessible. □

[39] To see this, let K be a compact subset of S with $C \subset K \times K$. (Such a K can be obtained by projecting C onto the first and second axis, and defining K as the union of these projections.) Since K is order bounded in S by assumption, we just choose $a, b \in S$ with $K \subset [a, b]$.

Proof of lemma 7.10. Let $\{\nu_n\}$ and ν be probability measures on S, and suppose that $\langle \nu_n, g\rangle \to \langle \nu, g\rangle$ for all $g \in \mathscr{G} \subset bS$. We claim that $\langle \nu_n, h\rangle \to \langle \nu, h\rangle$ for all $h \in \mathscr{H} \subset bS$. To see this, pick any $h \in \mathscr{H}$, and choose sequences $\{g_n^1\}$ and $\{g_n^2\}$ in $\mathscr{G}$ with $g_n^1 \uparrow h$ and $g_n^2 \downarrow h$. Clearly

$$\liminf_n \langle \nu_n, h\rangle \geq \liminf_n \langle \nu_n, g_k^1\rangle = \langle \nu, g_k^1\rangle \quad \text{for all } k.$$

$$\therefore \quad \liminf_n \langle \nu_n, h\rangle \geq \sup_k \langle \nu, g_k^1\rangle = \lim_k \langle \nu, g_k^1\rangle = \langle \nu, h\rangle.$$

A symmetric argument applied to $\{g_n^2\}$ yields $\limsup_n \langle \nu_n, h\rangle \leq \langle \nu, h\rangle$. □

Proof of lemma 7.11. Let A be the countable subset of S in assumption 3.1. For $a \in A$, let $I_a := \mathbb{1}\{y \in S : y \geq a\}$. Let $\mathscr{L}$ be the set of functions $\ell = rI_a$ for some $r \in \mathbb{Q} \cap [0,1]$ and $a \in A$. Let $\mathscr{G}_1$ be all functions $g = \sup_{\ell \in F} \ell$ where $F \subset \mathscr{L}$ is finite. Clearly $\mathscr{G}_1$ is countable, and, by theorem 3.3, every element of $\mathscr{G}_1$ satisfies the strong law of large numbers. We claim that for each $f \in ic(S, [0,1])$ there exists a sequence $\{g_n\}$ in $\mathscr{G}_1$ converging up to f. To verify this claim it suffices to show that

$$\sup\{\ell(x) : \ell \in \mathscr{L} \text{ and } \ell \leq f\} = f(x) \quad \text{for any } x \in S. \tag{25}$$

Indeed, if (25) is valid, then take $\{\ell_k\}$ to be an enumeration of all $\ell \in \mathscr{L}$ with $\ell \leq f$ and choose $g_n = \sup_{1 \leq k \leq n} \ell_k$.

To establish (25), fix $x \in S$ and $\epsilon > 0$. By continuity of f and assumption 3.1, we can find an $a \in A$ with $a \leq x$ and $f(x) - \epsilon < f(a)$. Let $r \in \mathbb{Q}$ be such that $f(x) - \epsilon < r < f(a)$ and let $\ell(x) := rI_a$. Since $\ell \leq f(a)I_a$ and f is increasing we have $\ell \leq f$. On the other hand, $f(x) - \epsilon < r = \ell(a) \leq \ell(x)$. Since ϵ was arbitrary we conclude that (25) is valid.

To complete the proof of lemma 7.11, we show existence of a class of functions $\mathscr{G}_2$ such that $\mathscr{G}_2$ is countable, every element of $\mathscr{G}_2$ satisfies the strong law of large numbers, and, for each $f \in ic(S, [0,1])$, there exists a sequence $\{g_n\}$ in $\mathscr{G}_2$ converging down to f. The claim in lemma 7.11 is then satisfied with $\mathscr{G} := \mathscr{G}_1 \cup \mathscr{G}_2$.

To construct $\mathscr{G}_2$, let $D_a := \mathbb{1}\{y \in S : y \leq a\}$, and let $\mathscr{L}$ be the set of functions $\ell \colon S \to [0,1]$ with $\ell(x) = 1 - rD_a$ for some $r \in \mathbb{Q} \cap [0,1]$ and $a \in A$. Let $\mathscr{G}_2$ be all functions $g = \inf_{\ell \in F} \ell$ where $F \subset \mathscr{L}$ is finite. The set $\mathscr{G}_2$ is countable and every element satisfies the LLN. We claim that for each $f \in ic(S, [0,1])$ there exists a sequence $\{g_n\}$ in $\mathscr{G}_2$ converging down to f. To verify this claim it suffices to show that

$$\inf\{\ell(x) : \ell \in \mathscr{L} \text{ and } \ell \geq f\} = f(x) \quad \text{for any } x \in S. \tag{26}$$

To establish (26), fix $x \in S$ and $\epsilon > 0$. By continuity of f and assumption 3.1, we can find an $a \in A$ with $x \leq a$ and $f(a) < f(x) + \epsilon$. Let $r \in \mathbb{Q}$ be such that $f(a) < 1 - r < f(x) + \epsilon$ and let $\ell(x) := 1 - rD_a$. It is now easy to check that $\ell \geq f$ and $\ell(x) < f(x) + \epsilon$. Since ϵ was arbitrary we conclude that (26) is valid. □

References

[1] P. Aghion, P. Bolton, A theory of trickle-down growth and development, Rev. Econ. Stud., 64, (1997) 151–172.

[2] R. Amir, Complementarity and diagonal dominance in discounted stochastic games, Ann. Oper. Res. 114, (2002) 39–56.

[3] R. Amir, Discounted supermodular stochastic games: theory and applications, mimeo, (2005) University of Arizona.

[4] A. R. Antunes, T. V. Cavalcanti, Start up costs, limited enforcement, and the hidden economy, Europ. Econ. Rev., 51 (2007) 203–224.

[5] A. R. Antunes, T. V. Cavalcanti, A. Villamil, Computing general equilibrium models with occupational choice and financial frictions, J. Math. Econ. 44 (2008) 553–568.

[6] L. Balbus, K. Reffett, L. Woźny, Stationary Markovian equilibrium in altruistic stochastic OLG models with limited commitment, mimeo (2010) ASU Department of Economics.

[7] R. N. Bhattacharya, O. Lee, Asymptotics of a class of Markov processes which are not in general irreducible, Ann. Probab., 16 (3) (1988) 1333–1347.

[8] R. N. Bhattacharya, M. Majumdar, On a class of stable random dynamical systems: theory and applications, J. Econ. Theory, 96 (2001) 208–229.

[9] R. N. Bhattacharya, M. Majumdar, N. Hashimzade, Limit theorems for monotone Markov processes, Sankhya, 72-A (1) (2010) 170–190.

[10] W. A. Brock, L. J. Mirman, Optimal economic growth and uncertainty: the discounted case, J. Econ. Theory, 4 (1972) 479–513.

[11] A. Cabrales, H. A. Hopenhayn, Labor-market flexibility and aggregate employment volatility, Carnegie-Rochester Conference Series on Public Policy 46 (1997) 189–228.

[12] A. H. Cabrillana, Endogenous capital market imperfections, human capital, and intergenerational mobility, J. Dev. Econ., 90 (2009) 285–298.

[13] B. A. Cardak, Ability, education, and income inequality, J. Pub. Econ. Theory, 6 (2004) 239–276.

[14] P. Chatterjee, M. Shukayev, Note on positive lower bounded of capital in the stochastic growth model, J. Econ. Dynam. Control, 32 (2008) 2137–2147.

[15] P. Chatterjee, M. Shukayev, A stochastic dynamic model of trade and growth: convergence and diversification, mimeo, (2010) National University of Singapore.

[16] T. F. Cooley, V. Quadrini, Financial markets and firm dynamics, Amer. Econ. Rev., 91 (2001) 1286–1310.

[17] K. A. Couch, O. F. Morand, Inequality, mobility, and the transmission of ability, J. Macroecon., 27 (2005) 365–377.

[18] S. P. Das, Trade, skill acquisition and distribution, J. Dev. Econ., (2006) 118–141.

[19] M. Datta, L. J. Mirman, O. F. Morand, K. L. Reffett, Monotone methods for Markovian equilibrium in dynamic economies, Ann. Oper. Res., 114 (2002) 117–144.

[20] W. D. Dechert, K. Nishimura, A complete characterization of optimal growth paths in an aggregated model with non-concave production function, J. Econ. Theory, 31 (1983) 332–354.

[21] P. A. de Hek, On endogenous growth under uncertainty, Int. Econ. Rev., 40 (1999) 727–744.

[22] A. Díaz, J. Pijoan-Mas, J. V. Ríos-Rull, Precautionary Savings and Wealth Distribution under Habit Formation Preferences, J. Monet. Econ., 50 (2003) 1257–1291.

[23] L. E. Dubins, D. A. Freedman, Invariant probabilities for certain Markov processes, Ann. Math. Stat., 37 (1966) 837–848.

[24] R. M. Dudley, Real analysis and probability, Cambridge Studies in Advanced Mathematics (2002) Cambridge University Press.

[25] O. Hernández-Lerma, J. B. Lasserre, Markov chains and invariant probabilities, Birkhäuser Verlag, (2003) Switzerland.

[26] L. Gong, X. Zhao, Y. Yang, Z. Hengfu, Stochastic growth with social-status concern: the existence of a unique stationary distribution, J. Math. Econ., 46 (2010) 505–518.

[27] H. A. Hopenhayn, Entry, exit, and firm dynamics in long run equilibrium, Econometrica, 60 (1992) 1127–1150.

[28] H. A. Hopenhayn, E. C. Prescott, Stochastic monotonicity and stationary distributions for dynamic economies, Econometrica, 60 (1992) 1387–1406.

[29] H. A. Hopenhayn, R. Rogerson, Job turnover and policy evaluation: a general equilibrium analysis, J. Polit. Economy, 101 (1993) 915–938.

[30] M. Huggett, The risk-free rate in heterogeneous agent economies, J. Econ. Dynam. Control, 17 (1993) 953–69.

[31] G. Joseph, T. Weitzenblum, Optimal unemployment insurance: transitional dynamics vs. steady state, Rev. Econ. Dynam., 6 (2003) 869–884.

[32] T. Kam, J. Lee, On a unique nondegenerate distribution of agents in the Huggett model, mimeo, (2011) Australian National University.

[33] T. Kamae, U. Krengel, Stochastic partial ordering, Ann. Probab., 6 (1978) 1044–1069.

[34] T. Kamae, U. Krengel, G. L. O'Brien, Stochastic inequalities on partially ordered spaces, Ann. Probab. 5 (1977) 899–412.

[35] T. Kamihigashi, Stochastic optimal growth with bounded or unbounded utiltity and bounded or unbounded shocks, J. Math. Econ., 43 (2007) 477–500.

[36] T. Kamihigashi, J. Stachurski, An order-theoretic mixing condition for monotone Markov chains, Stat. Probab. Letters, 82 (2012) 262–267.

[37] D. Kristensen, Uniform ergodicity of a class of Markov chains with applications to time series models, mimeo (2008) Columbia University.

[38] S. Lee, O. Linton and Y-J. Whang, Testing for stochastic monotonicity, Econometrica, 77 (2) (2009) 585–602.

[39] H. Lloyd-Ellis, Public education, occupational choice and the growth-inequality relationship, Int. Econ. Rev., 41 (2000) 171–201.

[40] H. Lloyd-Ellis, D. Bernhardt, Enterprise, inequality and economic development, Rev. Econ. Stud., 67 (2000) 147–168.

[41] M. Majumdar, T. Mitra and Y. Nyarko, Dynamic optimization under uncertainty: nonconvex feasible set, in Joan Robinson and Modern Economic Theory, (Ed.) G. R. Feiwel, MacMillan Press, (1989) New York.

[42] A. Marcet, F. Obiols-Homs and P. Weil, Incomplete markets, labor supply and capital accumulation, J. Monet. Econ., 54 (2007) 2621–2635.

[43] K. Matsuyama, Financial market globalization, symmetry-breaking, and endogenous inequality of nations, Econometrica, 72 (2004) 853-884.

[44] S. Meyn, R. L. Tweedie, Markov chains and stochastic stability, 2nd Edition, Cambridge University Press (2009) Cambridge.

[45] L. J. Mirman, L. F. Morand, K. L. Reffett A qualitative approach to Markovian equilibrium in infinite horizon economies with capital, J. Econ. Theory, 139, 75–98.

[46] T. Mitra, S. Roy, Optimal exploitation of renewable resources under uncertainty and the extinction of species, Econ. Theory, 28 (2006) 1-23.

[47] O. F. Morand, K. L. Reffett, Stationary Markovian equilibrium in overlapping generation models with stochastic nonclassical production and Markov shocks, J. Math. Econ., 43 (2007) 501–522.

[48] K. Nishimura, J. Stachurski, Stability of stochastic optimal growth models: a new approach, J. Econ. Theory, 122 (1) (2005) 100–118.

[49] L. J. Olson, Stochastic growth with irreversible investment, J. Econ. Theory, 47, (1989) 101–129.

[50] L. J. Olson, S. Roy, Dynamic efficiency of conservation of renewable resources under uncertainty, J. Econ. Theory, 95 (2000) 186–214.

[51] L. J. Olson, S. Roy, The theory of stochastic optimal economic growth, in Handbook on Optimal Growth 1: Discrete Time, ed. by R-A Dana, C. L. Van, T. Mitra, and K. Nishimura, Springer, (2006) Berlin.

[52] A. L. Owen, D. N. Weil, Intergenerational earnings mobility, inequality and growth, J. Monet. Econ., 41 (1998) 71–104.

[53] J. Pijoan-Mas, Precautionary savings or working long Hours? Rev. Econ. Dynam., 9 (2006) 326–352.

[54] T. Piketty, The dynamics of the wealth distribution and the interest rate with credit rationing, Rev. Econ. Stud., 64 (1997) 173–189.

[55] L. S. V. Portes, On the distributional effects of trade policy: dynamics of household saving and asset prices, Quart. Rev. Econ. Finance, 49 (2009) 944–970.

[56] P. Ranjan, Dynamic evolution of income distribution and credit-constrained human capital investment in open economies, J. Int. Econ., 55 (2001) 329–358.

[57] A. Razin, J. A. Yahav, On stochastic models of economic growth, Int. Econ. Rev., 20 (1979) 599–604.

[58] R. M. Samaniego, Industrial subsidies and technology adoption in general equilibrium, J. Econ. Dynam. Control, 30 (2006) 1589–1614.

[59] R. M. Samaniego, Can technical change exacerbate the effects of labor market sclerosis? J. Econ. Dynam. Control, 32 (2008) 497–528.

[60] M. S. Santos, A. Peralta-Alva, Accuracy of simulations for stochastic dynamic models, Econometrica, 73 (2005) 1939–1976.

[61] J. Stachurski, Stochastic optimal growth with unbounded shock, J. Econ. Theory, 106 (2002) 40–65.

[62] N. L. Stokey, R. E. Lucas, Jr., with E. C. Prescott, Recursive methods in economic dynamics, Harvard University Press, Cambridge, (1989) MA.

[63] W. Whitt, Uniform conditional stochastic order, J. Appl. Probab., 17 (1980) 112–123.

[64] Y. Zhang, Stochastic optimal growth with a non-compact state space, J. Math. Econ., 43, (2007) 115–129.