

INVARIANT PROBABILITY DISTRIBUTIONS IN ECONOMIC MODELS: A GENERAL RESULT

ALFREDO MEDIO

University Ca' Foscari of Venice

This paper discusses the asymptotic behavior of distributions of state variables of Markov processes generated by first-order stochastic difference equations. It studies the problem in a context that is general in the sense that (i) the evolution of the system takes place in a general state space (i.e., a space that is not necessarily finite or even countable); and (ii) the orbits of the unperturbed, deterministic component of the system converge to subsets of the state space which can be more complicated than a stationary state or a periodic orbit, that is, they can be aperiodic or chaotic. The main result of the paper consists of the proof that, under certain conditions on the deterministic attractor and the stochastic perturbations, the Markov process describing the dynamics of a perturbed deterministic system possesses a unique, invariant, and stochastically stable probability measure. Some simple economic applications are also discussed.

Keywords: Stochastic Stability, Markov Processes, Chaotic Dynamics, Sunspots

1. INTRODUCTION AND MOTIVATION

This paper discusses the asymptotic behavior of distributions of the state variables of Markov processes generated by first-order stochastic difference equations. The theory of stochastic dynamical systems in discrete time plays a very significant role in economic dynamics. The well-known treatise by Stokey and Lucas (1989) provides an extensive discussion of the relevant mathematical and statistical methods and includes many applications of the theory of Markov processes to economic models.

To fix ideas and put the mathematical background in place, we start with some formal definitions.

The basic stochastic dynamical system investigated in the paper can be described by the following equation:

$$x_{t+1} = T(x_t, \xi_{t+1}^\epsilon), \quad (1)$$

Financial support from the Italian Ministry of the University (MURST) and the National Council of Research (CNR) is gratefully acknowledged. The author is greatly indebted to Sean Meyn for a very stimulating e-mail correspondence and many insightful comments. Many thanks are also due to Alain Monfort and Sergio Invernizzi, whose comments prompted the author to tidy up the proof of Theorem 1, to an Associate Editor of MD, and two anonymous referees for many helpful comments and criticisms. The author is of course solely responsible for any remaining errors or misunderstandings. Address correspondence to: Professor Alfredo Medio, Department of Economics, University Ca' Foscari of Venice, 873 S. Giobbe, 30121 Venice, Italy; e-mail: medio@unive.it.

where the $\xi_t^\epsilon (t = 1, 2, \dots)$ are i.i.d. random vectors with values in W_ϵ , an open subset of \mathbf{R}^m ; the state variables x_t take values in M , an open subset of \mathbf{R}^n ; the initial vector x_0 is a given constant (not random). Alternatively, we can take x_0 as a random vector taking values in M , arbitrary but independent of ξ_t^ϵ for $t \geq 1$. In either case, x_t is independent of ξ_{t+1}^ϵ , for all $t \geq 0$. The index ϵ parameterizes the level of ξ perturbations. T is a measurable function mapping $M \times W_\epsilon$ to $M \subset \mathbf{R}^n$.

The fact that x_{t+1} is conditionally independent of x_{t-1}, x_{t-2}, \dots , given x_t , ensures that (1) has the Markovian property, that is, for any integrable function ϕ , we have

$$\mathbf{E}[\phi(\{x_\tau\}), \tau > t \mid \{x_\tau\}, \tau \leq t] = \mathbf{E}[\phi(\{x_\tau\}), \tau > t \mid x_t] \quad \forall t \geq 0$$

In words, this means that the present value of the state variable x contains all the information from its past history relevant for the prediction of its future. System (1) generates a collection $\mathbf{X} = \{x_t\} \mid t \in \mathbf{Z}_+$ of random variables, which is called *Markov chain* or *process*.

The dynamics of Markov processes such as (1) can be defined by the iterations of an one-step transition probability kernel

$$P(x, A) \quad x \in M, \quad A \in \mathcal{B}(M),$$

where, denoting by M the state space and by $\mathcal{B}(M)$ the Borel σ -algebra on M , we have

- (i) for each $A \in \mathcal{B}(M)$, $P(\cdot, A)$ is a nonnegative measurable function on M ;
- (ii) for each $x \in M$, $P(x, \cdot)$ is a probability measure on $\mathcal{B}(M)$;
- (iii)

$$P^n(x, A) = \int_M P(x, dy) P^{n-1}(y, A), \quad x, y \in M, \quad A \in \mathcal{B}(M), \quad n \geq 1,$$

with $P^0(x, A) = \delta_x(A)$, δ_x being the Dirac measure centered on x .

The formal relation between the map T of equation (1) and the corresponding transition probability kernel is the following:

$$P_\epsilon(x, A) = \int_{W_\epsilon} \chi_A[T(x, \xi^\epsilon)] \nu_\epsilon(d\xi^\epsilon),$$

where χ_A is the indicator function of the set A , i.e.,

$$\chi_A(x) = \begin{cases} 1 & \text{for } x \in A \\ 0 & \text{for } x \notin A, \end{cases}$$

and ν_ϵ is the probability measure

$$\nu_\epsilon(B) = \text{prob}(\xi_t^\epsilon \in B) \quad \text{for } B \in \mathcal{B}(W_\epsilon),$$

identical for all t .

In the investigation of systems represented by equation (1), we are especially interested in finding invariant configurations of the (random) state variables. Then we need the following.

DEFINITION 1. *We say that a probability measure $\pi : \mathcal{B}(M) \rightarrow [0, 1]$ is invariant with respect to P or, equivalently, that π is preserved by P , if we have*

$$\pi(A) = \int P(x, A)\pi(dx) \quad A \in \mathcal{B}(M).$$

An interesting special case of (1) obtains when we have $\{\xi_t^\epsilon\} = \{\hat{\xi}, \hat{\xi}, \hat{\xi}, \dots\}$, where $\hat{\xi}$ is a vector of constants and therefore the difference equation (1) is deterministic. In what follows, we assume that for $\epsilon = 0$, $\xi_t^0 = \hat{\xi}$ for all t , where $\hat{\xi}$ corresponds to a certain “normal” configuration of parameters. In each particular application, it is possible to normalize $\hat{\xi}$ to zero. Thus, we can write

$$F(x_t) \equiv T(x_t, \xi_{t+1}^0) = T(x_t, 0) \quad \forall t \geq 0,$$

where the map F denotes the deterministic component of the stochastic dynamical system (1), which we shall often call “deterministic core.”

Although the results presented below, and in particular Theorem 1, can be applied to any problem taking the form of a Markovian process, we had in mind especially two broad types of economic models giving rise to stochastic difference equation such as (1). Because the models in questions are well known, we describe them very briefly, ignoring variations and omitting many technical details, for which we refer the reader to the literature. Our purpose here is to relate the models in questions to the mathematical setup described above.

1.1. Type 1 Models: Stochastic Intertemporal Optimization

The first class of models deals with intertemporal optimization problems, in the presence of exogenous shocks perturbing fundamentals. It includes single-agent models of optimal growth, inventory accumulation, asset pricing, search unemployment and many others. Earlier results in this area are found in Brock and Mirman (1972), subsequently extended, by Radner (1973), Brock and Majumdar (1978), Majumdar and Zilcha (1987), and Joshi (1995), among others. A detailed discussion of the basic mathematical methods involved is provided by Stokey and Lucas (1989).

A general representation of these problems can be written as follows:

$$\begin{aligned} & \sup_{\{y_t\}_{t=0}^\infty} \mathbf{E} \left[\sum_{t=0}^\infty \beta^t f(k_t, y_t, \omega_t) \right] \\ & \text{s.t. } (k_t, y_t, \omega_t) \in D = [K \times Y \times \Omega \mid y_t \in \Gamma(k_t, \omega_t)]; \quad \text{(P1)} \\ & (k_0, \omega_0) \in K \times \Omega \text{ are given constants,} \end{aligned}$$

where k is the endogenous state variable; y is the control variable; $K \subseteq \mathbf{R}^n, Y \subseteq \mathbf{R}^d$; $\{\omega_1, \omega_2, \dots\}$ is a sequence of i.i.d. exogenous shocks taking values in $\Omega \subseteq \mathbf{R}^m$; $\beta \in (0, 1)$ is the constant discount factor. The correspondence Γ is the set of all possible control choices, given the known, endogenous and exogenous, state of the system; $f : D \rightarrow \mathbf{R}$ is the return function.

The Bellman equation for problem (P1) can be written as

$$V(k, \omega) = \sup_{y|(k,y,\omega) \in D} \left\{ f(k, y, \omega) + \beta \int_{\Omega} V[\phi(k, y, \omega'), \omega'] \nu(d\omega') \right\}, \tag{2}$$

where (k, ω) is the current, known state of the system; y the control variable; ω' is next period's shock, unknown at the moment of decision; and the function ϕ is the given "law of motion" relating the next period's value of the endogenous variable to its current state, the current decision, and the future exogenous shock.

Under appropriate, fairly standard assumptions on the spaces K, Y, Ω , the return function f , the feasibility constraint D , and the "law of motion" ϕ , there exists a unique, continuously differentiable function V satisfying (2) and an associated, continuous "policy function" $g(k, \omega)$, which determines the optimal value of the control variable y for any given pairs of the endogenous and exogenous state variables. Thus, from g and ϕ , we can derive a difference equation of the form

$$k_{t+1} = \Phi(k_t, \omega_t, \omega_{t+1}). \tag{3}$$

The sequence of the joint state variable $\{x_t = (k_t, \omega_t)\}$ taking values in the set $K \times \Omega$ follows a Markov process. If the sequence $\{\omega_t\}$ is i.i.d., from equation (3) we can derive a first-order stochastic difference equation of the form

$$x_{t+1} = T(x_t, \omega_{t+1}) \tag{4}$$

with the same properties as (1).

1.2. Type 2 Models: Sunspots in Sequential Markets Models

This class of models includes many variations, mostly of the overlapping generations (OLG) type. We shall provide an abstract characterization of it, referring the reader to the comprehensive survey by Chiappori and Guesnerie (1991) for details and a rich bibliography.

Consider a competitive economy in which, at each time t , the present state x_t is entirely determined by nonstochastic fundamentals, agents' expectations about the state one step ahead, x_{t+1} , and the equilibrium requirement that markets clear at all times. Suppose now that agents observe a Markov process of i.i.d. signals, ξ_t^ϵ , characterized by a probability distribution ν_ϵ , as described earlier, and that they commonly believe that these signals are perfectly correlated with the equilibrium values of the state variable x , in the sense that there exists a homeomorphism ϕ

such that

$$x_t = \phi(\xi_t^\epsilon) \quad \forall t. \tag{5}$$

Then, the equilibrium conditions can be written as

$$\mathbf{E}[\mathcal{G}(x_t, x_{t+1}) \mid x_t] = \int_M \mathcal{G}(x_t, x_{t+1}) P(x_t, dx_{t+1}) = 0, \tag{6}$$

where $P(x_t, dx_{t+1}) = v_\epsilon[\phi^{-1}(x_t, dx_{t+1})]$, and the function \mathcal{G} depends on (non-stochastic) fundamentals, for example, utility functions. We say that agents' expectations are self-fulfilling if the equilibrium values of x , as determined by fundamentals and beliefs, actually validate equations (5) and (6). In this case, the resulting sequence of equilibrium values $\{x_t\}$ is a Markov process with probability transition kernel $P(x, A)$, which is called *sunspot equilibrium (SE)*. An SE is *stationary (SSE)* if there exists a probability measure π , invariant for the Markov process, as in Definition 1.

As will be illustrated in Example 2, below, sufficient conditions for the validation of (5) and (6) can be represented by an equation $\Phi(x_t, x_{t+1}, \xi_{t+1}^\epsilon) = 0$. When the conditions for the implicit function theorem hold, Φ can be inverted locally with respect to x_{t+1} , yielding a difference equation $x_{t+1} = T(x_t, \xi_{t+1}^\epsilon)$ like (1). In sunspot OLG models, “locally” usually means in a neighborhood of a stationary, or periodic, solution of the perfect-foresight, “deterministic core.” Whether or not the map T is uniquely defined globally depends on the properties of the function \mathcal{G} and ultimately on the fundamental functions of the model. The results of this paper can therefore be applied both to the “invertible” and, locally, to the “noninvertible” case. Global analysis of the noninvertible case is difficult and involves discussion of delicate questions concerning the so-called “backward dynamics” that would take us far afield. Because our main interest here is to establish global results, the examples discussed in Section 3 refer to problems of Type 1 (stochastic dynamic optimization), and to problems of Type 2 (sunspots) limited to the invertible case.

2. MAIN RESULT

Our discussion of existence and stability of invariant probability distributions for Markov processes generated by equation (1) is general in the following, twofold sense:

- We discuss the question in the context of *general spaces*. In the present context, the term “general” refers to cardinality, not dimension, and means that the space in which the Markov process takes values is not necessarily finite (as almost always assumed in the sunspot literature), or even countable. Because endogenous state variables in deterministic economic models (e.g., consumption, labor supply, capital stock) are typically assumed to be continuous, we do not see why their stochastic counterparts should be confined to finite-state spaces. To make this problem tractable, some regularity assumptions are needed and are provided by Assumptions 1, 3, and 4, below.

- We discuss the relations between stability of invariant sets of the “deterministic core” of the system and existence and stability of invariant probability distributions, in the general case in which those sets are not necessarily fixed points or periodic orbits, but they can be aperiodic or chaotic. This is all the more interesting because there is strong analytical and numerical evidence that deterministic optimal growth models as well as overlapping generations models with perfect foresight can have complex or chaotic dynamics. See, for example, the survey by Boldrin and Woodford (1990) and the readings volume edited by Benhabib (1992). In particular, this paper provides a generalization of some of the results of the sunspot literature linking local (forward) stability of stationary or periodic states to local sunspots, showing that there exists an analogous relation between the existence of attractors of any kind and *global* sunspots.

Proving the existence of an invariant probability distribution for the Markov process under investigation may not be very interesting if that distribution only obtains for very special (random) initial conditions. Therefore, we explicitly discuss the question of uniqueness and stochastic stability of invariant distributions, making use of some recent, powerful results in this field of research.

Given the rather technical nature of the argument, we postpone a more detailed discussion of the relevant specialized literature until the end of the paper, when the necessary concepts and methods have been properly defined. For the unexplained concepts and the basic results used in the following pages and in particular in Appendices A and B, we refer the reader to the excellent treatise on stochastic stability by Meyn and Tweedie (1993), which contains the state of the art in this area.

In view of what we said before and without loss of generality, we study the basic stochastic difference equation (1) in the following decomposed form:

$$x_{t+1} = T(x_t, \xi_{t+1}^\epsilon) = F(x_t) + G(x_t, \xi_{t+1}^\epsilon), \tag{7}$$

where the map $F(x_t) = T(x_t, 0)$ is the deterministic core, which we assume to be known, and the map $G(x_t, \xi_{t+1}^\epsilon) = T(x_t, \xi_{t+1}^\epsilon) - F(x_t)$ denotes the perturbation term.

We also need to consider the related family of equations:

$$x_k = T_k(x_0, \xi_1^\epsilon, \dots, \xi_k^\epsilon), \tag{8}$$

where for each k the function $T_k : M \times W_\epsilon^k \rightarrow M$ (to be distinguished from the k th iterate of T , T^k !) can be determined inductively as follows:

$$T_k(x_0, \xi_1^\epsilon, \dots, \xi_k^\epsilon) = T[T_{k-1}(x_0, \xi_1^\epsilon, \dots, \xi_{k-1}^\epsilon), \xi_k^\epsilon] \quad k > 1$$

with $T_0(x_0) = x_0$; $T_1(x_0, \xi_1^\epsilon) = T(x_0, \xi_1^\epsilon)$.

Equation (8) can be interpreted by saying that, for any given initial state $x_0 \in M$, and any given set of values $(\xi_1^\epsilon, \dots, \xi_k^\epsilon) \in W_\epsilon^k$, the value of x at time k , x_k , is

determined by the function T_k . In that sense, equations (8) describe a deterministic system uniquely related to the Markov chain (1), known in the literature as an associated control system (ACS). We have more to say about ACS later.

Let us now introduce the following assumptions:

Assumption 1. The maps F and G (and therefore the map T) are continuously differentiable with bounded derivatives. This implies that those maps are Lipschitzian on bounded sets.

To formulate the next assumption, we need a few preliminary definitions.

For a set $A \subset M$ and a scalar constant $r > 0$, let $L(r, A) = [x \in M \mid d(x, A) < r]$, where $d(x, A)$ denotes the distance from x to A , namely,

$$d(x, A) = \inf_{y \in A} \{\|x - y\|\}$$

and $\|\cdot\|$ is any vector norm, such as the Euclidean distance. Then we have the following definition.

DEFINITION 2. A compact subset Λ of the state space M , invariant under a map f , is said to be Lyapunov stable if, for any $\epsilon > 0$, there exists $\delta > 0$ (depending on ϵ) such that $f^n[L(\delta, \Lambda)] \subset L(\epsilon, \Lambda)$, for $n \geq 0$. Λ is said to be asymptotically stable if

- (a) it is Lyapunov stable; and
- (b) the basin of attraction of Λ , i.e., the set

$$B(\Lambda) = \{x \in M \mid \lim_{n \rightarrow \infty} d[f^n(x), \Lambda] = 0\},$$

is open non-empty.

DEFINITION 3. A map $f : \Lambda \rightarrow \Lambda$ is said to be topologically transitive (t.t.) on Λ if, for any two open sets $U, V \subset \Lambda$, there exists an integer n such that $f^n(U) \cap V \neq \emptyset$.

Transitivity implies that orbits generated by the map f starting from any arbitrarily small open neighborhood visit any other arbitrarily small open neighborhood in Λ in finite time. Thus, the set Λ is dynamically indecomposable and must be studied as one piece.

Sometimes we need stronger forms of transitivity, as defined below.

DEFINITION 4. A map $f : \Lambda \rightarrow \Lambda$ is said to be strongly topologically transitive (s.t.t.) on Λ if for any integer $m > 0$ the map f^m is topologically transitive on Λ .

DEFINITION 5. A map $f : \Lambda \rightarrow \Lambda$ is said to be topologically mixing on Λ if for any two open non-empty sets $U, V \subset \Lambda$ there exists a positive integer N such that for every $n > N$, $f^n(U) \cap V \neq \emptyset$.

Topological mixing implies s.t.t. and each of these properties implies t.t.

DEFINITION 6. An asymptotically stable set Λ is said to be an attractor if it is “indecomposable” under the action of the map f in the sense that f is topologically transitive on Λ .

We can now write the Assumption 2.

Assumption 2. The deterministic system

$$x_{t+1} = F(x_t) \tag{9}$$

possesses a (locally unique) attractor $\Lambda \subset M$, which is in the interior of its basin of attraction $B(\Lambda)$. Moreover, the map F is strongly topologically transitive on Λ .

The stronger form of indecomposability is necessary to guarantee aperiodicity of the Markov chain generated by (7). The case in which F is not t.t., and the one in which F is t.t., but not strongly t.t., are discussed in the Remarks 2 and 3 below. Notice that we do *not* require that the attractor be exponentially stable.

Assumption 3. For all $\epsilon > 0$, the probability measure ν_ϵ characterizing the i.i.d. random variables ξ_t^ϵ (discussed earlier) possesses a density γ_ϵ that is lower semi-continuous (l.s.c.), and its support, that is, the set

$$W_\epsilon = \{x \mid \gamma_\epsilon(x) > 0 \text{ a.s.}\},$$

is an open, bounded subset of \mathbf{R}^m , containing the singleton $\{0\}$. (For $\epsilon = 0$, the measure ν_ϵ degenerates to the Dirac measure centered on $\{0\}$.)

This assumption is obviously satisfied for many distributions with continuous density, commonly assumed in the description of noise. However, continuity is a sufficient, but not necessary, condition for l.s.c. For example, uniform distributions on bounded, *open* sets, although not continuous, would satisfy Assumption 3. In fact, a discrete distribution with mass concentrated on k points (x_1, \dots, x_k) can be approximated by a uniform distribution with support on small open balls centered on those points.

Before stating Assumption 4, we need to define the concept of *forward accessibility*. Keeping in mind the earlier definition of ACS and, in particular, equation (8), we can write Definition 7.

DEFINITION 7. Let $A_+(x)$ denote the set of all states reachable from x at some time in the future for any admissible sequences of perturbations $\{\xi_1^\epsilon, \xi_2^\epsilon, \dots\}$. Then, an ACS is said to be forward accessible (FA) if, for each $x_0 \in M$, the set $A_+(x_0) \subset M$ has a non-empty interior.

Then, we can state the following assumption.

Assumption 4. The deterministic ACS (8) associated with the stochastic system (7) is forward accessible.

From Definition 7 (and the definition of interior), we gather that a Markov chain X , for which the corresponding ACS is FA, cannot be concentrated in some lower dimensional subset of the state space (e.g., if $M \subset \mathbf{R}$, the chain cannot be concentrated in a point; if $M \subset \mathbf{R}^2$, it cannot be concentrated in a line, and so forth). For additional comments on FA, together with a brief description of the method for ascertaining the presence of FA in simple models, see Appendix B.

Assumption 5. For all ξ^ϵ , the perturbation term $\|G(x, \xi^\epsilon)\|$ is bounded uniformly for $x \in B(\Lambda)$, i.e., for each $\epsilon \geq 0$ there exists a $L^\epsilon < \infty$, such that

$$\sup_{x \in B(\Lambda)} \|G(x, \xi^\epsilon)\| \leq L^\epsilon \quad \text{a.s.}$$

and

$$L^\epsilon \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0.$$

Thus, by reducing ϵ , we can make the perturbation level as small as we please and in the limit for $\epsilon \rightarrow 0$ the stochastic process (7) degenerates to its “deterministic core” (9).

We can now prove the following theorem.

THEOREM 1. *Let a Markov chain X be defined by (7), and by the associated transition probability kernel $P_\epsilon(x, A)$, and let assumptions 1–5 hold. Then, for any deterministic attractor Λ and a sufficiently small ϵ [depending on the size of the basin of attraction $B(\Lambda)$], there exists a set $O_\epsilon \subset B(\Lambda)$ that is absorbing; i.e., $P_\epsilon(x, O_\epsilon) = 1$ if $x \in O_\epsilon$. Moreover, for the chain X restricted to O_ϵ , the following properties hold:¹*

(a) X is uniformly ergodic; namely,

$$\sup_{x \in O_\epsilon} \|P_\epsilon^n(x, \cdot) - \pi_\epsilon\| \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

where, for any given ϵ , π_ϵ is a uniquely determined, finite (i.e., finite mass) probability measure, which is invariant with respect to P_ϵ , according to Definition 1.

(b) For all $x \in O_\epsilon$, there exists $r_\epsilon > 1$ and $R_\epsilon < \infty$, such that

$$\|P_\epsilon^n(x, \cdot) - \pi_\epsilon\| \leq R_\epsilon r_\epsilon^{-n};$$

i.e., the convergence of P_ϵ^n to the invariant measure π_ϵ takes place at a geometric rate.

(c) Moreover, π_ϵ is stable in the sense that, for any initial probability measure μ , we have

$$\left\| \int P_\epsilon^n(x, \cdot) \mu(dx) - \pi_\epsilon \right\| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

The proof of the Theorem above—which is the core of this paper—is rather long and quite technical and we have relegated it to Appendix A. In economic terms, the consequence of the mathematical results listed above can be summarized in the following proposition.

PROPOSITION 1. *Let equation (7) describe a problem of Type 1 (e.g., stochastic optimal growth), let equation (9) describe the corresponding deterministic core, and let the conditions of the Theorem hold. Then,*

- (a) *An invariant probability distribution π_ϵ for (7) exists, is unique and has a bounded support that is nontrivial in the sense that π_ϵ is not a Dirac measure. The support of π_ϵ is in the set O_ϵ defined in Theorem 1 and includes the attractor of (9).*
- (b) *Given any random initial distribution (with support in O_ϵ), arbitrary but independent of the shocks, the distribution will converge asymptotically to π_ϵ , which therefore properly characterizes the long-run behavior of the system.*

Analogously, in the case of Type 2 problems (sunspots), we have the following proposition.

PROPOSITION 2. *Let equation (7) describe a sunspot equilibrium, let equation (9) describe a (forward-moving) perfect-foresight equilibrium (PFE), and let the conditions of Theorem 1 hold. Then,*

- (a) *A stationary sunspot equilibrium (represented by π_ϵ) exists, is unique, and has a bounded, nontrivial support in the set O_ϵ . The support of π_ϵ includes the attractor of the perfect-foresight deterministic core (9).*
- (b) *For any initial distribution (with support in O_ϵ), chosen arbitrarily but independently of the perturbations, the distribution will converge asymptotically to π_ϵ , corresponding to the unique SSE.*

Remark 1. Strong topological transitivity is satisfied trivially when the deterministic attractor Λ is a fixed point. It is also satisfied when the dynamics of map F are chaotic and mixing on Λ . As we shall see in Example 3, strong transitivity may also be verified when the attractor is quasiperiodic (aperiodic but not chaotic).

Remark 2. When the deterministic map is t.t. on Λ , but not strongly t.t., Assumption 2 is violated, the Markov process need not be aperiodic or converge to a unique probability distribution. This situation obviously occurs when Λ is periodic and, less obviously, when Λ consists of two or more chaotic sets mapped into each other cyclically by F —the so-called periodic, or nonmixing chaos. However, this is not as serious a drawback as it seems, in view of the fact [cf. Meyn and Tweedie (1993, Proposition 5.4.6, p. 118)] that for an irreducible Markov process X periodic with period d , the state space can be decomposed as

$$M = \bigcup_{i=1}^d D_i \cup E,$$

where the d sets D_i , $i = 1, \dots, d$ are disjoint, absorbing, and irreducible for the process X^d generated by the probability transition kernel P^d (i.e., the d th iterate of P), and the process X^d on any of the sets D_i is aperiodic. The residual set E is transient and therefore negligible.² The conclusions of Theorem 1 still hold for the process X^d on any of the absorbing sets D_i . All of this means that, if the dynamics

generated by the map T of equation (1) or (7) and the corresponding Markov process X are d -periodic, then we can change the time unit and consider instead the map T^d (the d th iterate of T) and the associated kernel P^d . The corresponding Markov process will be aperiodic and, as $n \rightarrow \infty$, the probability measure defined by $P^{nd}(x, \cdot)$ will converge to one of d probability measures π_i (one for each set D_i) and the choice among them will depend on the initial conditions.³

Remark 3. When the deterministic map F is not t.t. on an attracting set Λ (e.g., Λ is decomposable into two invariant sets), Assumption 2 is violated, the Markov process is not irreducible, and Theorem 1 is no longer valid. However, for any reducible T-chain, there exists a finite decomposition of the state space

$$M = \bigcup_{k=0}^n H_k \cup E$$

where the sets H_k are disjoint, absorbing sets and the chain restricted to any of the H_k sets is uniformly ergodic, whereas E a transient set. For each of the absorbing, irreducible set H_k , there is a unique, invariant, probability distribution stochastically stable for appropriate initial conditions [cf. Tuominen and Tweedie (1979); Meyn and Tweedie (1993, p. 408)].

Remark 4. The hypothesis of independence of the exogenous perturbations is not essential. When perturbations are not independent, we can still write a stochastic difference equation in which there appear i.i.d. exogenous perturbations, but the equations need not be of first order. That is to say, we would have equations of the form

$$w_{t+1} = T(w_t, w_{t-1}, \dots, \xi_{t+1}),$$

where w_t denotes the state variable, the sequence $\{\xi_t\}$ is i.i.d., and the order of the difference equation depends on the structure of the dependence of perturbations. The resulting n th-order system of equations can then be reduced to an equivalent first-order system by extending the state space through the introduction of appropriate auxiliary variables. This procedure is illustrated in Example 4.

3. SOME SIMPLE ECONOMIC EXAMPLES

A brief discussion of simple economic models of Type 1 and Type 2 will help understand the nature and relevance of our results and the assumptions on which they are based. To avoid repetitions, let us start with some general considerations concerning the four models that follow. First, the functional relationships commonly adopted in the models considered here guarantee that the “smoothness assumption,” Assumption 1, is satisfied. Second, the r.v.s considered below comply with Assumption 3. This does not pose any particular restriction on the economic primitives of the model and, as mentioned on p. 10 above, it allows a wide choice of stochastic perturbations.

3.1. Example 1: Type 1—One-Dimensional Optimal Growth

Consider the well-known, one-good/two-sector optimal growth model in reduced form. In the deterministic version of the model, if we choose the single capital good k as the endogenous state variable, and the control variable is the amount y of output saved and invested, the map governing the dynamics of k along an optimal path coincides with the optimal policy function g ; that is, we have

$$k_{t+1} = y_t = g(k_t). \tag{10}$$

It is known that, for sufficiently small discount factor β (i.e., sufficiently large discount rate), any C^2 map can be a policy function for a problem of optimal growth satisfying the standard economic requirements—in a nutshell, convexity of technology and convexity of preferences [see Montrucchio (1986), Boldrin and Montrucchio (1986)]. In particular, there exist specifications of the return function f , satisfying those economic requirements, that yield the optimal policy map

$$k_{t+1} = \mu k_t (1 - k_t), \tag{11}$$

i.e., the “logistic” map much studied in the literature on chaotic dynamics [in the present context, besides the articles quoted above, see Deneckere and Pelikan (1986)].

Suppose now we introduce a stochastic perturbation such that a random proportion ξ_{t+1}^ϵ of output saved at each instant t is wasted *before* it can be used as production input, but *after* the optimal choice of how much to save has been made. Consequently, we have $k_{t+1} = y_t (1 - \xi_{t+1}^\epsilon)$ and

$$k_{t+1} = \mu k_t (1 - k_t) (1 - \xi_{t+1}^\epsilon), \tag{12}$$

which is a special case of problem (P1), above, where the return function $f(k, y)$ in this case can be interpreted as the “consumption frontier.”

We also assume that $\{\xi_t^\epsilon\}$ is a sequence of i.i.d. r.v., uniformly distributed over $(0, a)$, $0 < a < 1$. If we normalize by defining $\epsilon = a/2\sqrt{3}$, the standard deviation ϵ of the r.v. ξ_t^ϵ can be used as the index parameterizing the level of perturbations.

Putting $F(k_t) = \mu k_t (1 - k_t)$ and $G(k_t, \xi_{t+1}^\epsilon) = -\xi_{t+1}^\epsilon F(k_t)$, equation (12) can be written in the form of (7) as

$$k_{t+1} = \mu k_t (1 - k_t) - \xi_{t+1}^\epsilon \mu k_t (1 - k_t). \tag{13}$$

The dynamic behavior of the map F is extremely well documented in the mathematical literature on maps of the interval and we refer the reader to it for details [see, e.g., Whitley (1983), Sharkovsky et al. (1997)]. Broadly speaking, the properties of attractors depend on the single parameter μ , which is a decreasing function of the discount factor. In particular, it is known that, for $\mu \in (1, 3)$, attractors are all fixed points; for $\mu \in (3, \mu_\infty)$, $\mu_\infty \approx 3.57$, attractors are all periodic; for $\mu \in (\mu_\infty, 4)$, there exist periodic, quasiperiodic, or chaotic attractors. We

have already discussed the simpler cases (fixed-point and periodic attractors) in Remarks 1 and 2, and the quasiperiodic case is taken up in Example 3. Here, we concentrate on the case of attractors that are chaotic in the following sense: there exists a unique, absolutely continuous, ergodic F -invariant probability measure ρ with support in Λ , with respect to which F has a positive metric entropy (and a positive Lyapunov exponent of equal value). There are two possibilities here: (i) ρ is mixing, which implies that F is topologically mixing on Λ [see Katok and Hasselblatt (1995, p. 151)] and Assumption 2 is satisfied; (ii) F is topologically transitive, but not topologically mixing on Λ , and we have the case of periodic, or nonmixing, chaos, discussed in Remark 2.

Assumption 4 (forward accessibility) is verified because (see Appendix B)

$$\frac{\partial}{\partial \xi_1^\epsilon} T(k_0, \xi_1^\epsilon) = -F(k_0) = -\mu k_0(1 - k_0) \neq 0, \quad (14)$$

which is true for initial conditions $0 < k_0 < 1$ (initial conditions $k_0 = 0$ or $k_0 = 1$ are excluded because, for $\mu_\infty < \mu < 4$, they are not in the basin of attraction of Λ). Finally, for initial conditions $0 < k_0 < 1$, $0 < F(k_t) < 1$, $\forall t$. This and the said properties of the r.v., ξ_t , guarantees that, for $k_t \in B(\Lambda)$, the perturbations term $\|G(k_t, \xi_{t+1}^\epsilon)\| = \|-\xi_{t+1}^\epsilon F(k_t)\|$ is bounded and goes to zero with ϵ , and Assumption 5 is verified.

3.2. Example 2: Type 2—Pure Exchange, One-Dimensional OLG Model

The deterministic core of this example is the pure exchange OLG model discussed by Samuelson (1958), Gale (1973), and many others. Here, we use a modified version of the Benhabib and Day (1982) model. In this model, there is no production, but each agent receives endowments of a single, perishable consumption good. Because we concentrate on Gale's "classical case" (young are impatient and borrow from the old), for simplicity's sake we assume that only old agents receive a constant endowment w . At each time t , the young agent borrows a certain amount c_t of the good and, when old, must pay back an amount $w - g_{t+1}$ (and will accordingly consume an amount g_{t+1} of the good) that depends on the exchange rate between present and future consumption R_{t+1} . Finally, at each t , the good market clears; that is, $w = c_t + g_t$. (We would also need some intergenerational arrangements guaranteeing that young people's debts are always settled, as well as some *ad hoc* rules for the "time zero" of the model, but we cannot discuss them here.)

In the stochastic case, agents believe that in equilibrium the exchange rate is perfectly correlated with an extrinsic i.i.d. sunspot, ξ_t^ϵ —that is, $R_t = R(\xi_t^\epsilon) \forall t$. We assume that ξ_t^ϵ is characterized by a probability distribution ν_ϵ with open, bounded support on the real line including $\{0\}$ and the bound goes to zero with ϵ (thus Assumption 5 is satisfied).

Assuming separable utility function, $U(c_t, g_{t+1}) = v(c_t) + u(g_{t+1})$, the young agent's program is

$$\begin{aligned} &\max v(c_t) + \mathbf{E}[u(g_{t+1}) \mid c_t] \\ &\text{s.t. } w - g_{t+1} = R(\xi_{t+1}^\epsilon)c_t, \text{ or} \\ &\max_{c_t} v(c_t) + \mathbf{E}[u(w - R(\xi_{t+1}^\epsilon)c_t)], \end{aligned}$$

where expectation is taken w.r.t. the measure ν_ϵ and it is conditional to c_t . If agents' beliefs are self-fulfilling and markets always clear, i.e., $c_t = w - g_t \forall t$, we must have

$$\mathcal{V}(c_t) + \mathbf{E}[\mathcal{U}(c_{t+1}) \mid c_t] = 0, \tag{15}$$

where $\mathcal{V}(c) = v'(c)c$ and $\mathcal{U}(c) = u'(c)c$. If \mathcal{U} is invertible (and putting $\mathbf{E}(\xi_{t+1}^\epsilon) = 0$ for simplicity's sake), condition (15) is satisfied for

$$c_{t+1} = T(c_t, \xi_{t+1}^\epsilon) = \mathcal{U}^{-1}[-(\mathcal{V}(c_t) + \xi_{t+1}^\epsilon)]. \tag{16}$$

Notice that the deterministic core of (16) is $F(c_t) = \mathcal{U}^{-1}[-\mathcal{V}(c_t)]$, corresponding to the perfect-foresight no-sunspot version of the pure exchange OLG model, and $G(c_t, \xi_{t+1}^\epsilon) = T(c_t, \xi_{t+1}^\epsilon) - F(c_t)$. Let us now choose the utility functions $v(c) = A - we^{-c}$; $u(g) = g$, corresponding, respectively, to constant absolute risk aversion and risk neutrality, with A a positive constant. In this case, $F(c_t) = wc_t e^{-c_t}$ and $G(c_t, \xi_{t+1}^\epsilon) = \xi_{t+1}^\epsilon$. F is a much-studied member of the class of unimodal maps of the interval. Notice that the interval $I_w = [0, w/e]$ is invariant for F and therefore, if $c_0 \in I_w$, $g_t = w - c_t$ will never become negative along a deterministic orbit. In the complete stochastic system, for every given w the level of perturbations (the parameter ϵ) must be fixed so that $g_t \geq 0$ at all times. For $w > 1$, F has two nonnegative fixed points, namely, $\bar{c}_1 = 0$, unstable, and $\bar{c}_2 = \ln w$, which is asymptotically stable (with improper oscillations) for $1 < w < e^2$. If the endowment is increased, at $w = e^2$, a flip bifurcation occurs, leading to an initially stable period-2 cycle and then, if w is increased further, to a cascade of bifurcations with (initially) stable cycles of increasing periods. Analytical and numerical studies of the map F [see, e.g., May and Oster (1976)] indicate that, for larger values of the endowment w , complex (chaotic) attractors will appear. Thus, the typology of attractors of the map F is similar to that occurring for the logistic map discussed in Example 1, and we could repeat what we said there with regard to Assumption 2. (However, notice that, in this case, very large values of w will lead back to simpler, periodic attractors). Other choices of the utility function $v(c)$, such as $v(c) = ac - bc^2$, or $u(c) = \lambda(c + b)^{1-a}/(1 - a)$ would lead to similar results [for details, see Benhabib and Day (1982)]. Notice that Assumption 4 is satisfied because $(\partial/\partial \xi_1^\epsilon)T(c_0, \xi_1^\epsilon) = 1 \neq 0$.

3.3. Example 3: Type 2—Two-Dimensional OLG Model with Production

This model differs from the one discussed in Example 2 in some important and related aspects concerning both technology and agents' behavior. First, the single good can be both consumed and invested (with depreciation equal to one period). There are no endowments and, at each time t , the single good is produced by current labor, supplied by young agents, and capital (output – consumption), invested at time $t - 1$. Second, only old agents consume.

To fix ideas, let us consider the case in which (i) the technology is represented by a Leontief linear production function, $y_t = \min[l_t, bk_{t-1}]$ (where y denotes output, $k = y - c$ is capital, c is consumption, l is labor, and, for viable systems, $b > 1$); (ii) young agents' preferences are represented by a separable utility function of a Cobb-Douglas type, $u(c_{t+1}, l_t) = (1/\alpha)c_{t+1}^\alpha - (1/\beta)l_t^\beta$, $0 < \alpha < 1 < \beta$; (iii) uncertainty is represented by means of random, i.i.d. perturbations of the future price of consumption. Under these assumptions, equilibrium dynamics of consumption and labor supply are described by the following stochastic dynamical system:

$$\begin{pmatrix} c_{t+1} \\ l_{t+1} \end{pmatrix} = \begin{bmatrix} (l_t^\beta + \xi_{t+1}^\epsilon)^{1/\alpha} \\ b(l_t - c_t) \end{bmatrix}, \tag{17}$$

where $\{\xi_t^\epsilon\}$ is an i.i.d. process and ξ_t^ϵ is a zero-mean r.v. with open, bounded support on the real line, including $\{0\}$ and with bound going to zero with ϵ . If we denote by x_t the two-dimensional vector variable (c_t, l_t) , equation (17) has the same form as (1). The RHS can be split into a deterministic part

$$F(x_t) = \begin{bmatrix} l_t^{\beta/\alpha} \\ b(l_t - c_t) \end{bmatrix} \tag{18}$$

corresponding to the perfect-foresight version of the model and a stochastic perturbation

$$G(x_t, \xi_{t+1}^\epsilon) = T(x_t, \xi_{t+1}^\epsilon) - F(x_t) = \begin{bmatrix} (l_t^\beta + \xi_{t+1}^\epsilon)^{1/\alpha} - l_t^{\beta/\alpha} \\ 0 \end{bmatrix}.$$

There exist two fixed points for the map F , namely, E_1 located at the origin and $E_2: \bar{c}_2 = (1 - 1/b)^{\beta/(\beta-\alpha)}$; $\bar{l}_2 = (1 - 1/b)^{\alpha/(\beta-\alpha)}$, in the positive orthant. E_1 is always unstable. Stability of E_2 depends on two parameters, that is, b , measuring productivity (output/capital ratio), and (β/α) , the ratio between the utility elasticities. For sufficiently low values of b and β/α , E_2 is locally, asymptotically stable and its basin of attraction is a forward invariant subset of $M = \{(c, l) \in \mathbf{R}^2 \mid c, l > 0 \text{ and } l > c\}$. If we increase either (or both) of those parameters, E_2 loses its stability through a Neimark–Sacker bifurcation leading to the appearance of an invariant, closed curve around E_2 [for details, cf. Reichlin (1986), Medio (1992, ch. 12)]. Numerical investigation suggests that for this model the curve is indeed stable. The dynamics *on the curve* can be periodic or quasiperiodic. Here, we concentrate on

the quasiperiodic case. It is known that, under quite general conditions, in this case the dynamics of the map F on the invariant curve is homeomorphically equivalent to the fixed rotation of the circle, described by the map $\theta : \mathbf{R}/\mathbf{Z} \rightarrow \mathbf{R}/\mathbf{Z}$,

$$\theta_{t+1} = f(\theta_t) = \theta_t + \rho \pmod{1}, \quad \rho \text{ irrational} \tag{19}$$

Consider first that, for ρ irrational, the map f is t.t. on the circle and so are all the maps f^k for arbitrary $k > 1$. Consequently, f is strongly t.t. on the circle and, because of the homeomorphic equivalence, so is F on the invariant attracting curve. Thus, Assumption 2 is verified. FA can now be ascertained by means of the techniques explained in Appendix B. The determinant of the “controllability matrix” C_{c_0, l_0}^2 , evaluated as in Appendix B, is

$$\det C_{(c_0, l_0)}^2 = \frac{b}{\alpha^2} \left(\{ [b(l_0 - c_0)]^\beta + \xi_2^\epsilon \} [l_0^\beta + \xi_1^\epsilon] \right)^{\frac{1-\alpha}{\alpha}}$$

From an economic point of view, we are interested in solutions that stay in the subset $M \subset \mathbf{R}^2$ defined before. Therefore, for any initial conditions $(l_0, c_0) \in M$, we can always choose values of $\xi_1^\epsilon, \xi_2^\epsilon$ belonging to the support of ξ_t^ϵ , such that $\det C_{(c_0, l_0)}^2 > 0$, $\text{rank } C_{(c_0, l_0)}^2 = 2$. Consequently FA is verified and Assumption 4 holds. Finally, notice that, given the assumptions on ξ_t^ϵ and for all (c_t, l_t) in the basin of attraction, the perturbation term G defined before is bounded and goes to zero with ϵ . Thus Assumption 5 is verified.

3.4. Example 4: Type 1—One-Dimensional Optimal Growth with Dependent Perturbations

This is a simple variation of the well-known one-good/one-sector model of optimal growth.⁴ Suppose that the return function is given by utility function $u(c_t) = \ln c_t$ and the production function is $x_t = f(k_t, \omega_t) = (e^{\omega_t} k_t^\alpha)$, $0 < \alpha < 1$, where c, x, k denote, respectively, consumption, output, and capital stock; $y_t = x_t - c_t$ is the control variable (saving) and $k_t = y_{t-1}$; and e^{ω_t} is a technology shock observed *after* the optimal choice of how much to save has been made. In this case, we assume that the sequence of shocks $\{\omega_t\}$ is generated by the equation $\omega_{t+1} = r\omega_t + \xi_{t+1}^\epsilon$, with $0 \leq r < 1$; $\{\xi_t^\epsilon\}$ is a zero-mean, i.i.d. process and ω_t and ξ_t^ϵ are independent. Therefore, the sequence $\{\omega_t\}$ is a Markov chain stationary but not independent for $r > 0$.

Under the stated assumptions, the Bellman equation can be written as

$$V(x, \omega) = \sup_{0 \leq y \leq x} \left\{ \ln(x - y) + \beta \int_{\Omega} V(x', \omega') \mu(\omega, d\omega') \right\}, \tag{20}$$

where Ω is the support of ω' and μ is its probability distribution, conditional on ω and, as usual, for a generic variable, we adopt the notation $x = x_t$ and $x' = x_{t+1}$. Let us now try a solution $V(x, \omega) = m \ln x + n_1 \omega + n_2$, with m, n_1, n_2 undetermined

coefficients. Substituting into (20), we have

$$m \ln x + n_1 \omega + n_2 = \sup_y \left\{ \ln(x - y) + \beta \int_{\Omega} [m \ln x' + n_1 \omega' + n_2] \mu(\omega, d\omega') \right\}, \tag{21}$$

whence, using the facts that at each point in time, (i) $x' = e^{\omega'} (k')^\alpha = e^{\omega'} y^\alpha$ and (ii) $\int_{\Omega} \omega' \mu(\omega, d\omega') = r\omega$, we can write

$$m \ln x + n_1 \omega + n_2 = \sup_y \{ \ln(x - y) + \beta [(m + n_1)r\omega + m\alpha \ln y + n_2] \}. \tag{22}$$

Next, finding the value of y that maximizes the RHS of (22) and solving for the unknown coefficients m , we obtain the policy function $y = g(x) = \alpha\beta x$. Finally, considering that for all t , $y_t = k_{t+1}$, we can write the stochastic difference equation of the model; that is,

$$x_{t+1} = e^{\omega_{t+1}} (\alpha\beta x_t)^\alpha \tag{23}$$

or, taking logarithms,

$$\ln x_{t+1} = C + \alpha \ln x_t + \omega_{t+1} = C + \alpha \ln x_t + r\omega_t + \xi_{t+1}^\epsilon, \tag{24}$$

where $C = \alpha \ln(\alpha\beta)$ is a constant. Moving (24) one period back in time, solving for ω_t , and substituting into (24), we obtain

$$\ln x_{t+1} = B + (\alpha + r) \ln x_t - \alpha r \ln x_{t-1} + \xi_{t+1}^\epsilon, \tag{25}$$

where $B = C(1 - r)$.

Equation (25) is a one-variable, second-order stochastic difference equation with an i.i.d. perturbation. By introducing appropriate auxiliary variables, it can be transformed into a dynamically equivalent two-variable, first-order equation. Putting $w_t = \ln x_t - \bar{w}$; $u_t = \ln x_{t-1} - \bar{w}$, where $\bar{w} = B/[1 - (\alpha + r) + \alpha r]$, we have

$$\begin{pmatrix} w_{t+1} \\ u_{t+1} \end{pmatrix} = J \begin{pmatrix} w_t \\ u_t \end{pmatrix} + \begin{pmatrix} \xi_{t+1}^\epsilon \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha+r & -\alpha r \\ 1 & 0 \end{pmatrix} \begin{pmatrix} w_t \\ u_t \end{pmatrix} + \begin{pmatrix} \xi_{t+1}^\epsilon \\ 0 \end{pmatrix}, \tag{26}$$

which has the same form as (1), with the RHS already split into a deterministic and a stochastic part, and it satisfies Assumption 1. It also provides an illustration of our comment in Remark 3 concerning the non-essentiality of the assumption of independence of perturbations.

The deterministic part of (26) is linear and it has a unique fixed point located at the origin (actually it corresponds to a positive value of the state variable, $\bar{x} = e^{\bar{w}}$). Under the stated assumptions, and, in particular, the fact that $0 < \alpha < 1$, $0 \leq r \leq 1$ (both absolutely reasonable in this context), the fixed point is always globally, exponentially asymptotically stable. Thus (26) satisfies (trivially) Assumption 2.

Assumption 4 (forward accessibility) can again be proved by applying the same technique as in Example 3. Choosing $k = 2$, the controllability matrix is

$$C^2_{w_0, u_0} = \begin{pmatrix} \alpha + r & 1 \\ 1 & 0 \end{pmatrix};$$

hence $\det(C^2_{w_0, u_0}) = -1 \neq 0$ and $\text{rank } C^2_{w_0, u_0} = 2$ as required. Finally, in view of the stability properties of the fixed point and the fact that the noise is additive, the results of Theorem 1 can be established under weaker conditions on the perturbing term ξ_t^ϵ . In particular, Assumption 5 can be replaced by the condition $\mathbf{E} \|\xi_t^\epsilon\| < \infty$ [for details, see Tong (1990), pp. 127–129].

4. A NOTE ON THE LITERATURE

Apart from the general reference to the book by Meyn and Tweedie, mathematically the obvious reference is to Chan and Tong, to whose work [see, Tong (1990), in particular its Appendix 1, written by Chan; Chan and Tong (1994)], we would like to acknowledge our intellectual debt. In particular, these authors proved uniform ergodicity of a Markov chain under hypotheses similar to Assumptions 1 and 3–5 above [see Chan and Tong (1994)]. However, they assumed *exponential* stability of the deterministic attractor, an assumption considerably stronger than Assumption 2 and unnecessarily restrictive. As will be seen in Appendix A, our proof of Theorem 1 (and the choice of the Lyapunov function) is accordingly different.

Moreover, Chan and Tong do not relate these results on stochastic stability to the economic literature in general, or to the sunspot question in particular, and so we shall do it here briefly.

Consider first that, for irreducible, aperiodic Markov chains, uniform ergodicity is equivalent (it implies and is implied by) the celebrated Doeblin condition. This condition requires that there exist a (finite-valued) measure ϕ on $\mathcal{B}(M)$, an integer $n \geq 1$, and a positive δ , such that

$$P^n(x, A) \leq 1 - \delta \quad \text{if} \quad \phi(A) \leq \delta$$

for every $x \in M$. Roughly speaking, this requirement means that there exists a measure ϕ such that the process X is not concentrated on ϕ -small sets.⁵

Uniform ergodicity (or, equivalently, the Doeblin condition) also implies that the operator $\mathcal{P} : \mathcal{M} \rightarrow \mathcal{M}$, defined by

$$\mathcal{P}[\mu(A)] = \int_M P(x, A)\mu(dx) \tag{27}$$

(where \mathcal{M} is the space of finite probability measures) is quasi compact. Hence, uniformly ergodic chains are sometimes called quasi compact. A formal definition and a thorough discussion of the concepts of compactness and quasi compactness of operators is out of the question here. Broadly speaking, we can say that if the operator \mathcal{P} on \mathcal{M} is quasi compact and the associated Markov chain is weak Feller

(which in the case of our model is implied by Assumptions 1 and 3), then the sequences generated by \mathcal{P} , starting from any initial condition on \mathcal{M} , converge to a unique invariant probability measure.

The relation between the Doeblin condition and quasi compactness of operators, and the associated property of convergence of probability measures, was introduced in the economic literature by Carl Futia's (1982) excellent mathematical survey. Several applications of these ideas to economic problems are discussed by Stokey and Lucas (1989, chs. 11–13). Futia's results were also employed in the sunspot literature (cf. Farmer and Woodford (1997), originally circulated in 1984 as a CARESS W.P.; Chiappori and Guesnerie (1991), p. 1708).

A sufficient condition for uniform ergodicity of an irreducible, aperiodic T-chain is that the state space can be reduced to a compact invariant set [see Meyn and Tweedie (1993, Th. 16.2.5, p. 395)]. The compact set argument was apparently introduced in the economic literature by Blume (1982), generalized by Duffie et al. (1994), and applied in a number of contexts, including the sunspot models [cf. Chiappori and Guesnerie (1991, pp. 1708–1710)]. See also Stokey and Lucas (1989, ch. 12).

The distinctive advantage of the approach adopted in this paper is that we do not *assume* compactness of the state space (or any other equivalent condition), but deduce uniform ergodicity from assumptions directly concerning the deterministic core of the dynamical system, on the one hand, and its stochastic perturbations, on the other. In the context of the sunspots problem, this means that we relate those assumptions to the properties of the perfect-foresight equilibrium and the agents' random beliefs, respectively.

NOTES

1. Notice that, here, $\|\cdot\|$ denotes the *total variation norm* defined as

$$\|\mu\| \equiv \sup_{A \in \mathcal{B}(O_\varepsilon)} \mu(A) - \inf_{A \in \mathcal{B}(O_\varepsilon)} \mu(A).$$

Convergence in the total variation norm implies weak convergence. We say that a sequence of probability measures $\{\mu_k\}$ converges weakly to μ if, for any bounded continuous function f , we have

$$\lim_{k \rightarrow \infty} \int f d\mu_k = \int f d\mu$$

[cf. Meyn and Tweedie (1993, pp. 311, 521)].

2. A *transient* set E is negligible in the sense that the expected number of times that an infinite chain X starting in E returns to it is finite.

3. For further details on this point, see Doob (1953, pp. 190–218) and Stokey and Lucas (1989, pp. 334–351).

4. Here we use the version discussed by Cugno and Montrucchio (1998, pp. 178–179).

5. The Doeblin condition, under the name of “Hypothesis D,” and its implications are extensively discussed by Doob (1953, pp. 190–234). See also Stokey and Lucas (1989, pp. 344–351), and Meyn and Tweedie (1993, Th. 16.0.2, pp. 384–385). (Meyn and Tweedie, however, use a somewhat different definition of the Doeblin condition, which is equivalent to the traditional one for irreducible Markov chains.)

6. Meyn and Tweedie do not give a full proof of Proposition 7.1.5, but it can be found in Meyn and Caines (1991).

7. Our converse lemma is an extension of Gordon's converse theorem (Gordon, 1972, Th. 3, p. 79). The main difference is that Gordon considers stability of a fixed point and we deal with the more general case of a compact attractor. (Of course, our dynamical system is autonomous and Gordon's is not, but this is irrelevant here.) We do not want to insist too much on the originality of this result. It is possible that a generalization of Gordon's theorem exists already in the very vast mathematical literature on stability, but we could not find it.

8. This lemma is mentioned by Gordon in the quoted article, but its source is misquoted. See also Hahn (1963, p. 70).

9. These are conditional expectations for a given value of $x_t \in O_\epsilon$. Therefore, we have

$$E[V(x_{t+1})] = E \left\{ V \left[T \left(x_t, \xi_{t+1}^\epsilon \right) \right] \right\} = \int_{W_\epsilon} V \left[T \left(x_t, \xi_{t+1}^\epsilon \right) \right] v_\epsilon(d\xi^\epsilon).$$

10. The fact that, in Meyn and Tweedie's definition of Condition V(4), the range of $V(x)$ is $[1, \infty)$, rather than $[0, \infty)$, is irrelevant here.

REFERENCES

- Benhabib, J. (ed.) (1992) *Cycles and Chaos in Economic Equilibrium*. Princeton, NJ: Princeton University Press.
- Benhabib, J. & R.H. Day (1982) A characterization of erratic dynamics in the overlapping generations model. *Journal of Economic Dynamics and Control* 4, 37–55.
- Blume, L.E. (1982) New techniques for the study of stochastic equilibrium processes. *Journal of Mathematical Economics* 9, 71–82.
- Boldrin, M. & L. Montrucchio (1986) On the indeterminacy of capital accumulation paths. *Journal of Economic Theory* 40, 26–39.
- Boldrin, M. & M. Woodford (1990) Equilibrium models displaying endogenous fluctuations and chaos: A survey. *Journal of Monetary Economics* 25, 189–223.
- Brock, W.A. & M. Majumdar (1978) Global asymptotic stability results for multisector models of optimal growth under uncertainty when future utilities are discounted. *Journal of Economic Theory* 18, 225–243.
- Brock, W.A. & L.J. Mirman (1972) Optimal economic growth and uncertainty: the discounted case. *Journal of Economic Theory* 4, 479–513.
- Chan, K.S. & H. Tong (1994) A note on noisy chaos. *Journal of the Royal Statistical Society B* 56, 301–311.
- Chiappori, P.A. & R. Guesnerie (1991) Sunspot equilibria in sequential markets models. In W. Hildenbrand & H. Sonnenschein (eds.), *Handbook of Mathematical Economics*, vol. IV, pp. 1683–1762. New York: Elsevier.
- Cugno, F. & L. Montrucchio (1998) *Scelte Intertemporal: Teoria e Modelli*. Rome: Carocci Editore.
- Deneckere, R. & S. Pelikan (1986) Competitive chaos. *Journal of Economic Theory* 40, 13–25.
- Doob, J.L. (1953) *Stochastic Processes*. New York: John Wiley & Sons.
- Duffie, D., J. Genakoplos, A. Mas-Colell, & A. MacLennan (1994) Stationary Markov equilibria. *Econometrica* 62, 745–783.
- Farmer R.E.A. & M. Woodford (1997) Self-fulfilling prophecies and the business cycle. *Macroeconomic Dynamics* 1, 740–770.
- Futia, C. (1982) Invariant distributions and the limiting behavior of Markov economic models. *Econometrica* 50, 377–407.
- Gale, D. (1973) Pure exchange equilibrium of dynamic economic model. *Journal of Economic Theory* 6, 12–36.

- Gordon, S.P. (1972) On converses of the stability theorems for difference equations. *SIAM Journal of Control* 10, 76–81.
- Hahn, W. (1963) *Theory and Applications of Liapunov's Direct Method*. Englewood Cliffs, NJ: Prentice-Hall Int.
- Hahn, W. (1967) *Stability of Motion*. Berlin: Springer-Verlag.
- Joshi, S. (1995) Recursive utility and optimal growth under uncertainty. *Journal of Mathematical Economics* 24, 601–617.
- Katok, A. & B. Hasselblatt (1995) *Introduction to the Modern Theory of Dynamical Systems*. Cambridge, UK: Cambridge University Press.
- Majumdar, M. & I. Zilcha (1987) Optimal growth in a stochastic environment: Some sensitivity and turnpike results. *Journal of Economic Theory* 43, 116–133.
- Massera, J.L. (1949) On Liapounoff's condition of stability. *Annales of Mathematics (2)* 50, 705–721.
- May, R.M. & G.F. Oster (1976) Bifurcations and dynamic complexity in simple ecological models. *American Naturalist* 110, 573–594.
- Medio, A. (1992) *Chaotic Dynamics. Theory and Applications to Economics*. Cambridge, UK: Cambridge University Press.
- Meyn, S.P. & P.E. Caines (1991) Asymptotic behavior of stochastic systems processing Markov realizations. *SIAM Journal of Control and Optimization* 29, 535–561.
- Meyn, S.P. & R.L. Tweedie (1993) *Markov Chains and Stochastic Stability*. London: Springer-Verlag.
- Montrucchio, L. (1986) Optimal decisions over time and strange attractors: An analysis by the Bellman Principle. *Mathematical Modelling* 7, 341–352.
- Radner, R. (1973) Optimal stationary consumption with stochastic production and resources. *Journal of Economic Theory* 6, 68–90.
- Reichlin, P. (1986) Equilibrium cycles and stabilization policies in an overlapping generations models with production. *Journal of Economic Theory* 40, 89–102.
- Samuelson, P.A. (1958) An exact consumption–loan model of interest with or without the social contrivance of money. *Journal of Political Economy* 66, 467–482.
- Sharkovsky, A.N., S.F. Kolyada, A.G. Sivak, & V.V. Fedorenko (1997) *Dynamics of One-Dimensional Maps*. Dordrecht, The Netherlands: Kluwer Academic.
- Stokey, N.L. & R.E. Lucas, Jr. (1989) *Recursive Methods in Economic Dynamics*. Cambridge, MA: Harvard University Press.
- Tong, H. (1990) *Non-Linear Time Series. A Dynamical System Approach*. New York: Oxford University Press.
- Tuominen P. & R.L. Tweedie (1979) Markov chains with continuous components. *Proceedings of the London Mathematical Society (3)* 38, 89–114.
- Whitley, D. (1983) Discrete dynamical systems in dimensions one and two. *Bulletin of London Mathematical Society* 15, 177–217.

APPENDIX A

A.1. PROOF OF THEOREM 1

Step 1. The Regularity Assumption 1, the Density Assumption 3, and the Forward accessibility Assumption 4, together imply that the Markov chain X is a T-chain {cf. Meyn and Tweedie [1993, Prop. 7.1.2, pp. 152–154 (scalar case) and 7.1.5, p. 157 (multidimensional case)]}.⁶

Step 2. From step 1 and the assumptions,

- (i) the set Λ is asymptotically stable and the map F of the deterministic system (9) is strongly topologically transitive on Λ (Assumption 2),

- (ii) $\{0\}$ is in the support of the (density of the) r.v. ξ_t^ϵ (Assumption 3), and
- (iii) the ACS is “forward accessible” (Assumption 4),

we conclude that any point $x \in \Lambda$ is *reachable*, in the sense that for every neighborhood N of x , $\sum_n P_\epsilon^n(y, N) > 0$, $y \in B(\Lambda)$. But from Meyn and Tweedie (1993, Prop. 6.2.1, p. 133), we gather that, if X is a T -chain and the state space contains a reachable point x^* , then X is ψ -irreducible with $\psi = T(x^*, \cdot)$. Aperiodicity follows from the above argument as well as the second part of Assumption 2 (F is strongly topologically transitive on Λ). This implies that, for any integer m , the Markov chain $X^{(m)}$ associated with the kernel P^m is also ψ -irreducible. Thus, the Markov chain X is aperiodic.

Step 3. The fact that X is a ψ -irreducible, aperiodic T-chain also implies that every compact set is petite [cf. Meyn and Tweedie (1993, Th. 6.2.5, p. 134)].

Step 4. To complete our proof, we now need to prove the following auxiliary lemma.⁷

LEMMA 1 (*Converse Stability Lemma*). *If a compact subset of the state space $\Lambda \subset M$ is uniformly asymptotically stable for the deterministic dynamical system (9), then there exists a real scalar function $V(x)$ which satisfies the following properties for x in any bounded set $U \subseteq B(\Lambda)$:*

- (a) $V(x)$ is positive definite if $x \notin \Lambda$,
- (b) $V(x)$ admits an infinitesimally small upper bound,
- (c) $V(x)$ is Lipschitzian,
- (d) $\Delta_{\text{DET}} V(x_t) \equiv [V(F(x_t)) - V(x_t)] = [V(x_{t+1}) - V(x_t)]$ is negative definite if $x \notin \Lambda$.

Proof. First, we need to recall a lemma by Massera (1949, pp. 716–717),⁸ which, with some inessential modifications, states that, for a scalar function $\sigma(t) > 0$, defined for $t \geq 0$ such that $\lim_{t \rightarrow +\infty} \sigma(t) = 0$, and a scalar function $\rho(t)$, positive, continuous, and nondecreasing for $t \geq 0$, there exists a positive, continuous, and increasing function $\Gamma(r)$, defined for $r \geq 0$, with continuous, increasing derivative, $\Gamma(r) > 0$ for $r > 0$ and $\Gamma(0) = \Gamma'(0) = 0$, and such that for any fixed number $c > 0$ and any function $\sigma^*(t)$ satisfying the inequality $0 < \sigma^*(t) \leq c\sigma(t)$, the integrals

$$\int_0^\infty \Gamma[\sigma^*(t)] dt \quad \text{and} \quad \int_0^\infty \Gamma'[\sigma^*(t)]\rho(t) dt \tag{A.1}$$

converge. Notice that $\Gamma \in C^1$ and therefore it is Lipschitzian on a bounded set.

This result can be extended to the discrete-time case [cf. Hahn (1967, p. 235) and Gordon (1972, p. 78)]. If $\sigma^*(t)$ is a nonincreasing function, the application of the Massera Lemma and the integral test for convergence of a series guarantees the uniform convergence of the sums

$$\sum_{t=0}^\infty \Gamma[\sigma^*(t)] \quad \text{and} \quad \sum_{t=0}^\infty \Gamma'[\sigma^*(t)] \tag{A.2}$$

where we put $\rho(t) = 1$, Γ is the same as in (A.1), but we now have $t \in \mathbf{Z}_+$.

Next, let us consider that the scalar function $\sup_{s \geq t} \{d[F^s(x), \Lambda]\}$ (where F^s denotes the s th iterate of F) is defined for $t \geq 0$, positive for $x \notin \Lambda$, and nonincreasing. Moreover, because the set Λ is uniformly asymptotically stable for any initial condition $x \in B(\Lambda)$, that function converges to 0 as $t \rightarrow \infty$. If we now put $c = 1$, $\sigma(t) = \sigma^*(t) = \sup_{s \geq t} \{d[F^s(x), \Lambda]\}$, from the Massera Lemma and the integral test for sequences, we conclude that the infinite

sums

$$\sum_{t=0}^{\infty} \Gamma \left(\sup_{s \geq t} \{d[F^s(x), \Lambda]\} \right) \quad \text{and} \quad \sum_{t=0}^{\infty} \Gamma' \left(\sup_{s \geq t} \{d[F^s(x), \Lambda]\} \right)$$

converge.

Let us now define the scalar function

$$V(x) = \sum_{t=0}^{\infty} \frac{1}{k^t} \Gamma\{d[F^t(x), \Lambda]\} \tag{A.3}$$

Obviously, $d[F^t(x), \Lambda] \leq \sup_{s \geq t} \{d[F^s(x), \Lambda]\}$, and therefore, for any scalar $k \geq 1$, (A.3) converges.

Moreover, for any two values $x_1, x_2 \in U$, we have

$$\begin{aligned} |V(x_1) - V(x_2)| &= \left| \sum_{t=0}^{\infty} \frac{1}{k^t} \Gamma\{d[F^t(x_1), \Lambda]\} - \sum_{t=0}^{\infty} \frac{1}{k^t} \Gamma\{d[F^t(x_2), \Lambda]\} \right| \\ &\leq \sum_{t=0}^{\infty} \frac{1}{k^t} |\Gamma\{d[F^t(x_1), \Lambda]\} - \Gamma\{d[F^t(x_2), \Lambda]\}| \\ &\leq \sum_{t=0}^{\infty} \frac{M_{\Gamma}}{k^t} |d[F^t(x_1), \Lambda] - d[F^t(x_2), \Lambda]| \\ &\leq \sum_{t=0}^{\infty} \frac{M_{\Gamma}}{k^t} \|F^t(x_1) - F^t(x_2)\| \\ &\leq \sum_{t=0}^{\infty} \frac{M_{\Gamma} M_F^t}{k^t} \|x_1 - x_2\|, \end{aligned}$$

where M_{Γ}, M_F are the Lipschitz constants of the functions Γ and F , respectively. We have used here the known fact that, for a given set $A \subset \mathbf{R}^n$, the map $d(x, A) : \mathbf{R}^n \rightarrow \mathbf{R}$ is Lipschitzian with Lipschitz constant equal to one. If we now put

$$k = \max(1, M_F),$$

the infinite sum $\sum_{t=0}^{\infty} [(M_{\Gamma} M_F^t)/k^t]$ converges and we can write

$$|V(x_1) - V(x_2)| \leq M \|x_1 - x_2\|,$$

where $M < \infty$ is a constant. Property (c) then obtains.

If we now choose x_2 as the point in Λ closest to x_1 , we shall have $d(x_2, \Lambda) = 0, V(x_2) = 0, d(x_1, \Lambda) = \|x_1 - x_2\|$ and

$$V(x_1) \leq M d(x_1, \Lambda) \tag{A.4}$$

and property (b) follows.

Moreover, consider that, obviously, $V(x) \geq \Gamma[d(x, \Lambda)]$ and property (a) follows. Finally, consider that $\Delta_{\text{DET}} V(x_t) = -\Gamma[d(x_t, \Lambda)] < 0$ for $x \notin \Lambda$, which establishes property (d).

Step 5. Applying the result just established to the complete, stochastic system (7), we have almost surely

$$\begin{aligned}
 V(x_{t+1}) - V(x_t) &= V[F(x_t) + G(x_t, \xi_{t+1}^\epsilon)] - V(x_t) \\
 &= V[F(x_t) + G(x_t, \xi_{t+1}^\epsilon)] - V[F(x_t)] + \Delta_{\text{DET}}V(x_t) \\
 &\leq \Delta_{\text{DET}}V(x_t) + M \|G(x_t, \xi_{t+1}^\epsilon)\| \\
 &= -\Gamma[d(x_t, \Lambda) + M \|G(x_t, \xi_{t+1}^\epsilon)\|]
 \end{aligned}
 \tag{A.5}$$

where $M < \infty$ is a constant. If the perturbation term $G(x_t, \xi_{t+1}^\epsilon)$ is sufficiently small (i.e., ϵ is sufficiently small), we can find a positive constant $\delta_\epsilon > 0$ (which can be made as small as we please by reducing ϵ) such that

$$V(x_{t+1}) - V(x_t) \leq -\Gamma(\delta_\epsilon) + M \|G(x_t, \xi_{t+1}^\epsilon)\| < 0.
 \tag{A.6}$$

From (A.4), it follows that $[V(x) > M\delta_\epsilon] \Rightarrow [d(x, \Lambda) > \delta_\epsilon]$. Then, from (A.6) and the fact that Γ is increasing in its argument, it follows that, for sufficiently small ϵ , we can find a positive scalar constant $\eta_\epsilon > M\delta_\epsilon$ such that the set $O_\epsilon \equiv \{x \mid V(x) \leq \eta_\epsilon\} \subset B(\Lambda)$ is absorbing; η_ϵ , too, can be made arbitrarily small by reducing ϵ .

We can now show that, for the Markov chain constrained to the absorbing set O_ϵ , the function $V(x)$ defined above satisfies the so-called “geometric drift condition” for stochastic stability [Condition V(4), Meyn and Tweedie (1993, pp. 255, 367)]. Taking expectations⁹ on both sides of equation (A.6), we can write

$$\Delta V(x) \equiv E[V(x_{t+1}) - V(x_t)] \leq -\Gamma[d(x, \Lambda)] + ME[\|G(x, \xi_{t+1}^\epsilon)\|].
 \tag{A.7}$$

Consider now that, in view of Assumption 5, $ME[\|G(x, \xi_{t+1}^\epsilon)\|] < a_\epsilon < \infty$. Next, define the scalar quantity $\beta_\epsilon = [\Gamma(\delta_\epsilon) - a_\epsilon]/\eta_\epsilon$ and the set $C_\epsilon = \{x \mid d(x, \Lambda) \leq \delta_\epsilon\}$, and notice that: (i) we can always fix ϵ (and thereby δ_ϵ and a_ϵ) so that $\beta_\epsilon > 0$, and (ii) C_ϵ is compact and therefore, because X is a T-chain, petite (cf. Step 3 above). Hence, $V(x)$ satisfies the Condition V(4),¹⁰ namely,

$$\Delta V(x) \leq -\beta_\epsilon V(x) + b_\epsilon \chi_{C_\epsilon}(x),$$

where χ_{C_ϵ} denotes the characteristic function of the set C_ϵ and $b_\epsilon < \infty$. Because C_ϵ is petite, from Meyn and Tweedie (1993, Lemma 15.2.2, p. 367), we deduce that the function V is unbounded off petite set; that is, the “level set,” O_ϵ defined above is also petite.

Then, from Meyn and Tweedie (1993, Th. 16.2.2, pp. 390–391, establishing that a ψ -irreducible, aperiodic Markov chain is uniformly ergodic if (and only if) the state space is petite, we conclude that the Markov chain X generated by (7), restricted to the absorbing set O_ϵ , is uniformly ergodic. This proves point (a) of the Theorem. Point (b) of the Theorem follows because Meyn and Tweedie (1993, Th. 16.0.2, pp. 384–385) proves that (a) is equivalent to (b) (the former implies and is implied by the latter).

Finally, point (c) of the Theorem follows from Meyn and Tweedie (1993, Th. 13.0.1, pp. 309–310; A. 3, p. 500; and Th. 13.3.3, p. 323), establishing that for an aperiodic, ergodic (“positive Harris”) Markov chain, any initial probability measure converges in the total variation norm to the unique, finite invariant measure. ■

APPENDIX B

In this Appendix, we provide some additional explanations of the concept of forward accessibility and briefly discuss a method for ascertaining FA in simple models, such as those of Examples 1–4. For further technical details and proofs, see Meyn and Tweedie (1993, pp. 150–155). Unexplained symbols and notions are as in the main text of the paper.

Let us first recall from the text that to each system of stochastic difference equations like (1) we can associate a control system (ACS) whose trajectories are determined by the equations

$$x_k = T_k(x_0, \xi_1^\epsilon, \dots, \xi_k^\epsilon), \tag{B.1}$$

where $x_k \in \mathbf{R}^n$ is the state reached at time k starting from x , given a certain sequence $\{\xi_1^\epsilon, \dots, \xi_k^\epsilon\}$. Therefore,

$$A_+^k(x) = \left\{ T_k(x, \xi_1^\epsilon, \dots, \xi_k^\epsilon) \mid \xi_t^\epsilon \in W_\epsilon, 1 \leq t \leq k \right\}$$

defines the set of all the states reachable from an initial state x for any admissible sequence of perturbations, and

$$A_+(x) = \bigcup_{k=0}^{\infty} A_+^k(x)$$

is the set of all states reachable from x at some time in the future.

Consider now equation (B.1) and the $(n \times k)$ matrix $C_{x_0}^k$ of the partial derivative of the function T_k with respect to its ξ -arguments.

Then [cf. Meyn and Tweedie (1993, Prop. 7.1.4, p. 156)], a necessary and sufficient condition for an ACS defined by (B.1) to be FA is that, for each initial condition $x_0 \in M \subset \mathbf{R}^n$, there exists an integer $k \geq 1$ and a sequence of given values $\{\xi_1^\epsilon, \dots, \xi_k^\epsilon\} \in W_\epsilon^k$ such that

$$\text{rank } C_{x_0}^k(\xi_1^\epsilon, \dots, \xi_k^\epsilon) = n$$

The matrix $C_{x_0}^k$ is called the controllability matrix.

For scalar models, that is, those for which $M \subset \mathbf{R}$, the “rank condition” for the associated ACS reduces to

$$\left[\frac{\partial}{\partial \xi_1^\epsilon} T_k(x_0, \xi_1^\epsilon, \dots, \xi_k^\epsilon), \dots, \frac{\partial}{\partial \xi_k^\epsilon} T_k(x_0, \xi_1^\epsilon, \dots, \xi_k^\epsilon) \right] \neq 0.$$

Notice that, to prove the existence of the rank condition, and thereby FA, we must find an integer $k \geq 1$ and an admissible sequence such that the condition holds for *any* initial condition.

Let us conclude by actually calculating the controllability matrix for the two-dimensional model discussed in Example 3—overlapping generations model with production.

If we now choose $k = 2$, from equation (17) we have

$$\begin{pmatrix} c_2 \\ l_2 \end{pmatrix} = T_2 \left[\begin{pmatrix} c_0 \\ l_0 \end{pmatrix}, \begin{pmatrix} \xi_1^\epsilon \\ 0 \end{pmatrix}, \begin{pmatrix} \xi_2^\epsilon \\ 0 \end{pmatrix} \right] = \begin{bmatrix} \{ [b(l_0 - c_0)]^\beta + \xi_2^\epsilon \}^{1/\alpha} \\ b^2(l_0 - c_0) - b(l_0^\beta + \xi_1^\epsilon)^{1/\alpha} \end{bmatrix}. \tag{B.2}$$

Therefore, the controllability matrix is the (2×2) matrix

$$C_{(c_0, l_0)}^2 = \begin{pmatrix} 0 & \frac{1}{\alpha} \{ [b(l_0 - c_0)]^\beta + \xi_2^\epsilon \}^{\frac{1-\alpha}{\alpha}} \\ -\frac{b}{\alpha} (l_0^\beta + \xi_1^\epsilon)^{\frac{1-\alpha}{\alpha}} & 0 \end{pmatrix}$$

and $\text{rank } C_{(c_0, l_0)}^2 = n = 2$, iff $\det C_{(c_0, l_0)}^2 \neq 0$.