View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

ISSN 0819-2642
ISBN 0 7340 2555 6

THE UNIVERSITY OF MELBOURNE

DEPARTMENT OF ECONOMICS

RESEARCH PAPER NUMBER 899

FEBRUARY 2004

PARAMETRIC CONTINUITY OF STATIONARY DISTRIBUTIONS

by

Cuong Le Van
&
John Stachurski

Department of Economics
The University of Melbourne
Melbourne Victoria 3010
Australia.

PARAMETRIC CONTINUITY OF STATIONARY DISTRIBUTIONS

CUONG LE VAN AND JOHN STACHURSKI

Abstract. The paper gives conditions under which stationary distributions of Markov models depend continuously on the parameters. It extends a well-known parametric continuity theorem for compact state space to the unbounded setting of standard econometrics and time series analysis. Applications to several theoretical and estimation problems are outlined.

1. Introduction

In macroeconomic dynamics, time series econometrics and other related fields, one frequently considers economies where the sequence of state variables $(X_t)_{t=0}^{\infty}$ is stationary. Here X_t is a vector of endogenous and exogenous variables, jointly following a Markov process generated by some underlying model. By stationary is meant the existence of a "stationary distribution" μ, such that if X_t has law μ, then so does X_{t+j} for all $j \in \mathbb{N}$. If such a μ exists, is unique and has some stability properties, then it naturally becomes the focus of equilibrium analysis. For example, in these settings a law of large numbers result often holds, in which case sample moments from the series $(X_t)_{t=0}^{\infty}$ can be identified with integrals of the relevant functions with respect to the stationary distribution μ.

Typically, the underlying laws which drive the process $(X_t)_{t=0}^{\infty}$ depend on a vector of parameters, which may for example be policy instruments, or regression coefficients to be estimated from the data. In this case the parameters themselves determine the stationary distribution. The study of how this distribution varies with the parameters is a stochastic analogue of standard comparative dynamics. Our paper investigates conditions under which the functional relationship between parameters and stationary distribution is continuous.

The main results extend directly a well-known and useful argument in Stokey, Lucas and Prescott for Markov processes on a compact state space (1989, Theorem 12.13), which is apparently due to R. Manuelli. Although compactness of the state space is indeed convenient where it can be assumed, for many empirical studies it fails. An elementary example is the scalar AR(1) models with normal shocks.

Instead of compactness, various uniform stability and monotonicity conditions are adopted. The conditions are stated in terms of the transition laws and distributions of the shocks, with a view to simple verification.

On related literature, another study of parametric continuity of stationary distributions in the compact case is Santos and Peralta-Alva (2003). They obtain detailed error bounds when the transition rule is uniformly contracting on average, as well as addressing the implications for accuracy of numerical simulations. We do not pursue the compact state case here.

The remainder of the paper is structured as follows. In Section 2 three examples illustrating the importance of parametric continuity are given. In Section 3 the general problem is formulated and results are given. Section 4 concludes with proofs.

2. Examples

2.1. **Simulation and Estimation.** Parametric continuity of stationary distributions is a component of various problems in estimation, simulation, numerical dynamic programming and economic theory. A well-known example is the Simulated Moments Estimator of Duffie and Singleton (1993), which can be described as follows. Let $(X_t)_{t=0}^T$ be observable data—a vector of state variables that are taken as generated by some dynamic general equilibrium model. Let φ be a function on the state space S which associates to current state X_t real number $\varphi(X_t)$. For concreteness, suppose the problem is one of asset pricing, and $\varphi(X_t)$ is the price of a given asset, the value of which depends on the current state. The transition rule for the model which is assumed to generate the data has reduced form

$$X_{t+1} = H_\alpha(X_t, \xi_t), \quad t = 0, 1, \ldots, \tag{1}$$

where $\alpha \in W$, a parameter space, and $(\xi_t)_{t=0}^{\infty}$ is a sequence of independent shocks.

The true parameter α is not known. However, the econometrician observes the sequence $(X_t)_{t=0}^{T}$, can compute H_β for given $\beta \in W$ as the solution to the model, and has access to a sequence of shocks $(\hat{\xi}_t)$ with the same (joint) distribution as (ξ_t). From these one can simulate for each $\beta \in W$ a sequence $(X_t^\beta)_{t=0}^{N}$ defined recursively by

$$X_{t+1}^\beta = H_\beta(X_t^\beta, \hat{\xi}_t), \quad t = 0, 1, \ldots$$

The simulated moments estimator is the value of β which minimizes the distance between the moments of the observed price process $\varphi(X_t)$ and the simulated process $\varphi(X_t^\beta)$.

In order to prove consistency of the estimator, Duffie and Singleton require that the map $\beta \mapsto \mathbb{E}_\beta(\varphi) := \int \varphi \, d\mu_\beta$ is continuous on W, where μ_β is the stationary distribution corresponding to β. Providing sufficient conditions for such continuity to hold in standard econometric settings is precisely the problem with which the present paper is concerned.

2.2. **Numerical Dynamic Programming.** Another example is the accuracy of simulated time series from general equilibrium models solved numerically by dynamic programming. Recently important progress in measuring the accuracy of numerical solutions for these models has been made by Santos and Vigo-Aguiar (1998). The relationship between these bounds and accuracy of simulation is studied in Santos and Peralta-Alva (2003).

The dynamic programming problem is to choose a policy function σ mapping state space S into control space A which maximizes a discounted sum of period utilities, or minimizes discounted loss. The law of motion for the system is given by some rule

$$X_{t+1} = F(X_t, U_t, \xi_t), \quad t = 0, 1, \ldots, \tag{2}$$

where for now (ξ_t) is an i.i.d. sequence and $U_t \in A$ is the current value of the control. Suppose that we obtain numerically a candidate solution $\hat{\sigma}_n$ for the optimal policy σ, where n is some index of dedicated CPU cycles. Suppose further that theoretical bounds are available for the distance $\|\hat{\sigma}_n - \sigma\|$ in terms of n, where $\|\cdot\|$ is the norm in some appropriate function space $\mathscr{F}$. Substituting $\hat{\sigma}_n$ into (2) gives the Markov process

$$X_{t+1}^n = F(X_t^n, \hat{\sigma}_n(X_t^n), \xi_t). \tag{3}$$

Let $T_{\hat{\sigma}_n}(X, \xi) := F(X, \hat{\sigma}_n(X), \xi)$. The point of this notation is that since in the theory below the parameter space W is required only to be a metric space, we can treat the candidate policy $\hat{\sigma}_n$ as a parameter. Set $W := (\mathscr{F}, \|\cdot\|)$, which includes all feasible policies.

The ultimate objective is to simulate time series at the stationary distribution μ under the optimal policy σ. The method is to substitute $\hat{\sigma}_n$ for σ, as in (3), and use the approximating sequence $(X_t^n)_{t=0}^{\infty}$. If the process is ergodic, X_t^n will be nearly distributed according to the stationary distribution μ_n corresponding to (3) when t is large. If $\|\hat{\sigma}_n - \sigma\| \to 0$ as $n \to \infty$ we hope that $\mu_n \to \mu$ in some topology, in which case, by taking both t and n large, X_t^n is approximately distributed according to μ. This problem again reduces to the kind of parametric continuity issue considered in the paper.

2.3. **The Solow-Phelps Stochastic Golden Rule.** Next we consider a more specific example, which is of independent interest, and also helps to illustrate how one might verify the conditions we impose. The example is the Solow-Phelps stochastic golden rule problem for a one-sector economy. The framework is a Solow growth model, to which a random component is added. The problem is to choose a savings rate which maximizes expected utility of consumption at the steady state. Surprisingly the existence of a solution to this problem has not been treated to our knowledge.[1]

Since the savings rate takes values in compact set $[0, 1]$, determining existence of a solution reduces in effect to proving continuity of the map from savings rate to stationary distribution.

In more detail, suppose initially a fixed savings rate s from current income. At each period t production takes place, deriving from capital stock k_t random output $y_t = f(k_t)\,\xi_t$; where $f\colon S \to S$ is the production function, $S := (0, \infty)$, and ξ_t is a productivity shock with distribution ν on S. The parameter is the savings rate, and we take the parameter space W to be $(0, 1]$. Here $s = 0$ is excluded so as to eliminate trivialities.

For simplicity, depreciation is taken to be total. Given savings rate s, next period capital is

$$k_{t+1} = s \cdot f(k_t)\,\xi_t. \tag{4}$$

The process then repeats, generating Markov chain $(k_t)_{t=0}^{\infty}$ starting from $k_0 \in S$ given. In order that this process have a stationary distribution, some restrictions on f and ν are necessary:

[1] Schenk-Hoppé (2002) considers maximization of expected consumption in a very general stochastic environment. This is a direct generalization of the deterministic problem. However in a stochastic setting maximization of expected *utility* of consumption is perhaps the more natural criterion.

Assumption 2.1. The function f is continuously differentiable on S, increasing, concave, and satisfies $f(k) = 0 \iff k = 0$ and the Inada conditions $\lim_{k \downarrow 0} f'(k) = \infty$, $\lim_{k \uparrow \infty} f'(k) = 0$.

Assumption 2.2. Both $\mathbb{E}[\xi_t]$ and $\mathbb{E}[1/\xi_t]$ are finite. The Lebesgue measure is absolutely continuous with respect to ν (non-Lebesgue-null sets have positive ν-measure).

The restrictions on f are presented in the current form because they are so familiar, but a quick reading of the proofs shows they are much stronger than what is actually required. The assumptions $\mathbb{E}[\xi_t] < \infty$ and $\mathbb{E}[1/\xi_t] < \infty$ restrict the right and left-hand tails of the shock respectively.[2] The assumption that ν is positive on sets of positive Lebesgue measure—which is used to prove uniqueness—is also much stronger than required, but simplifies the analysis, and holds for many standard econometric shocks. For example, all lognormal distributions satisfy the three restrictions in Assumption 2.2.

The golden rule problem is to maximize expected utility of consumption at the steady state. Let $u\colon [0, \infty) \to \mathbb{R}$ be a utility function.

Assumption 2.3. The function u is continuous, bounded, strictly increasing and, by adding or subtracting a constant if necessary, satisfies $u(0) = 0$.

Suppose that for each savings rate s, (4) has a unique stationary distribution μ_s (for a precise definition of stationary distribution see Definition 3.2 below). Consumption c_t is just $(1 - s)f(k_t)\xi_t$. Recalling that k_t and ξ_t are independent, the distribution of c_t in equilibrium is

$$\varphi_s(B) := \text{Prob}\{c_t \in B\} = \int \int \mathbf{1}_B[(1 - s)f(k)z]\nu(dz)\mu_s(dk), \tag{5}$$

[2]With unbounded shocks some restrictions of this nature are necessary for the stationary distribution to exist—the first restriction prevents unbounded growth; the second collapse to the origin.

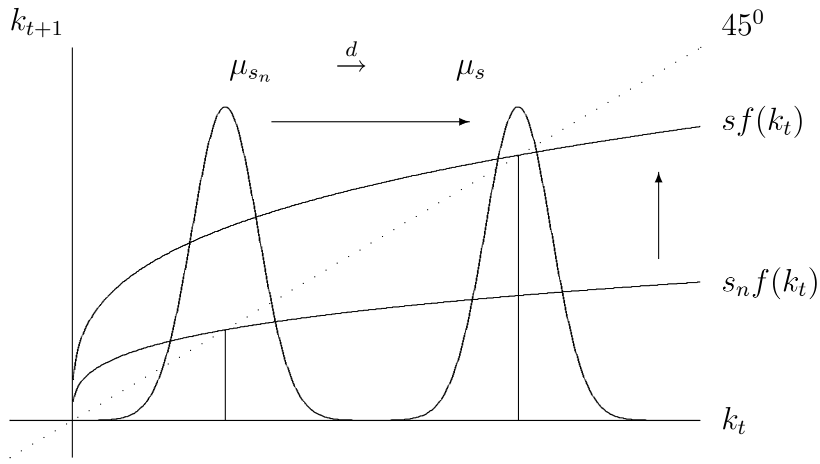

FIGURE 1. Convergence of stationary distributions.

where $\mathbf{1}_B[(1-s)f(k)z] = 1$ when $(1-s)f(k)z \in B$ and zero otherwise. The golden rule problem is to select the "lottery" φ_s which maximizes expected utility:

$$\max_{0 \le s \le 1} \mathbb{E}_s u(c), \quad \mathbb{E}_s u(c) := \int u \, d\varphi_s. \tag{6}$$

We now have the following result. All definitions are made precise in Section 3. The proofs are in Section 4. Part 1 of Theorem 2.1 is illustrated in Figure 1.

Theorem 2.1. *Regarding the Solow–Phelps golden rule problem, let Assumptions 2.1, 2.2 and 2.3 hold. The following statements are true.*

1. *For each savings rate s, the process* (4) *has a unique stationary distribution, which depends continuously on s.*
2. *Moreover, an interior solution s^* to the stochastic Phelps golden rule problem* (6) *exists, at which the expected utility of long run equilibrium consumption is maximized.*

3. THE GENERAL MODEL

In what follows, for measurable space $(E, \mathscr{E})$ let $\mathcal{P}(E)$ denote the probabilistic measures on $(E, \mathscr{E})$, and $b\mathscr{M}(S)$ the finite signed measures.

Functional notation is often used for integration: if $h\colon E \to \mathbb{R}$ is $\mathscr{E}$-measurable and $\mu \in b\mathscr{M}(E)$ then $\mu(h) := \int h d\mu$ when the latter is defined. If E has a topology, the σ-algebra $\mathscr{E}$ is always the Borel sets. Throughout the paper we use the convention that if X is a random variable taking values in measurable space $(E, \mathscr{E})$, then $\mathcal{L}(X)$ is the law (distribution) of X, an element of $\mathcal{P}(E)$.

Let E have a topology, and let $C_b(E)$ be the continuous bounded functions on E. To each $h \in C_b(E)$ there corresponds the linear functional on $b\mathscr{M}(E)$ defined by $\mu \mapsto \mu(h) \in \mathbb{R}$. The weak topology $w(b\mathscr{M}(E), C_b(E))$ on $b\mathscr{M}(E)$ is the smallest topology which makes all such functionals continuous. Also, $w(\mathcal{P}(E), C_b(E))$ is the weak topology on $\mathcal{P}(E)$ (i.e., the relative topology inhereted from $b\mathscr{M}(E)$). For a collection of probabilities $\{\mu_n\}_{n=1}^{\infty} \cup \{\mu\} \subset \mathcal{P}(E)$, we write $\mu_n \xrightarrow{d} \mu$ when the sequence (μ_n) converges to μ in $w(\mathcal{P}(E), C_b(E))$—that is when $\mu_n(h) \to \mu(h)$ in $\mathbb{R}$ for every $h \in C_b(E)$. We say that (μ_n) converges to μ in distribution (Stokey, Lucas and Prescott, 1989, Chapter 12).

3.1. **The Model.** Now let $\alpha \in W$ be a parameter. Consider an economy that evolves according to the rule

$$X_{t+1} = T_\alpha(X_t, \xi_t), \quad X_0 \equiv x_0 \in S \text{ given.} \tag{7}$$

The sequence of shocks $(\xi_t)_{t=0}^{\infty}$ is assumed to be uncorrelated and identically distributed.[3] Together, the transition rule T_α and the distribution of the shock—call it ν—determine a Markov chain $(X_t)_{t=0}^{\infty}$. The timing is that X_t is observed at the start of time t, planning of economic activity takes place on the basis of this information, and is then implemented through the period. Implementation is perturbed by a shock ξ_t, giving rise to X_{t+1} at the start of the next period via (7). It is clear from the timing that X_t and ξ_t are independent.

[3]As is well-known, correlations of finite order can be treated in this framework by redefining state variables and, if necessary, adjusting the dimensions of the state space.

The random variable X_t evolves in state space S. The state space has a topology, and $\mathscr{B}(S)$ denotes the Borel sets of S. Also, let $(Z, \mathscr{Z})$ be any measurable space, with the shocks that perturb the economy in each period taking values in Z, so that $T_\alpha \colon S \times Z \to S$.[4] The parameter space W is any metric space.

3.2. **Basic Assumptions.** Some basic requirements are now given which will at least assure the existence and uniqueness of stationary distributions.

Assumption 3.1. The space S is separable and completely metrizable.

Such spaces are called Polish. For example, in the golden rule problem the state space is $S = (0, \infty)$ with the usual topology. This space is Polish. In fact every open (indeed, every G_δ) subset of a complete metric space is completely metrizable by Alexandrov's Lemma (c.f., e.g., Aliprantis and Border, 1999, Theorem 3.22).

Assumption 3.2. For each fixed $\alpha \in W$, $x \mapsto \int h[T_\alpha(x, z)]\nu(dz)$ is continuous and bounded whenever $h \colon S \to \mathbb{R}$ is.

While it may not immediately appear so, this assumption is easy to verify in applications. For example, it holds when $x \mapsto T_\alpha(x, z)$ is continuous for each fixed α and z, as follows from Dominated Convergence. The continuity provided by Assumption 3.2 is paired with a compactness requirement below to obtain existence of stationary distributions via a fixed point argument. To define the compactness requirement, first we introduce norm-like functions, which play a role somewhat analogous to Liapunov functions in dynamical systems theory.

Definition 3.1. A continuous function $V \colon S \to [0, \infty)$ is called norm-like if there is an increasing sequence of compact sets $(K_j)_{j=1}^\infty$ with $\cup_{j=1}^\infty K_j = S$ and $\inf_{x \notin K_j} V(x) \to \infty$ as $j \to \infty$.

[4] Of course T_α is required to be $(\mathscr{B}(S) \otimes \mathscr{Z}, \mathscr{B}(S))$-measurable, $\forall\, \alpha \in W$.

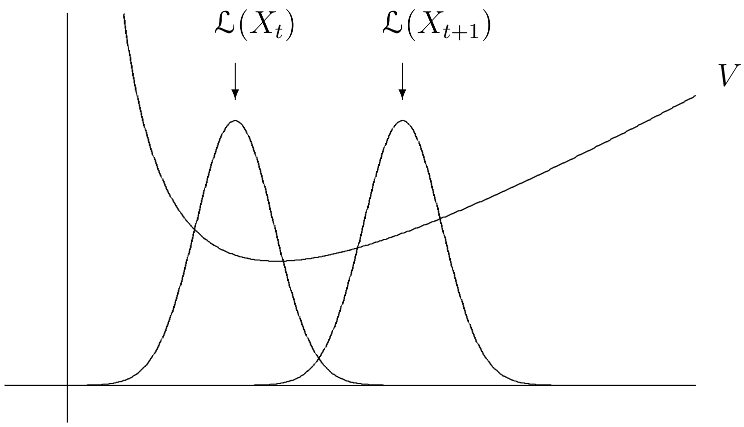

FIGURE 2. A norm-like function on $(0,\infty)$.

A norm-like function V gets larger and larger at the "edges" of the state space, so we can be sure that $\mathcal{L}(X_t)$ stays near the "center" of the space by checking that the expectation of $V(X_t)$ is bounded in t (Figure 2). The next condition will guarantee precisely this. It means that there is on average a drift of the state variable towards parts of the state space on which V is remains small.

Assumption 3.3. For each $\alpha \in W$, there exists a norm-like function V and constants $\lambda, b \in [0,\infty)$ with $\lambda < 1$ and

$$\int V[T_\alpha(x,z)]\nu(dz) \leq \lambda V(x) + b, \quad \forall x \in S. \tag{8}$$

Finally, the question of parametric continuity of stationary distributions is most interesting when this distribution is unique (for each parameterization). Fortunately, uniqueness is far more common in stochastic models than in deterministic ones, as a result of the mixing provided by noise. In order to generate uniqueness we assume throughout the paper the following, although any other uniqueness condition will do.

Assumption 3.4. For each $\alpha \in W$, the process (7) is ψ-irreducible. In other words, there is a $\psi \in \mathcal{P}(S)$, possibly depending on the parameter

α, such that $\psi(B) > 0$ implies

$$\text{Prob}_x\{X_t \in B \text{ for some } t \in \mathbb{N}\} > 0, \quad \forall x \in S,$$

where Prob_x is the distribution of $(X_t)_{t=0}^{\infty}$ when $X_0 \equiv x$.

Again, this assumption is often easier to verify in practice than it looks. See the proof for the golden rule example below.

3.3. **Stationary Distributions.** Conditions for the existence of stationary distributions for Markov processes on uncountable state space have been investigated by many authors. In economics, a well known result is that Feller processes on compact state space have at least one stationary distribution (Stokey, Lucas and Prescott, 1989, Theorem 12.10). The compactness requirement can be weakened to the condition that at least one "trajectory" of the process (7) be precompact. Intuitively, if this is the case then up to an epsilon, probability mass stays in a compact set at least for one starting point—by Prohorov's theorem, see below—and the compact state proof can be generalized appropriately. The details follow.

The long run equilibrium of the economy (7) is identified—when it exists—with the stationary distribution μ_α of the Markov chain $(X_t) = (X_t^\alpha)$.

Definition 3.2. Distribution $\mu_\alpha \in \mathcal{P}(S)$ is called stationary for (7) if

$$\mu_\alpha(B) = \int_S \left[\int_Z \mathbf{1}_B[T_\alpha(x,z)]\nu(dz) \right] \mu_\alpha(dx), \quad \forall B \in \mathscr{B}(S). \tag{9}$$

Here $\mathbf{1}_B[T_\alpha(x,z)] = 1$ if $T_\alpha(x,z) \in B$ and zero otherwise, so the inner integral has the interpretation $\text{Prob}\{X_{t+1} \in B \mid X_t = x\}$. Equation (9) states that if $\mathcal{L}(X_t) = \mu_\alpha$, then so does $\mathcal{L}(X_{t+1})$.

To clarify further, these ideas can be recast in a more general form.[5] Define for each $\alpha \in W$ a map $\mathbf{P}_\alpha \colon \mathcal{P}(S) \to \mathcal{P}(S)$ by

$$(10) \qquad \mathbf{P}_\alpha \mu(B) = \int_S \left[\int_Z \mathbf{1}_B[T_\alpha(x,z)]\nu(dz) \right] \mu(dx), \quad B \in \mathscr{B}(S).$$

The operator $\mathbf{P}_\alpha$ is called the Markov (or Foias) operator associated with (7).[6] It is well-known that if $\mathcal{L}(X_t) = \mu$, then $\mathcal{L}(X_{t+1}) = \mathbf{P}_\alpha \mu$ (c.f., e.g. Stokey, Lucas and Prescott, 1989, p. 219, or Lasota and Mackey, 1994, p. 414). In particular,

$$\text{if } \mathcal{L}(X_0) = \delta_x, \text{ then } \mathcal{L}(X_t) = \mathbf{P}_\alpha^t \delta_x,$$

where $\delta_x \in \mathcal{P}(S)$ is the distribution concentrated at x, and the superscript t means t iterations of $\mathbf{P}_\alpha$. Comparing (10) with (9), clearly $\mu_\alpha \in \mathcal{P}(S)$ is a stationary distribution of (7) if and only if it is a fixed point of $\mathbf{P}_\alpha$.

The Markov operator (10) can be thought of as a map not just on $\mathcal{P}(S)$, but also from $b\mathscr{M}(S)$ to itself. It satisfies $\mathbf{P}_\alpha \mathcal{P}(S) \subset \mathcal{P}(S)$. More generally, let us define a Markov operator $\mathbf{P}$ on $(S, \mathscr{B}(S))$ to be any linear map from $b\mathscr{M}(S)$ into itself such that $\mathbf{P}\mathcal{P}(S) \subset \mathcal{P}(S)$.

Definition 3.3. A Markov operator $\mathbf{P}$ on $(S, \mathscr{B}(S))$ is said to have the Feller property if it is $w(b\mathscr{M}(S), C_b(S))$-continuous.

We now give the main existence result, which can be thought of as a direct generalization of the well-known compact state result given in Stokey, Lucas and Prescott (1989, Theorem 12.10). The result is not essentially new. It can be deduced in a similar setting from Meyn and Tweedie (1993, Theorem 12.1.2, Proposition 12.1.3 and Lemma D.5.3). We provide a simple direct proof.

[5] Most of the following is standard in economics. See for example Futia (1982), or Stokey, Lucas and Prescott (1989).

[6] In some literatures it is the adjoint of $\mathbf{P}_\alpha$ which is called the Markov operator. Our terminology and notation closely follows Lasota and Mackey (1994).

Theorem 3.1. *Let* $\mathbf{P}$ *be a Markov operator on* $(S, \mathscr{B}(S))$*. If* $\mathbf{P}$ *is Feller and there is an* $x \in S$ *such that* $\{\mathbf{P}^t \delta_x : t \in \mathbb{N}\}$ *is precompact, then* $\mathbf{P}$ *has a fixed point* μ *in* $\mathcal{P}(S)$*.*[7]

Proof. Let A be the closure of the convex hull of $\{\mathbf{P}^t \delta_x : t \in \mathbb{N}\}$. From the fact that $\mathbf{P}$ is invariant on $\{\mathbf{P}^t \delta_x : t \in \mathbb{N}\}$, Feller and linear, it is easy to see that $\mathbf{P}$ maps A into A. By the Schauder fixed point theorem, $\mathbf{P}$ has a fixed point $\mu \in A \subset \mathcal{P}(S)$.[8] □

We now explore the properties of the Markov operator generated by (7) under Assumptions 3.1–3.4.

Lemma 3.1. *Let* $\alpha \in W$*, and let* $\mathbf{P}_\alpha$ *be defined by* (10)*. If Assumption 3.2 holds, then* $\mathbf{P}_\alpha$ *has the Feller property.*

This result is elementary and well-known (c.f., e.g., Stokey, Lucas and Prescott, 1989, p. 376).

Lemma 3.2. *Let* $\alpha \in W$*, and let* $\mathbf{P}_\alpha$ *be defined by* (10)*. If Assumptions 3.1 and 3.3 hold, then* $\{\mathbf{P}_\alpha^t \delta_x : t \in \mathbb{N}\}$ *is precompact for every* $x \in S$*.*

The proof is given in Section 4. The intuition is that from any $X_0 \equiv x$, Assumption 3.3 bounds the growth of the expectation of $V(X_t)$, which—from the definition of norm-like functions—means that up to an epsilon, $\mathcal{L}(X_t)$ is concentrated on a compact set for all t. This property is called tightness, which is equivalent to precompactness in the Polish setting.

[7] In the statement, precompact means having compact closure. The topology is $w(b\mathscr{M}(S), C_b(S))$.

[8] To be more precise, $\mathcal{P}(S)$ is a subset of $b\mathscr{M}(S)$, the space of finite signed Borel measures on S with the $w(b\mathscr{M}(S), C_b(S))$ topology. This space is locally convex Hausdorff—here we need the Polish assumption—so Schauder's fixed point theorem applies (c.f., e.g. Aliprantis and Border, 1999, Theorem 16.52). That the closed convex hull of a precompact set is compact in this setting follows from the same reference, Theorem 5.20 and the remark thereafter.

We now collect these results and combine them with Assumption 3.4 to obtain existence and uniqueness.

Theorem 3.2. *If Assumptions 3.1–3.4 hold, then the Markov chain* (7) *has a unique stationary distribution* μ_α *for each* $\alpha \in W$.

Proof. Existence follows from Theorem 3.1 and Lemmas 3.1 and 3.2. Uniqueness under Assumption 3.4 is well-known. See, for example, Meyn and Tweedie (1993, Theorem 10.0.1 and Proposition 10.1.1). □

3.4. **Parametric Continuity.** We can now progress to the main results of the paper, concerning parametric continuity of stationary distributions. First, though, in order to have any hope of getting parametric continuity in the stationary distributions, we at least need continuity in the primitives:

Assumption 3.5. The transition rule T_α is continuous in α. That is, $\alpha \mapsto T_\alpha(x, z)$ is continuous for each $x \in S$ and $z \in Z$.

Two different parametric continuity results will be presented. The first requires two additional conditions. Specifically, we strengthen Assumptions 3.2 and 3.3 to hold in a uniform way over the parameters. In the statement of the theorem, ϱ is any metric compatible with the topology on S.

Theorem 3.3. *Let Assumptions 3.1–3.5 hold, and let* $\{\alpha_n\}_{n=1}^\infty \cup \{\alpha\} \subset W$. *Suppose that for each compact* $K \subset S$, *there is a constant* $M < \infty$ *which can be chosen independent of* α_n *to satisfy*

$$\int \varrho[T_{\alpha_n}(x, z), T_{\alpha_n}(x', z)]\nu(dz) \leq M\varrho(x, x') \quad \text{whenever } x, x' \in K. \tag{11}$$

Suppose further that V, λ *and* b *in Assumption 3.3 can be chosen independent of* α_n. *Let* μ_{α_n} *and* μ_α *in* $\mathcal{P}(S)$ *be the unique stationary distributions of* (7) *for each* α_n *and* α *respectively. In this case, if* $\alpha_n \to \alpha$, *then* $\mu_{\alpha_n} \xrightarrow{d} \mu_\alpha$.

The proof is in Section 4. The condition (11) requires that the model (7) satisfy a local Lipschitz condition "on average," where the Lipschitz constant can be chosen independent of the points in the sequence α_n. For an example of how to check this and the other hypotheses of Theorem 3.3, see the proof of Theorem 2.1 (the golden rule problem) in Section 4.

The second result (Theorem 3.4, below) depends on various order properties, which replace the uniform continuity and uniform boundedness assumptions in Theorem 3.3. In particular, the set of stationary distributions may form a (stochastic dominance) order interval in $\mathcal{P}(S)$. These order intervals are compact in the topology of weak convergence for certain spaces S. The compactness of this set can substitute in some sense for the compactness of the state space assumed in earlier results.

A number of order concepts are introduced.[9] Recall that a relation $\leq_E$ on a space E is called an order if it is reflexive, transitive and antisymmetric. If $(X, \leq_X)$ and $(E, \leq_E)$ are two ordered spaces, a map $T \colon X \to E$ is called increasing if $Tx \leq_E Tx'$ whenever $x \leq_X x'$, decreasing if $Tx' \leq_E Tx$ whenever $x \leq_X x'$, and monotone if it is either increasing or decreasing. A subset of $(E, \leq_E)$ is called increasing (decreasing) if its indicator function is an increasing (a decreasing) function (into $\mathbb{R}$ with the usual order). Given $A \subset E$, $i(A)$ is the smallest increasing set containing A, and $d(A)$ is the smallest decreasing set. If $(E, \mathscr{E})$ is a measure space with order $\leq_E$, then $ib\mathscr{E}$ denotes the increasing bounded $\mathscr{E}$-measurable real functions on E.

When E has a topology, $\leq_E$ is called separating if, whenever $x \nleq_E y$, there is an increasing continuous bounded $h \colon E \to \mathbb{R}$ with $h(x) > h(y)$. For example, $\mathbb{R}^n$ with the usual topology and order is separating.

[9]For more details we refer to the excellent survey of Torres (1990).

If E is metrizable with separating order $\leq_E$, then $\leq_E$ induces on $\mathcal{P}(E)$ the order $\leq_{\mathcal{P}(E)}$ of stochastic dominance.[10] It is defined by

$$\mu \leq_{\mathcal{P}(E)} \mu' \text{ iff } \mu(h) \leq \mu'(h), \ \forall h \in ib\mathscr{E}, \qquad \mathscr{E} \text{ the Borel sets.}$$

In what follows let the state space S and the parameter space W be given orders $\leq_S$ and $\leq_W$ respectively. We require some restrictions on the order $\leq_S$ and its interaction with the topology of S:

Assumption 3.6. The order $\leq_S$ on S is separating. In addition, if K is a compact subset of S, then $i(K) \cap d(K)$ is again compact in S.[11]

For example, Assumption 3.6 holds on the space $\mathbb{R}^n$ with the usual order and topology.

We have the following preliminary result, which extends Hopenhahn and Prescott (1992, Corollary 3) to a noncompact setting. The proof uses a result of Huggett (2003) on monotone comparative dynamics, a convergence result of Meyn and Tweedie (1993), and the fact that separating orders are closed.

Proposition 3.1. *Let Assumptions 3.1—3.6 hold. Suppose further that for fixed $\alpha \in W$ and $z \in Z$, $x \mapsto T_\alpha(x, z)$ is increasing, and for fixed $x \in S$ and $z \in Z$, $\alpha \mapsto T_\alpha(x, z)$ is increasing. Let $\alpha, \beta \in W$, and let μ_α and μ_β be the corresponding stationary distributions. In this case, if $\alpha \leq_W \beta$, then $\mu_\alpha \leq_{\mathcal{P}(S)} \mu_\beta$.*

In other words, the stationary distribution is (stochastic dominance)-increasing in the parameter. The second parametric continuity result is as follows.

[10] Without some restrictions on the topology and order $\leq_E$, the induced relation $\leq_{\mathcal{P}(E)}$ may not be antisymmetric. (See Torres, 1990, Theorem 5.3.)

[11] This last condition is used in Torres (1990, p 33) to generate $w(\mathcal{P}(S), C_b(S))$-compact $\leq_{\mathcal{P}(S)}$-order intervals in $\mathcal{P}(S)$.

Theorem 3.4. *Let the hypotheses of Proposition 3.1 hold, and let* $\{\alpha_n\}_{n=1}^{\infty} \cup \{\alpha\} \subset W$*. Let* μ_{α_n} *and* μ_α *in* $\mathcal{P}(S)$ *be the unique stationary distributions of* (7) *for each* α_n *and* α *respectively. In this case, if* $\alpha_n \to \alpha$ *monotonically in* W*, then* $\mu_{\alpha_n} \xrightarrow{d} \mu_\alpha$*.*

The conditions (i) $x \mapsto T_\alpha(x,z)$ increasing for fixed $\alpha \in W$ and $z \in Z$, and (ii) $\alpha \mapsto T_\alpha(x,z)$ increasing for fixed $x \in S$ and $z \in Z$ hold for many economic models. For example, the transition rule for the stochastic optimal growth model is well-known to be increasing in the state variable (c.f., e.g., Hopenhayn and Prescott, 1992). Also, the transion rule is increasing in the discount factor (a parameter) for each value of the state and shock (c.f., e.g., Danthine and Donaldson, 1981).

Corollary 3.1. *If, in addition,* $W \subset \mathbb{R}$*, then the conclusion holds even when the sequence* $(\alpha_n)_{n=1}^{\infty}$ *is not monotone.*

Proof. Write μ_n for μ_{α_n} and μ for μ_α. Since $\xrightarrow{d}$ is convergence in a topology, it is sufficient to show that every subsequence $(\mu_{n(k)})_{k=1}^{\infty}$ of μ_n has a subsubsequence $(\mu_{n(k(j))})_{j=1}^{\infty}$ converging to μ. So pick any $(\mu_{n(k)})_{k=1}^{\infty}$. Since every sequence in $\mathbb{R}$ has a monotone subsequence, and since every subsequence of $(\alpha_n)_{n=1}^{\infty}$ converges to α, there is a subsequence of $(\alpha_{n(k)})_{k=1}^{\infty}$—call it $(\alpha_{n(k(j))})_{j=1}^{\infty}$—which converges monotonically to α. In which case $\mu_{n(k(j))} \xrightarrow{d} \mu$ by Theorem 3.4, completing the proof. □

4. Proofs

4.1. **The Golden Rule Problem.** First we give the proof of Theorem 2.1 based on the parametric continuity results in Section 3. The proof is based on Theorem 3.3. (Alternatively, the same results can be established through Corollary 3.1.) The hypotheses of Theorem 3.3 are now progressively verified.

That Assumption 3.1 holds for $S = (0, \infty)$ was verified after the statement of the assumption. Assumption 3.2 is immediate from the Dominated Convergence Theorem and continuity of f: For all $s \in W :=$

$(0,1]$ and $h \in C_b(S)$,

$$\int h[sf(k_n)z]\nu(dz) \to \int h[sf(k)z]\nu(dz) \text{ whenever } k_n \to k \in S.$$

Regarding Assumption 3.3, pick any $s \in W$. On $S = (0, \infty)$ let $V(k) := k + k^{-1}$. Clearly V is norm-like on S. By Assumption 2.2, we can choose a real number γ strictly larger than $\mathbb{E}(1/\xi)$. By the Inada conditions and differentiability there is a $\delta > 0$ such that $sf(k) \geq \gamma k$ on $(0, \delta)$, and then a $c > 0$ such that $sf(k) \geq c$ on $[\delta, \infty)$ by monotonicity, in which case

$$\frac{1}{sf(k)} \leq \frac{1}{\gamma k} + \frac{1}{c} \text{ for all } k \text{ in } (0, \infty).$$

In addition, by the diminishing returns Inada condition, concavity and the finite mean of the shock, there is an $a_1 < 1$ and $b_1 < \infty$ such that $sf(k)\mathbb{E}\xi \leq a_1 k + b_1$, $\forall k$. From these bounds we get

$$\begin{aligned} \int V[sf(k)z]\nu(dz) &= \int sf(k)z\nu(dz) + \int \frac{1}{sf(k)z}\nu(dz) \\ &\leq a_1 k + b_1 + \frac{\mathbb{E}(1/\xi)}{\gamma k} + \frac{\mathbb{E}(1/\xi)}{c}. \end{aligned}$$

Setting $\lambda := a_1 \vee [\mathbb{E}(1/\xi)/\gamma] < 1$ and $b := b_1 + [\mathbb{E}(1/\xi)/c]$ gives (8).

Regarding Assumption 3.4, fix $s \in (0,1] =: W$ and let ψ be any probability on $(0, \infty)$ with a density. Now take any $B \subset (0, \infty)$ with positive ψ-measure and any $k \in (0, \infty) =: S$. It is easy to check that the set $B/sf(k)$ has positive Lebesgue measure, so that from Assumption 2.2 we have

$$\begin{aligned} \text{Prob}_k\{k_t \in B \text{ for some } t \in \mathbb{N}\} &\geq \text{Prob}_k\{k_1 \in B\} \\ &= \nu\{z : sf(k)z \in B\} = \nu[B/sf(k)] > 0. \end{aligned}$$

Here Prob_k is the distribution of $(k_t)_{t=0}^{\infty}$ when $k_0 \equiv k$.

From Theorem 3.1 we now already know that $(k_t)_{t=0}^{\infty}$ has a unique stationary distribution. For Theorem 3.3 to hold, however, there are some additional conditions. Let $\{s_n\}_{n=1}^{\infty} \cup \{s\} \subset W$. First, pick any compact $K \subset S = (0, \infty)$. Since f is continuously differentiable, there is a number A such that $|f(k) - f(k')| \leq A|k - k'|$ for all $k, k' \in K$. But then

$$\begin{aligned}\int |s_n f(k)z - s_n f(k')z|\nu(dz) &\leq \int z\nu(dz)|f(k) - f(k')| \\ &\leq \int z\nu(dz)A|k - k'|.\end{aligned}$$

Thus $M := \int z\nu(dz)A$ verifies the bound in (11) independent of the value s_n.

To finish checking the hypotheses of Theorem 3.3, it remains only to verify that V, λ and b in Assumption 3.3 can be chosen independent of s_n. Let $V(k) = 1/k + k$ as before. By restricting attention to the tail of the sequence (s_n) if necessary we can assume that $r := \inf_n s_n > 0$, because $s_n \to s \in W = (0, 1]$. As above, we can choose a real number γ strictly larger than $\mathbb{E}(1/\xi)$. By the Inada conditions and differentiability there is a $\delta > 0$ such that $rf(k) \geq \gamma k$ on $(0, \delta)$, and then a $c > 0$ such that $rf(k) \geq c$ on $[\delta, \infty)$ by monotonicity, in which case

$$\frac{1}{s_n f(k)} \leq \frac{1}{rf(k)} \leq \frac{1}{\gamma k} + \frac{1}{c} \qquad \forall\, n \in \mathbb{N},\ \forall\, k \in (0, \infty).$$

In addition, by the diminishing returns Inada condition, concavity and the finite mean of the shock, there is an $a_1 < 1$ and $b_1 < \infty$ such that $s_n f(k)\mathbb{E}\xi \leq f(k)\mathbb{E}\xi \leq a_1 k + b_1$, $\forall k$. From these bounds we get

$$\begin{aligned}\int V(s_n f(k)z)\nu(dz) &= \int s_n f(k)z\nu(dz) + \int \frac{1}{s_n f(k)z}\nu(dz) \\ &\leq a_1 k + b_1 + \frac{\mathbb{E}(1/\xi)}{\gamma k} + \frac{\mathbb{E}(1/\xi)}{c}.\end{aligned}$$

Setting $\lambda := a_1 \vee [\mathbb{E}(1/\xi)/\gamma] < 1$ and $b := b_1 + [\mathbb{E}(1/\xi)/c]$ gives (8).

We have now verified all the hypotheses of Theorem 3.3, and hence established Part 1 of Theorem 2.1.

Regarding Part 2, we first establish continuity of the objective function (6). Take any $s \in W$. Let $\mu(s)$ be the corresponding stationary distribution for capital. Let $k(s)$ denote a random variable with this distribution. Steady state consumption given s is

$$c(s) = (1-s)f(k(s))\xi. \tag{12}$$

Now take any $s_n \to s$. It follows from (12), the continuity of f and the Slutsky's Theorem (Dudley, 2002, Theorem 11.7.1) that if $\mathcal{L}(k(s_n)) \stackrel{d}{\to} \mathcal{L}(k(s))$, then $\mathcal{L}(c(s_n)) \stackrel{d}{\to} \mathcal{L}(c(s))$. Further, if $\mathcal{L}(c(s_n)) \stackrel{d}{\to} \mathcal{L}(c(s))$, then $\mathbb{E}u[c(s_n)] \to \mathbb{E}u[c(s)]$ by the definition of weak convergence and the assumption that u is continuous and bounded. This will complete the proof of continuity of the objective function.

That $\mathcal{L}(k(s_n)) \stackrel{d}{\to} \mathcal{L}(k(s))$ when $s \in W = (0,1]$ was proved in Part 1. It only remains then to check that $\mathcal{L}(k(s_n)) \stackrel{d}{\to} \mathcal{L}(k(0)) = \delta_0$ when $s_n \to 0$.

So take such a sequence $(s_n) \subset (0,1]$. Write μ_n for $\mu_{s_n} := \mathcal{L}(k(s_n))$, which we now regard as measures on $[0,\infty)$ putting zero mass on $\{0\}$. By the Portmanteau Theorem (Aliprantis and Border, 1999, Theorem 14.2) $\mu_n \stackrel{d}{\to} \delta_0$ if and only if $\liminf_{n\to\infty} \mu_n(G) \geq \delta_0(G)$ for every relatively open set $G \subset [0,\infty)$. Evidently this is equivalent to $\lim_n \mu_n(G) = 1$ for all open G containing 0, which in turn is equivalent to

$$\lim_{n\to\infty} \mu_n([a,\infty)) = 0, \quad \forall a > 0. \tag{13}$$

In other words, the probability that in the steady state the capital stock exceeds a converges to zero when savings converges to zero.

So pick any $a > 0$. From (9),

$$\begin{aligned}\mu_n([a,\infty)) &= \int \left[\int \mathbf{1}_{[a,\infty)}[s_n f(k) z] \nu(dz) \right] \mu_n(dk) \\ &= \int \nu([a/(s_n f(k)), \infty)) \mu_n(dk).\end{aligned}$$

Without loss of generality, $s_n \leq s_1$ for all n. In that case $\mu_n \leq_{\mathcal{P}(S)} \mu_1$ by Proposition 3.1. Since $k \mapsto \nu([a/(s_n f(k)), \infty))$ is increasing and bounded, by the definition of stochastic dominance

$$\mu_n([a,\infty)) \leq \int \nu([a/(s_n f(k)),\infty))\mu_1(dk). \tag{14}$$

In addition, $\lim_{n\to\infty} \nu([a/(s_n f(k)),\infty)) = 0$ by Dominated Convergence, so applying Dominated Convergence to (14) gives (13) as required.

We conclude that the objective function (6) is continuous for $s \in [0,1]$, a compact set. Therefore a solution to the golden rule problem exists. That it is interior is immediate from (12), since u is strictly positive on $(0,\infty)$ and integrals of strictly positive functions are strictly positive. (Therefore $s = 0$ and $s = 1$ are not solutions, as in both cases the value of the objective function is zero.)

4.2. **Remaining Proofs.** We first give some general discussion.

It is known that when S is Polish, $\mathcal{P}(S)$ with topology $w(\mathcal{P}(S), C_b(S))$ is itself Polish, completely metrized by the Fortet-Mourier metric described below. Let $BL(S,\varrho)$ be the collection of bounded Lipschitz functions on (S,ϱ). This space is given the norm

$$\|h\|_{BL} := \sup_{x\in S}|h(x)| + \sup_{x\neq y}\frac{|h(x)-h(y)|}{\varrho(x,y)}. \tag{15}$$

Now set $d_{FM}(\mu,\nu) := \sup|\mu(h)-\nu(h)|$, where the supremum is over all $h \in BL(S,\varrho)$, $\|h\|_{BL} \leq 1$. Then d_{FM} metrizes $w(\mathcal{P}(S), C_b(S))$ (c.f., e.g., Dudley 2002, Theorem 11.3.3).

As usual, we say that

Definition 4.1. A collection $\mathcal{P}_0 \subset \mathcal{P}(S)$ is tight if, for every $\varepsilon > 0$, there is a compact set $K \subset S$ such that $\sup_{\mu\in\mathcal{P}_0} \mu(S \setminus K) < \varepsilon$. Also, $\mathcal{P}_0 \subset \mathcal{P}(S)$ is called precompact if it has compact closure.

We will make extensive use of Prohorov's Theorem (c.f., e.g., Aliprantis and Border, 1999, Theorem 14.22): For S Polish and $\mathcal{P}(S)$ with the $w(\mathcal{P}(S), C_b(S))$ topology, a subset of $\mathcal{P}(S)$ is tight if and only if it is precompact.

Proof of Lemma 3.2. Fix $\alpha \in W$. By Meyn and Tweedie (1993, Lemma D.5.3), a set $\mathcal{P}_0 \subset \mathcal{P}(S)$ is tight if (and only if) there is a norm-like function V such that $\sup_{\mu \in \mathcal{P}_0} \mu(V) < \infty$. This is true for $\mathcal{P}_0 = \{\mathbf{P}_\alpha^t \delta_x\}_{t=1}^\infty$, because if V, λ and b are as in Assumption 3.3, then

$$\begin{aligned}\mathbf{P}_\alpha^t \delta_x(V) &= \mathbf{P}_\alpha \mathbf{P}_\alpha^{t-1} \delta_x(V) \\ &= \int_S \left[\int_Z V[T_\alpha(u,z)]\nu(dz) \right] \mathbf{P}_\alpha^{t-1} \delta_x(du).\end{aligned}$$

(Substitute $\mathbf{P}_\alpha^{t-1}\delta_x$ for μ and V for $\mathbf{1}_B$ in (10).) But

$$\begin{aligned}&\int_S \left[\int_Z V[T_\alpha(u,z)]\nu(dz) \right] \mathbf{P}_\alpha^{t-1} \delta_x(du) \\ &\qquad \leq \int_S [\lambda V(u) + b] \mathbf{P}_\alpha^{t-1} \delta_x(du) = \lambda \mathbf{P}_\alpha^{t-1} \delta_x(V) + b.\end{aligned}$$

Repeating this argument t times gives

$$\mathbf{P}_\alpha^t \delta_x(V) \leq \lambda^t V(x) + \frac{b}{1-\lambda} \leq V(x) + \frac{b}{1-\lambda}. \tag{16}$$

$$\therefore \quad \sup_{t \in \mathbb{N}} \mathbf{P}_\alpha^t \delta_x(V) < \infty.$$

Thus $\{\mathbf{P}_\alpha^t \delta_x\}_{t=1}^\infty$ is tight, and hence precompact by Prohorov's Theorem. □

In order to prove Theorems 3.3–3.4 and Proposition 3.1 we will make use of the following lemma.

Lemma 4.1. *Fix $\beta \in W$. Under Assumptions 3.1—3.4, not only does $\mathbf{P}_\beta$ have a unique fixed point $\mu_\beta \in \mathcal{P}(S)$, but in addition the averages of the marginal distributions converge to it in distribution. That is,*

$$\frac{1}{t} \sum_{j=1}^t \mathbf{P}_\beta^j \delta_x \xrightarrow{d} \mu_\beta, \quad \forall x \in S. \tag{17}$$

Proof. Immediate from Lemma 3.2, Theorem 3.2 and Meyn and Tweedie (1993, Theorem 12.1.4). □

Proof of Theorem 3.3. For arbitrary $\beta \in W$, we know by Theorem 3.2 that the process $(X_t)_{t=0}^\infty = (X_t^\beta)_{t=0}^\infty$ has a unique stationary distribution. In other words, $\mathbf{P}_\beta$ has a unique fixed point $\mu_\beta \in \mathcal{P}(S)$. Let (α_n)

and α be as in the statement of the theorem. In what follows, write μ_n for μ_{α_n}, μ for μ_α, and analogously for T_{α_n}, T_α and so on.

We first claim that the set of stationary distributions $\{\mu_n\}_{n=1}^\infty$ is precompact in $w(\mathcal{P}(S), C_b(S))$. To see this, note that for every $x \in S$, (i) $P_x := \{\mathbf{P}_n^t \delta_x : t, n \in \mathbb{N}\}$ is a tight subset of $\mathcal{P}(S)$; and (ii) so is $Q_x := \{\frac{1}{t}\sum_{j=1}^t \mathbf{P}_n^j \delta_x : t, n \in \mathbb{N}\}$. The reason for (i) is as follows. By hypothesis, the function V and constants λ and b on the right hand side of (16) can be chosen independent of α_n (as well as t), so that $\sup_{t,n} \mathbf{P}_n^j \delta_x(V) < \infty$. This implies tightness as discussed in the proof of Lemma 3.2. From (i), (ii) is immediate. By Prohorov's theorem, tightness of Q_x implies precompactness. Finally, note that $\{\mu_n\}_{n=1}^\infty$ is in the closure of Q_x for some fixed $x \in S$ by (17). The claim follows.

As a result, every subsequence of (μ_n) has a convergent subsubsequence, written here as (μ_j). Denote the limit of (μ_j) by μ^0. Suppose we can show for this arbitrary subsequence that the limit point μ^0 is equal to μ. In this case every subsequence of (μ_n) has a subsubsequence converging to μ, and—since convergence in distribution is convergence in a topology—it must be that the sequence (μ_n) itself converges to μ, which is what we wish to prove.

So let (μ_j) be the above subsubsequence, $\mu_j \xrightarrow{d} \mu^0$. Pick any $h \in BL(S, \varrho)$, $\|h\|_{BL} \leq 1$. Fix $\varepsilon > 0$. We have $|\mathbf{P}\mu^0(h) - \mu^0(h)|$ dominated by

$$|\mathbf{P}\mu^0(h) - \mathbf{P}\mu_j(h)| + |\mathbf{P}\mu_j(h) - \mu_j(h)| + |\mu_j(h) - \mu^0(h)|.$$

$$\therefore \quad |\mathbf{P}\mu^0(h) - \mu^0(h)| \leq |\mathbf{P}\mu_j(h) - \mu_j(h)| + \varepsilon, \quad j \text{ large}, \tag{18}$$

where we have used Lemma 3.1. Now consider the remaining term. We have

$$\begin{aligned} |\mathbf{P}\mu_j(h) - \mu_j(h)| &= |\mathbf{P}\mu_j(h) - \mathbf{P}_j\mu_j(h)| \\ &\leq \int \left| \int h(T_j(x,z))\nu(dz) - \int h(T(x,z))\nu(dz) \right| \mu_j(dx). \end{aligned}$$

(Substitute h for $\mathbf{1}_B$ in (10).) Since $\{\mu_n\}$ and hence $\{\mu_j\}$ is precompact it is tight, and there is a compact $K \subset S$ with $\sup_j \mu_j(S \setminus K) < \varepsilon$. Define real functions g_j and g on K by

$$g_j(x) := \int h[T_j(x,z)]\nu(dz), \quad g(x) := \int h[T(x,z)]\nu(dz).$$

It follows from Assumption 3.5 and Dominated Convergence that g_j converges to g pointwise on K. Also, $\{g_j : j \in \mathbb{N}\}$ as a collection of real functions is uniformly bounded by 1 and uniformly equicontinuous, as, for any $x, x' \in K$,

$$\begin{aligned} |g_j(x) - g_j(x')| &\leq \int |h(T_j(x,z)) - h(T(x',z))|\nu(dz) \\ &\leq \int \varrho[T_j(x,z), T_j(x',z)]\nu(dz) \quad (\because \ \|h\|_{BL} \leq 1) \\ &\leq M\varrho(x,x'). \end{aligned}$$

From the Arzelà-Ascoli Theorem, $\{g_j\}$ is precompact in the supremum norm topology, and therefore has a *uniformly* convergent subsequence. Obviously the limit of this subsequence is g, so that, for some $j \in \mathbb{N}$, $|g_j(x) - g(x)| \leq \varepsilon$ for all $x \in K$. But

$$\begin{aligned} |\mathbf{P}\mu_j(h) - \mathbf{P}_j\mu_j(h)| &\leq \int_K |g_j(x) - g(x)|\mu_j(dx) \\ &+ \int_{S\setminus K} \left[\int |h(T_j(x,z)) - h(T(x',z))|\nu(dz) \right] \mu_j(dx) \leq \varepsilon + 2\varepsilon. \end{aligned}$$

Combining this with (18) gives

$$|\mathbf{P}\mu^0(h) - \mu^0(h)| < 4\varepsilon, \quad \forall \varepsilon > 0.$$

$$\therefore \quad |\mathbf{P}\mu^0(h) - \mu^0(h)| = 0, \quad \forall h \in BL(S,\varrho), \ \|h\|_{BL} \leq 1.$$

$$\therefore \quad \sup\{|\mathbf{P}\mu^0(h) - \mu^0(h)| : \|h\|_{BL} \leq 1\} = d_{FM}(\mathbf{P}\mu^0, \mu^0) = 0.$$

$$\therefore \quad \mathbf{P}\mu^0 = \mu^0 \quad (\because \ d_{FM} \text{ a metric}).$$

Since $\mathbf{P}$ has only one fixed point in $\mathcal{P}(S)$, we conclude that $\mu^0 = \mu$. This completes the proof. □

Proof of Proposition 3.1. Let α and β be as in the statement of the proposition. Pick any $x \in S$. Consider the Markov chains $(X_t^\alpha)_{t=0}^\infty$ and $(X_t^\beta)_{t=0}^\infty$ when $X_0^\alpha = X_0^\beta \equiv x$. From Huggett (2003, Theorem 2) we have

$$\mathcal{L}(X_t^\alpha) = \mathbf{P}_\alpha^t \delta_x \leq_{\mathcal{P}(S)} \mathbf{P}_\beta^t \delta_x = \mathcal{L}(X_t^\beta), \quad \forall t \in \mathbb{N} \tag{19}$$

whenever the following two conditions hold:

(i) $x \leq_S x'$ implies $\mathbf{P}_\gamma \delta_x \leq_{\mathcal{P}(S)} \mathbf{P}_\gamma \delta_{x'}$, all $\gamma \in W$.
(ii) $\alpha \leq_W \beta$ implies $\mathbf{P}_\alpha \delta_x \leq_{\mathcal{P}(S)} \mathbf{P}_\beta \delta_x$, all $x \in S$.

Both of these are easy to verify from the hypotheses of the proposition.

$$\therefore \quad \frac{1}{t}\sum_{j=1}^{t} \mathbf{P}_\alpha^j \delta_x \leq_{\mathcal{P}(S)} \frac{1}{t}\sum_{j=1}^{t} \mathbf{P}_\beta^j \delta_x \quad \forall t \in \mathbb{N}_0.$$

Taking limits with respect to t now gives $\mu_\alpha \leq_{\mathcal{P}(S)} \mu_\beta$, where we have used (17) and the fact that $\leq_{\mathcal{P}(S)}$ is a $w(\mathcal{P}(S), C_b(S))$-closed order on $\mathcal{P}(S)$ (Assumption 3.6—separating implies closed—and Torres, 1990, Theorem 6.1). □

Proof of Theorem 3.4. Once again, for each $\beta \in W$, $\mathbf{P}_\beta$ has a unique fixed point and (17) holds. Take (α_n) monotone and converging to $\alpha \in W$. We can and do assume that the sequence is increasing: $\alpha_n \uparrow \alpha$.[12] As before we write μ_n for μ_{α_n}, μ for μ_α, and analogously for T_{α_n}, T_α and so on.

The first important observation is that By Proposition 3.1,

$$\mu_1 \leq_{\mathcal{P}(S)} \mu_n \leq_{\mathcal{P}(S)} \mu, \quad \forall\, n \in \mathbb{N}.$$

Moreover, in the present setting and order intervals are compact (Assumption 3.6 and Torres ,1990, Theorem 6.6). Indeed, it follows from Assumption 3.6, and Torres (1990, Proposition 6.7) that $\mu_n \xrightarrow{d} \mu^0$ for some $\mu^0 \in \mathcal{P}(S)$. If we can show that $\mu^0 = \mu$ then the proof is done.

[12]Otherwise define another order $\leq'$ on W by $\alpha \leq' \beta$ iff $\beta \leq_W \alpha$.

From Assumption 3.6 and Torres (1990, Proposition 6.2), the increasing functions in $C_b(S)$ separate $\mathcal{P}(S)$, in the sense that if μ' and μ'' are distinct elements of $\mathcal{P}(S)$, then there is an $h \in ib\mathscr{B}(S) \cap C_b(S)$ with $\mu'(h) \neq \mu''(h)$. Given that μ is the only fixed point of $\mathbf{P}$, then, it suffices to show that

$$\mathbf{P}\mu^0(h) = \mu^0(h), \quad \forall\, h \in ib\mathscr{B}(S) \cap C_b(S). \tag{20}$$

So take $h \in ib\mathscr{B}(S) \cap C_b(S)$ and $\varepsilon > 0$. As in the proof of Theorem 3.3, $|\mathbf{P}\mu^0(h) - \mu^0(h)|$ is dominated by

$$|\mathbf{P}\mu^0(h) - \mathbf{P}\mu_n(h)| + |\mathbf{P}\mu_n(h) - \mu_n(h)| + |\mu_n(h) - \mu^0(h)|.$$

$$\therefore \quad |\mathbf{P}\mu^0(h) - \mu^0(h)| \leq |\mathbf{P}\mu_n(h) - \mu_n(h)| + \varepsilon, \quad n \text{ large}, \tag{21}$$

using Lemma 3.1. Consider the remaining term. We have

$$\begin{aligned} |\mathbf{P}\mu_n(h) - \mu_n(h)| &= |\mathbf{P}\mu_n(h) - \mathbf{P}_n\mu_n(h)| \\ &\leq \int \left| \int h(T_n(x,z))\nu(dz) - \int h(T(x,z))\nu(dz) \right| \mu_n(dx). \end{aligned}$$

(Substitute h for $\mathbf{1}_B$ in (10).) Since $\{\mu_n\}$ is precompact by Prohorov's theorem it is tight, so there is a compact $K \subset S$ with $\sup_n \mu_n(S \setminus K) < \varepsilon$. Define real functions g_n and g on K by

$$g_n(x) := \int h[T_n(x,z)]\nu(dz), \quad g(x) := \int h[T(x,z)]\nu(dz).$$

It follows from Assumption 3.5 and Dominated Convergence that g_n converges to g pointwise on K. Since h is increasing, the pointwise convergence is monotone. Since $h \in C_b(S)$ and Assumption 3.2 holds, g_n and g are continuous. From Dini's Lemma, then, g_n also converges uniformly to g on K, so that, for some $n \in \mathbb{N}$, $|g_n(x) - g(x)| \leq \varepsilon$ for all $x \in K$. But then

$$\begin{aligned} |\mathbf{P}\mu_n(h) - \mathbf{P}_n\mu_n(h)| &\leq \int_K |g_n(x) - g(x)|\mu_n(dx) \\ &+ \int_{S\setminus K} \left[\int |h(T_n(x,z)) - h(T(x',z))|\nu(dz) \right] \mu_n(dx) \leq \varepsilon + 2B\varepsilon, \end{aligned}$$

where $B := \sup_{x \in S} |h(x)|$. Combining this with (21) gives

$$|\mathbf{P}\mu^0(h) - \mu^0(h)| < (2 + 2B)\varepsilon.$$

This verifies (20), which completes the proof. $\square$

References

[1] Aliprantis, C. D. and K. C. Border (1999): *Infinite Dimensional Analysis,* Second Edition, Springer-Verlag, New York.

[2] Danthine, J-P. and J. B. Donaldson (1981): Stochastic Properties of Fast vs. Slow Growth Economies, *Econometrica,* 49 (4), 1007–1033.

[3] Dudley, R. M. (2002): *Real Analysis and Probability,* Cambridge Studies in Advanced Mathematics No. 74, Cambridge University Press.

[4] Duffie, D. and K. J. Singleton (1993): "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica,* 61 (4) 929–952.

[5] Futia, C. A. (1982): "Invariant Distributions and the Limiting Behavior of Markovian Economic Models," *Econometrica,* 50, 377–408.

[6] Hopenhayn, H. A. and E. C. Prescott (1992): "Stochastic Monotonicity and Stationary Distributions for Dynamic Economies," *Econometrica,* 60, 1387–1406.

[7] Huggett, M. (2003): "When are comparative dynamics monotone?" *Review of Economic Dynamics,* 6 (1), 1–11.

[8] Lasota, A. and M. C. Mackey (1994): *Chaos, Fractals and Noise: Stochastic Aspects of Dynamics,* 2nd edition, Springer-Verlag, New York.

[9] Meyn, S. P. and Tweedie, R. L. (1993): *Markov Chains and Stochastic Stability,* Springer-Verlag: London.

[10] Phelps, E.S. (1961): "The Golden Rule of Accumulation: A Fable for Growthmen," *American Economic Review,* 51, 638–643

[11] Santos, M. S. and A. Peralta-Alva (2003): "Accuracy of Simulations for Stochastic Dynamic Models." Manuscript.

[12] Santos, M. and Vigo-Aguiar J. T. (1998): "Analysis of a Numerical Dynamic Programming Algorithm Applied to Economic Models," *Econometrica,* 66, 2, 409–426.

[13] Schenk-Hoppé, K. R (2002): "Is There a Golden Rule for the Stochastic Solow Growth Model?" *Macroeconomic Dynamics* 6, 457–475.

[14] Stokey, N. L., R. E. Lucas and E. C. Prescott (1989): *Recursive Methods in Economic Dynamics* Harvard University Press, Massachusetts.

[15] Torres, R. (1990): “Stochastic Dominance.” DP n. 905, September 1990, Center for Mathematical Studies in Economics and Management Science, Northwestern University

Center for Operations Research and Econometrics, Université Catholique de Louvain, 34 Voie du Roman Pays, B-1348 Louvain-la-Neuve, Belgium