ACCURACY OF SIMULATIONS FOR STOCHASTIC DYNAMIC MODELS

By Manuel S. Santos and Adrian Peralta-Alva[1]

This paper is concerned with accuracy properties of simulations of approximate solutions for stochastic dynamic models. Our analysis rests upon a continuity property of invariant distributions and a generalized law of large numbers. We then show that the statistics generated by any sufficiently good numerical approximation are arbitrarily close to the set of expected values of the model's invariant distributions. Also, under a contractivity condition on the dynamics we establish error bounds. These results are of further interest for the comparative study of stationary solutions and the estimation of structural dynamic models.

Keywords: Stochastic Dynamic Model, Invariant Distribution, Numerical Solution, Approximation Error, Simulated Moments, Convergence.

1 Introduction

Economists perform computational experiments to analyze quantitative properties of equilibrium solutions, but relatively little is known about the biases that may result from numerical approximations. The purpose of this paper is to set forth easily verifiable conditions for the accuracy of statistics obtained from numerical simulations. This work provides theoretical foundations for the widespread use of numerical techniques in the simulation and estimation of dynamic economic models.

In the simulation of dynamic models approximation errors may cumulate over time, and hence they may change drastically the evolution of sample paths. In order to establish accuracy properties of the statistics obtained from numerical simulations we address the

[1]We have benefitted from several discussions with various seminar participants at Arizona State University, Universidad Carlos III of Madrid, the University of Minnesota, the Federal Reserve Bank of Minneapolis, and the SED meetings (Paris, June 2003). We are grateful to Jorge Aseff and to three anonymous referees for several detailed comments to an earlier draft. This research was partially supported by the Spanish Ministerio de Ciencia y Tecnologia, Grant SEC 2002-4318.

following issues: (i) Effects of numerical errors on the asymptotic equilibrium dynamics, and (ii) laws of large numbers for Markovian stochastic processes. Our results are relevant for most quantitative studies using calibration methods. An early example of this literature is Kydland and Prescott (1982) on the quantitative effects of productivity shocks on the business cycle. Calibration methods span now many areas of economics. Recent contributions include Chari, Kehoe and McGrattan (2002), and Arellano and Mendoza (2005) in international finance; Boldrin, Christiano and Fisher (2001), and Cooley and Quadrini (2004a) in financial economics; Castaneda, Diaz-Gimenez and Rios-Rull (2003), Restuccia and Urrutia (2004), and De Nardi (2004) in wealth and income inequality; and Cooley and Quadrini (2004b), and Christiano, Eichembaum and Evans (2005) in monetary economics. A major methodological problem of this literature is to presume –without further justification– that the statistics generated from computer simulations are sufficiently close to the true values of an invariant distribution of the original, calibrated model.

Our results are also of direct application to a growing body of research that uses numerical simulation for the estimation of structural dynamic models. These models have become central in various areas of economics such as industrial organization, health economics, development, demography, and public finance. Among the many examples of simulation-based estimation, let us mention Wolpin (1984), Pakes (1986), Rust (1987), Bajari, Benkard and Levin (2004), Tan (2004), and Crawford and Shum (2005). Numerical errors may lead to biases in parameter estimates. Therefore, without further control of the approximation error it becomes awkward to proceed to model estimation by numerical simulation. Several methods have been put forward for simulation-based estimation: The method of simulated moments [e.g., Duffie and Singleton (1993)], score methods [e.g., Gallant and Tauchen (1996)], indirect inference [e.g., Gourieroux and Monfort (1996)] and the likelihood function [e.g., Fernandez-Villaverde and Rubio-Ramirez (2004)]. Our results are just a first step to derive asymptotic properties of simulation-based estimators from primitive assumptions on economic models.

Our analysis rests upon a continuity property of the correspondence of invariant distributions and a generalized law of large numbers. The continuity property of the correspondence of invariant distributions implies that quantitative properties of invariant distributions of a sufficiently good numerical approximation are close to those of the original model. This result justifies the use of approximation methods to explore steady-state predictions of dynamic models. Further, under a simple contractivity property of the equilibrium solution, we establish upper bounds. These bounds hold globally for any given numerical approximation (even if this approximation is not close to the true solution of the model). We show that the errors in the moments are proportional to the expected error of the computed solution; moreover, the constants bounding these errors can be easily estimated. These theoretical bounds can significantly be sharpened for solution methods that are already known to work well along a "typical" sample path.

Available laws of large numbers for non-linear dynamical systems rely on Doeblin's condition and its various generalizations [e.g., Hypothesis D in Doob (1953), Chapter V, and Stokey, Lucas and Prescott (1989), Chapter 11]. These conditions are difficult to check for both the economic model and its numerical approximations. We shall establish a generalized law of large numbers that dispenses completely with such technical conditions. Our method of proof builds on some fundamental results developed by Crauel (1991, 2002) on convergence properties of invariant distributions for finite-horizon approximations. These methods do not translate directly into a law of large numbers for Markovian distributions, and hence some further technical work is involved in our contribution. We also allow for some types of discontinuities that arise naturally in various economic settings whereas Crauel assumes that the dynamical system is continuous in the vector of state variables.

The main assumption underlying our results is a weak continuity property on the exact and numerical solutions. As discussed below, this assumption is indispensable. A wide variety of dynamic economic models satisfies our continuity condition; and so for these models, it seems adequate to approximate quantitative properties of invariant distributions by numerical methods. There are some other models, however, that may lack continuous Markovian solutions and for which our accuracy results may not hold. Non-continuous Markovian solutions are a common property of dynamic games and contracts, and of competitive economies with heterogeneous agents, incomplete financial markets, taxes, externalities, and other market distortions [e.g., Duffie *et al*. (1994) and Santos (2002)].

A considerable part of the literature in computational economics has focused on the performance of numerical algorithms [e.g., see Taylor and Uhlig (1990), Judd (1998), Santos (1999), and Aruoba, Fernandez-Villaverde and Rubio-Ramirez (2003)]. These algorithms are evaluated in some test cases or by using some accuracy checks [e.g., den Haan and Marcet (1994) and Santos (2000)]. But given the complex dynamic behavior involved in random sample paths, error bounds on the computed solution are generally not operative to assess steady-state predictions of stochastic dynamic models. What seems to be missing is a theory that links the approximation error of the numerical solution to the corresponding error in the statistics generated by this solution. In the absence of this theory, error bounds or accuracy checks on the numerical solution fall short of what is required for the calibration, estimation, and testing of an economic model.

In the next section we describe our framework of analysis. We consider a stochastic Markov process that may represent the equilibrium law of motion of a dynamic economic model or a numerical approximation. In Section 3, we establish an upper semicontinuity property for the correspondence of invariant distributions and a generalized law of large numbers. These results imply that the simulated statistics from a good enough numerical approximation are close to the corresponding expected values of the model's invariant distributions. In Section 4, we derive error bounds for these approximations under an additional contractivity condition on the dynamics. In Section 5, we present several examples to illustrate the role of the assumptions and the applicability of our analysis to economic models. We conclude in Section 6 with a further discussion of our main findings. Proofs of our main results are contained in the Appendix.

2 Stochastic Dynamics

In several economic models [e.g. Stokey, Lucas and Prescott (1989), Rust (1994), Fernandez-Villaverde and Rubio-Ramirez (2004), and Fershtman and Pakes (2005)] the equilibrium law of motion of the state variables can be specified by a dynamical system of the following form

$$\begin{aligned} z_{n+1} &= \Psi(z_n, \varepsilon_{n+1}) \\ k_{n+1} &= g(z_n, k_n, \varepsilon_{n+1}), \qquad n = 0, 1, 2, \cdots . \end{aligned} \tag{2.1}$$

Here, z is a finite vector made up of exogenous stochastic variables such as some indices of factor productivity or international market prices. This random vector belongs to a set Z in Euclidean space R^m and it evolves according to a function Ψ and an *iid* shock ε in a set of "events" E. The distribution of the shock ε is given by a probability measure Q defined on a measurable space $(E, \mathbb{E})$. Vector k lists the endogenous state variables which may correspond to several types of capital stocks and measures of wealth. The evolution of k is determined by an equilibrium decision rule g taking values in a set $K \subset R^l$. Hence, $s = (z, k)$ represents a generic vector in the state space $S = Z \times K$. We endow S with its relative Borel σ-algebra, which we denote by $\mathbb{S}$.

For expository purposes, let us write (2.1) in the simple form:

$$s_{n+1} = \varphi(s_n, \varepsilon_{n+1}), \qquad n = 0, 1, 2, \cdots . \tag{2.2}$$

We maintain the following basic assumptions:

ASSUMPTION 1 *The set S is compact.*

ASSUMPTION 2 *Function $\varphi : S \times E \to S$ is bounded and jointly measurable. Moreover, for every continuous function $f : S \to R$,*

$$(2.3) \qquad \int f(\varphi(s_j,\varepsilon))Q(d\varepsilon) \to_j \int f(\varphi(s,\varepsilon))Q(d\varepsilon) \;\; as \;\; s_j \to_j s.$$

Assumption 1 is standard in the numerical literature. Assumption 2 requires continuity in the state variable after integrating over the space of shocks. By the bounded convergence theorem Assumption 2 holds if $\varphi(\cdot,\varepsilon)$ is a continuous function of s for each given ε. This assumption is also satisfied in the presence of discrete jumps and further discontinuities of $\varphi(\cdot,\varepsilon)$ that are smoothed out when integrating over ε.

Stochastic systems can generate very complex dynamics. To analyze the average behavior of sample paths it is useful to define the transition probability function

$$(2.4) \qquad P(s,A) = Q(\{\varepsilon|\varphi(s,\varepsilon) \in A\}).$$

For any given initial condition μ_0 on $\mathbb{S}$, the evolution of future probabilities, $\{\mu_n\}$, can be specified by the following operator T^* that takes the space of probabilities on $\mathbb{S}$ into itself

$$(2.5) \qquad \mu_{n+1}(A) = (T^*\mu_n)(A) = \int P(s,A)\mu_n(ds),$$

for all A in $\mathbb{S}$ and $n \geq 0$. Condition (2.3) guarantees a weak continuity property of operator T^*; hence, this condition is often referred to as that the mapping φ is a *Feller map* or that P satisfies the *Feller property* [Stokey, Lucas and Prescott (1989), Chapter 8].

An invariant probability measure or invariant distribution μ^* is a fixed point of operator T^*, i.e., $\mu^* = T^*\mu^*$. Therefore, an invariant distribution is a stationary solution of the original system (2.1). The analysis of invariant distributions seems then a very first step to investigate the dynamics of the system.[2]

In many economic applications an explicit solution for the equilibrium function φ is not available. Then, the most one can hope for is to get a numerical approximation $\widehat{\varphi}$. Moreover, using some accuracy checks [cf. Santos (2000)] we may be able to bound the distance between functions φ and $\widehat{\varphi}$. Every numerical approximation $\widehat{\varphi}$ satisfying Assumptions 1-2 will give rise to a transition probability $\widehat{P}$ on $(S,\mathbb{S})$. But even if $\widehat{\varphi}$ is an arbitrarily good approximation of function φ, the asymptotic dynamics under transition functions P and $\widehat{P}$ may be quite different. Indeed, both functions may not possess the same number of invariant distributions, and the moments of these invariant distributions may be quite far apart. As the following

[2] In this paper we are concerned with invariant distributions μ^* on $\mathbb{S}$ of the Markov transition function P. As discussed below, function φ may contain some other invariant distributions ν which are not Markovian and which are jointly defined on $\mathbb{S}$ and the infinite product space of sequences of shocks $\{\varepsilon_n\}$.

simple example illustrates, without further restrictive assumptions we cannot expect good stability properties.

EXAMPLE 2.1: This example studies a change in the asymptotic dynamics from a numerical error on the transition probability P, but analogous examples can be constructed for numerical approximations of mapping φ. The state space S is a discrete set with three possible states, s_1, s_2, s_3. Transition function P is defined by the following Markov matrix

$$\Pi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix}.$$

Each row i specifies the probability of moving from state s_i to any state in S, so that an element π_{ij} corresponds to the value $P(s_i, \{s_j\})$, for $i, j = 1, 2, 3$. Note that $\Pi^n = \Pi$ for all $n \geq 1$. Hence, $p = (1, 0, 0)$, and $p = (0, 1/2, 1/2)$ are invariant distributions of Π, and $\{s_1\}$ and $\{s_2, s_3\}$ are the ergodic sets.[3] All other invariant distributions of Π are convex combinations of these two probabilities.

Assume now that P is calculated subject to some approximation error δ. Let

$$\widehat{\Pi} = \begin{bmatrix} 1-2\delta & \delta & \delta \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix} \text{ for } 0 < \delta < 1/2.$$

Then, as $n \to \infty$ the sequence of stochastic matrices $\{\widehat{\Pi}^n\}$ converges to

$$\begin{bmatrix} 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix}.$$

Hence, $p = (0, 1/2, 1/2)$ is the only possible long-run distribution for the system. Moreover, $\{s_1\}$ is a transient state, and $\{s_2, s_3\}$ is the only ergodic set. Therefore, a small error in P may lead to a pronounced change in its invariant distributions. Indeed, small errors may propagate over time and alter substantially the existing ergodic sets.

If we consider the correspondence that takes Π to the set of its invariant distributions $\{p | p\Pi = p\}$, then the present example shows that such correspondence fails to be lower semi-

[3] We say that a set $A \in \mathbb{S}$ is invariant if $P(s, A) = 1$ for all s in A. An invariant set A is called ergodic if there is no proper invariant subset $B \subset A$. An invariant distribution μ^* is called ergodic if $\mu^*(A) = 0$ or $\mu^*(A) = 1$ for every invariant set A. The support $\sigma(\mu^*)$ of a probability measure μ^* is the smallest closed set such that $\mu^*(\sigma(\mu^*)) = 1$.

continuous. A key step below is to establish that this correspondence is upper semicontinuous for a general class of Markov processes.

3 Asymptotic Convergence

In this section we establish an upper semicontinuity property of the correspondence of invariant distributions and a generalized law of large numbers. These results entail that the moments computed from numerical simulations are close to the set of moments of the model's invariant distributions as the approximation error of the computed solution converges to zero.

3.1 Upper Semicontinuity of the Correspondence of Invariant Distributions

We begin with some basic definitions in probability theory. Let $C(S)$ be the space of all continuous real-valued functions f on S. The integral $\int f(s)\mu(ds)$ or expected value of f over μ will be denoted by $E(f)$ whenever distribution μ is clear from the context. The weak topology is the coarsest topology such that every linear functional in the set $\{\mu \to \int f(s)\mu(ds), f \in C(S)\}$ is continuous. A sequence $\{\mu_j\}$ of probability measures on $\mathbb{S}$ is said to converge weakly to a probability measure μ if $\int f(s)\mu_j(ds) \to_j \int f(s)\mu(ds)$ for every $f \in C(S)$. The weak topology is metrizable [e.g., see Billingsley (1968)]. Hence, every weakly convergent sequence $\{\mu_j\}$ of probability measures has a unique limit point.

The existence of an invariant distribution follows from the above assumptions.

Theorem 1 *Under Assumptions 1-2, there exists a probability measure μ^* such that $\mu^* = T^*\mu^*$.*

For further reference, here is a sketch of the proof of this well-known result [e.g., Futia (1982), page 382]. Define the linear operator $(Tf)(s) = \int f(s')P(s, ds')$ all $s \in S$. Then by Assumption 2 operator T maps the space $C(S)$ into itself. Hence, the adjoint operator T^* must be weakly continuous. Moreover, by Assumption 1 the set of all probability measures on $\mathbb{S}$ is compact in the weak topology, and it is obviously a convex set. Therefore, by the Markov-Kakutani fixed-point theorem [Dunford and Schwartz (1958), Theorem V.10.5, page 456] there exists μ^* such that $\mu^* = T^*\mu^*$.

Let $\|\cdot\|$ be the max norm in R^l. Then, for any two vector-valued functions φ and $\widehat{\varphi}$ let

$$d(\varphi, \widehat{\varphi}) = \max_{s\in S}\left[\int \|\varphi(s,\varepsilon) - \widehat{\varphi}(s,\varepsilon)\| \, Q(d\varepsilon)\right]. \tag{3.1}$$

Convergence of a sequence of functions $\{\varphi_j\}$ should be understood in this notion of distance. Note that by Assumptions 1-2 each φ_j defines the associated pair (P_j, T_j^*); moreover, by Theorem 1 there always exists an invariant distribution $\mu_j^* = T_j^* \mu_j^*$.

Our first main result shows that under Assumptions 1-2 every sequence of invariant distributions of approximate solutions converges weakly to an invariant distribution of the model as the sequence of approximations converges to the true solution. Hence, the moments of an invariant distribution of a sufficiently good numerical approximation must be arbitrarily close to the moments of some invariant distribution of the model.

THEOREM 2 *Let $\{\varphi_j\}$ be a sequence of functions converging to φ. Let $\{\mu_j^*\}$ be a sequence of probabilities on $\mathbb{S}$ such that $\mu_j^* = T_j^* \mu_j^*$ for each j. Under Assumptions 1-2, if μ^* is a weak limit point of $\{\mu_j^*\}$ then $\mu^* = T^* \mu^*$.*

Theorem 2 asserts the bilinear convergence of $T_j^* \mu_j^*$ to $T^* \mu^*$ in the weak topology. This result is stronger than the standard notion of weak convergence, and entails that the invariant distributions correspondence is closed [cf. Hildenbrand (1974), page 23]. Therefore, this correspondence must be upper semicontinuous (see *opt. cit.*) since as asserted above the set of all probability measures on $\mathbb{S}$ is compact in the weak topology. An early result on the upper semicontinuity of the correspondence of invariant distributions appears in Dubins and Freedman (1966), Theorem 3.4. Further extensions can be found in Manuelli (1985) and Stokey, Lucas and Prescott (1989), Theorem 12.13. All these authors validate the upper semicontinuity of this correspondence under assumptions on the transition function P, but P is not a computable object and it is usually derived from knowledge of φ. The assumptions on P postulated by these authors do not seem to have a simple characterization in terms of equivalent conditions on φ. Their method of proof is directly applicable to our setting if for each ε the mapping $\varphi(s, \varepsilon)$ is continuous in s and the sequence of functions $\{\varphi_j\}$ converges to φ in the sup norm, but it cannot accommodate the weaker continuity condition (2.3). Our strategy of proof (see the Appendix) relies on a carefully chosen metric for the weak topology to gain control of the approximation errors. Note that for non-continuous functions φ and $\widehat{\varphi}$ the distance in (3.1) seems more appropriate than the usual functional metric induced by the sup norm, since (3.1) considers the average distance over the space of shocks rather than over each single value.

COROLLARY 1 *Let f belong to $C(S)$. Then under the conditions of Theorem 2, for every $\eta > 0$ there is $\delta > 0$ such that if μ_j^* is an invariant distribution of φ_j and $d(\varphi, \varphi_j) < \delta$ there exists an invariant distribution μ^* of φ such that*

$$|\int f(s)\mu_j^*(ds) - \int f(s)\mu^*(ds)| < \eta. \tag{3.2}$$

This is an alternative formulation of the upper semicontinuity of the correspondence of invariant distributions [cf. Hildenbrand (1974), page 23]. Hence, this result makes clear that function φ_j does not have to be an element of a converging sequence. Observe that function f refers to a characteristic of an invariant distribution such as the mean, the variance or any other quantitative property.

Theorem 2 can also be restated in terms of expectations operators for functions f in $C(S)$ over P-invariant distributions. This result is useful for later purposes. Let

$$E^{max}(f) = \max_{\{\mu^* | \mu^* = T^* \mu^*\}} \int f(s) \mu^*(ds) \tag{3.3a}$$

$$E^{min}(f) = \min_{\{\mu^* | \mu^* = T^* \mu^*\}} \int f(s) \mu^*(ds). \tag{3.3b}$$

Remark 1 As the set of invariant distributions $\{\mu^* | \mu^* = T^* \mu^*\}$ is weakly compact and convex, we get $[E^{min}(f), E^{max}(f)] = \{\int f(s)\mu^*(ds) | \mu^* = T^*\mu^*\}$. Therefore, for any continuous function f the interval $[E^{min}(f), E^{max}(f)]$ is conformed by the range of expected values $E(f)$ over the set of all the invariant distributions μ^*.

Corollary 2 *Let f belong to $C(S)$. Then under the conditions of Theorem 2, for every $\eta > 0$ there exists J such that*

$$E^{min}(f) - \eta < \int f(s) \mu_j^*(ds) < E^{max}(f) + \eta \tag{3.4}$$

for all μ_j^* with $j \geq J$.

It seems worth mentioning the following two applications of these results. First, for some computational methods [e.g., see Santos (1999)] one can obtain a sequence of numerical approximations $\{\varphi_j\}$ that converges to the exact solution φ. Then, Theorem 2 implies that the invariant distributions generated by these approximations will eventually be contained in an arbitrarily small weak neighborhood of the set of all invariant distributions generated by the original model. Second, various economic settings involve a family of solutions $s_{t+1} = \Phi(s_t, \varepsilon_{t+1}, \theta)$ parameterized by a vector θ in a space Θ; moreover, under mild regularity conditions it follows that $\Phi(s, \varepsilon, \theta_j) \to_j \Phi(s, \varepsilon, \theta)$ as $\theta_j \to \theta$. Then, as above for each θ one could define the functional mappings $E_\theta^{min}(f)$ and $E_\theta^{max}(f)$. Hence, (3.4) implies that $E_\theta^{min}(f)$ is a lower semicontinuous function of θ and $E_\theta^{max}(f)$ is an upper semicontinuous function of θ. If there exists but a unique invariant distribution μ_θ^* so that $E_\theta(f) = E_\theta^{min}(f) = E_\theta^{max}(f)$ for all θ, then the expectations operator $E_\theta(f)$ varies continuously with θ. This latter continuity property is usually assumed for the estimation of dynamic models [e.g., Duffie and Singleton (1993)], and it is of interest to derive this property from primitive conditions on φ.

3.2 A Generalized Law of Large Numbers

In applied work, laws of large numbers are often invoked to compute statistics of invariant distributions. For stochastic dynamical systems, usual derivations of laws of large numbers proceed along the lines of the ergodic theorem [e.g., see Krengel (1985)]. But to apply the ergodic theorem the initial state s_0 must lie inside an ergodic set. A certain technical condition – known as Hypothesis D from Doob (1953) – ensures that for every initial value s_0 the dynamical system will enter one of its ergodic sets almost surely.[4] Hence, using this condition standard laws of large numbers can be extended to incorporate all initial values s_0, since such points will eventually fall into one ergodic set. Hypothesis D, however, is usually difficult to verify in economic applications [Stokey, Lucas and Prescott (1989), Chapter 11].

Under conditions similar to Assumptions 1-2 above, Breiman (1960) proves a law of large numbers that is valid for all initial values s_0. This author dispenses with Hypothesis D, but requires the Markov process to have a unique invariant distribution. Uniqueness of the invariant distribution seems, however, a rather limiting restriction for numerical approximations. Primitive conditions may be imposed on the original model that guarantee the existence of a unique invariant distribution, but these conditions may not be preserved by the discretization procedure leading to the numerical approximation. Indeed, uniqueness of the invariant distribution is not robust to continuous perturbations of the model. This is illustrated in Figure 1 that portrays a deterministic policy function with a unique stationary point. The dotted line depicts a close approximation that contains a continuum of stationary points.[5]

Our goal is then to establish a law of large numbers that holds true for all initial values s_0 but without imposing the technical Hypothesis D assumed in Doob (1953) or the existence of a unique invariant distribution assumed in Breiman (1960). Our proof builds on some fundamental work by Crauel (1991, 2002) that bounds the average behavior of sample paths from the expected values over the set of all φ-invariant distributions. As discussed below, Crauel's bounds are not operative in applied work since function φ may contain some non-Markovian invariant distributions which are usually non-computable. We also dispense with a strong continuity condition which is replaced by condition (2.3). In preparation for our analysis, we define a new probability space comprising all infinite sequences $\omega = (\varepsilon_1, \varepsilon_2, \cdots)$.

[4] As explained in Doob (1953), Hypothesis D is a generalization of Doeblin's condition. Related conditions are ϕ-irreducibility and Harris' recurrence [e.g, see Jain and Jamison (1967) and Meyn and Tweedie (1993)]. All these conditions are difficult to verify in economic applications.

[5] Similar examples can be constructed for stochastic models. In these latter models uniqueness of the approximate solution may be preserved under certain regularity conditions (e.g., see example 5.4 below).

Let $\Omega = E^{\infty}$ be the countably infinite cartesian product of copies of E. Let $\mathbb{F}$ be the σ-field in E^{∞} generated by the collection of all cylinders $A_1 \times A_2 \times \cdots \times A_n \times E \times E \times E \times \cdots$ where $A_i \in \mathbb{E}$ for $i = 1, \cdots, n$. A probability measure λ can be constructed over these finite-dimensional sets as

$$\lambda\{\omega : \varepsilon_1 \in A_1, \varepsilon_2 \in A_2, \cdots, \varepsilon_n \in A_n\} = \prod_{i=1}^{n} Q(A_i). \tag{3.5}$$

This measure λ has a unique extension on $\mathbb{F}$. Hence, the triple $(\Omega, \mathbb{F}, \lambda)$ denotes a probability space. Finally, for every initial value s_0 and sequence of shocks $\omega = \{\varepsilon_n\}$, let $\{s_n(s_0, \omega)\}$ be the sample path generated by function φ, so that $s_{n+1}(s_0, \omega) = \varphi(s_n(s_0, \omega), \varepsilon_{n+1})$ for all $n \geq 1$ and $s_1(s_0, \omega) = \varphi(s_0, \varepsilon_1)$.

For expositional convenience, in the Appendix we provide a formal presentation of Crauel (2002), Proposition 6.21, page 95. We establish a more useful result for the simulation of dynamic economic models under two further weak assumptions. As above, we assume that $\{\varepsilon_n\}$ for $n \geq 0$ is an *iid* process whereas Crauel assumes that $\{\varepsilon_n\}$ for $-\infty < n < \infty$ is a stationary ergodic process. There is little loss of generality under *iid* shocks, since Markov processes offer enough flexibility to be suitable in most economic applications. Also, we consider moment functions $f(s)$ that only condition on the vector of state variables s rather than on both s and ω. Again, this is a reasonable restriction in applied work since the vector of noise variables ε is usually not observable. Under these two mild assumptions we extend Crauel's result in the following directions: (*i*) The mapping φ is only required to satisfy the weak continuity condition (2.3) rather than the more restrictive condition that $\varphi(\cdot, \varepsilon)$ must be continuous for each ε; and (*ii*) our bounds are sharper. Equalities (3.6a)-(3.6b) below hold for $E^{min}(f)$ and $E^{max}(f)$, and these bounds are the extreme points of the expected values $E(f)$ over the set of all P-invariant distributions [see (3.3a)-(3.3b)]. Crauel's bounds are the extreme points of the expected values $E(f)$ over the set of all φ-invariant distributions.[6] Therefore, we exclude from our analysis those other φ-invariant, non-Markovian distributions which do not admit a simple representation as a product measure.

THEOREM 3 *Let f belong to $C(S)$. Then, under Assumptions 1-2 for λ-almost all ω,*

[6] A measure ν on $(\mathbb{S}, \mathbb{F})$ is called φ-invariant if $\nu(\varphi^{-1}(D)) = \nu(D)$ for all sets $D \in (\mathbb{S}, \mathbb{F})$. Therefore, a φ-invariant distribution is defined on $(\mathbb{S}, \mathbb{F})$, whereas a P-invariant distribution is defined on the much simpler domain $\mathbb{S}$. We should stress that if μ^* is an invariant distribution of P, then $\mu^* \times \lambda$ is an invariant distribution of φ. But even if $\{\varepsilon_n\}$ is an *iid* process for $-\infty < n < \infty$ the mapping φ may contain some other invariant distributions ν that cannot be represented as a product measure $\mu^* \times \lambda$ [e.g., Arnold (1998), p. 56]. These other invariant distributions are usually non-computable, since they are jointly defined over S and the set of infinite-dimensional sequences $\{\varepsilon_n\}$.

$$(3.6a) \qquad (i) \quad \lim_{N\to\infty}(inf_{s_0\in S}[\frac{1}{N}\sum_{n=1}^{N} f(s_n(s_0,\omega))]) = E^{min}(f)$$

$$(3.6b) \qquad (ii) \quad \lim_{N\to\infty}(sup_{s_0\in S}[\frac{1}{N}\sum_{n=1}^{N} f(s_n(s_0,\omega))]) = E^{max}(f).$$

By the convexity of the set of invariant distributions μ^* (see Remark 1) this result implies that for every limit point $m(f(s_0,\omega))$ of $\{\frac{1}{N}\sum_{n=1}^{N} f(s_n(s_0,w))\}$ there exists an invariant distribution μ^* such that $m(f(s_0,\omega)) = \int f(s)\mu^*(ds)$ for *all* s_0 over a subset of Ω of full measure. Therefore, Theorem 3 entails that $\frac{1}{N}\sum_{n=1}^{N} f(s_n(s_0,\omega))$ approaches the interval of expected values $[E^{min}(f), E^{max}(f)]$ *uniformly* in s_0 for λ-almost all ω, and the bounds $E^{min}(f)$ and $E^{max}(f)$ are tight. Example 5.1 below illustrates that in the absence of the continuity condition (2.3) these equalities may no longer hold.

If there exists a unique invariant distribution μ^*, then $E^{min}(f) = E^{max}(f)$. Therefore, both limits in (3.6a)-(3.6b) are the same and correspond to the unique expected value $E(f)$. Thus, as a special case of Theorem 3 we obtain a standard formulation of the law of large numbers for a system with a unique invariant distribution.

Corollary 3 [cf. Breiman (1960)] *Assume that there exists a unique $\mu^* = T^*\mu^*$. Then under the conditions of Theorem 3 for all s_0 and for λ-almost all ω,*

$$(3.7) \qquad \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} f(s_n(s_0,\omega)) = E(f).$$

3.3 Accuracy of Numerical Simulations

We now apply the above results to the numerical simulation of stochastic dynamic models. A researcher is concerned with the predictions of a stochastic dynamic model whose equilibrium law of motion can be specified by a function φ. Usually, the Markovian equilibrium solution φ does not have an analytical representation, and so this function is approximated by numerical methods. Moreover, the invariant distributions or stationary solutions of the numerical approximation cannot be calculated analytically. Hence, numerical methods are again brought up into the analysis. This time in connection with some law of large numbers.

As is typical in the simulation of stochastic models we suppose that the researcher can draw sequences $\{\widehat{\varepsilon}_n\}$ from a generating process that can mimic the distribution of the shock process $\{\varepsilon_n\}$. A probability measure λ is defined over all sequences $\omega = (\varepsilon_1, \varepsilon_2, ...)$. Once a numerical approximation φ_j is available, it is generally not so costly to generate sample paths $\{s_{jn}(s_0, \omega)\}$ defined recursively as $s_{jn+1}(s_0, \omega) = \varphi_j(s_{jn}(s_0, \omega), \varepsilon_{n+1})$ for every $n \geq 0$ for fixed s_0 and ω. Averaging over these sample paths we get sequences of simulated statistics $\{\frac{1}{N}\sum_{n=1}^{N} f(s_{jn}(s_0, \omega))\}$ as defined by some function f. We next show that for a sufficiently good numerical approximation φ_j and for a sufficiently large N the series $\{\frac{1}{N}\sum_{n=1}^{N} f(s_{jn}(s_0, \omega))\}$ is close (almost surely) to the expected value $E(f) = \int f(s)\mu^*(ds)$ of some invariant distribution μ^* of the original equilibrium function φ.

THEOREM 4 *Under the conditions of Theorem 2, for every $\eta > 0$ there are functions $N_j(w)$ and an integer J such that for all $j \geq J$ and $N \geq N_j(\omega)$,*

$$E^{min}(f) - \eta < \frac{1}{N}\sum_{n=1}^{N} f(s_{jn}(s_0, \omega)) < E^{max}(f) + \eta \tag{3.8}$$

for all s_0 and λ-almost all ω.

Theorem 4 is a simple consequence of Corollary 2 and Theorem 3. Indeed, by (3.3a)-(3.3b) and (3.6a)-(3.6b) the term $\int f(s)\mu_j^*(ds)$ in (3.4) can be replaced (up to an arbitrarily small error) by $\frac{1}{N}\sum_{n=1}^{N} f(s_{jn}(s_0, \omega))$. Moreover, by a further application of (3.6a)-(3.6b) we can rewrite (3.8) as

(3.9)

$$inf_{s_0 \in S}[\frac{1}{N}\sum_{n=1}^{N} f(s_n(s_0, \omega))] - \eta < \frac{1}{N}\sum_{n=1}^{N} f(s_{jn}(s_0, \omega)) < sup_{s_0 \in S}[\frac{1}{N}\sum_{n=1}^{N} f(s_n(s_0, \omega))] + \eta$$

for all j and N large enough. Hence, statistical, long-run average properties of sample paths of approximate solutions are always (up to a small error) within the feasible interval of corresponding values of sample paths of equilibrium solutions.

As already discussed, the convergence in (3.6a)-(3.6b) is uniform in s_0 over λ-almost all ω. Furthermore, by the convexity of the set of invariant distributions (see Remark 1) the whole interval of values $[E^{min}(f), E^{max}(f)]$ is equal to the range of the linear functional $\mu^* \to \int f(s)\mu^*(ds)$ over the set $\{\mu^* | \mu^* = T^*\mu^*\}$. Therefore, Theorem 4 could be loosely restated as follows: If φ_j is sufficiently close to φ then for all s_0 and for every ω in a set of full measure the series $\{\frac{1}{N}\sum_{n=1}^{N} f(s_{jn}(s_0, \omega))\}$ must be arbitrarily close to the expected value $E(f) = \int f(s)\mu^*(ds)$ of some invariant distribution μ^* for all large enough N. Of course, if there exists a unique invariant distribution μ^* then $\{\frac{1}{N}\sum_{n=1}^{N} f(s_{jn}(s_0, \omega))\}$ must lie in a small neighborhood of $E(f) = \int f(s)\mu^*(ds)$ for sufficiently large N.

Corollary 4 *Assume that there exists a unique invariant distribution $\mu^* = T^*\mu^*$. Then under the conditions of Theorem 4 for all $j \geq J$ and $N \geq N_j(\omega)$,*

$$(3.10) \qquad |\frac{1}{N}\sum_{n=1}^{N} f(s_{jn}(s_0,\omega)) - E(f)| < \eta$$

for all s_0 and λ-almost all ω.

Observe that each approximating function φ_j may contain multiple invariant distributions μ_j^*. In fact, this corollary is a good application of Theorem 3, since we do not have to assume Hypothesis D or the existence of a unique invariant distribution for every approximating function φ_j. Also, (3.9) now reads

$$(3.11) \qquad |\frac{1}{N}\sum_{n=1}^{N} f(s_{jn}(s_0,\omega)) - \frac{1}{N}\sum_{n=1}^{N} f(s_n(s_0,\omega))| < \eta$$

for all N large enough. Hence, if the model contains a unique invariant distribution, then generically up to a small approximation error all sample paths of both the model and the numerical approximation share the same statistical, long-run average behavior.

Theorems 2 and 3 and all their corollaries have been presented in a framework of numerical simulation, but they may be recasted under alternative perturbations of the model such as variation of parameter values and of the distribution of the stochastic shock ε. These results are essential to establish asymptotic properties of simulation-based estimators [Santos (2004)].

4 Error Bounds

In numerical applications it is often desirable to bound the size of the approximation error. Computations must stop in finite time, and hence error bounds can dictate efficient stopping criteria. In this section we provide various upper error estimates for the above convergence results under the following simple contractivity condition on the dynamics.

CONDITION C: There exists a constant $0 < \gamma < 1$ such that $\int \|\varphi(s,\varepsilon) - \varphi(s',\varepsilon)\| \, Q(d\varepsilon) \leq \gamma \|s - s'\|$ for all pairs s, s'.

Condition C is familiar in the literature on Markov chains [e.g., see Norman (1972) for an early analysis and applications, and Stenflo (2001) for a recent update of the literature]. Related contractivity conditions are studied in Dubins and Freedman (1966), Schmalfuss (1996) and Bhattacharya and Majumdar (2004). In the macroeconomics literature, Condition C arises naturally in one-sector growth models [Schenk-Hoppe and Schmalfuss (2001)] and in models of asset trading [Foley and Hellwig (1975)]. Stochastic contractivity properties are also encountered in learning models [Schmalensee (1975), and Ellison and Fudenberg (1993)] and in certain types of stochastic games [Sanghvi and Sobel (1976)].

As shown in Stenflo (2001), Condition C implies the Feller property. Also, this condition guarantees the existence of a unique invariant distribution μ^* for system (2.5). The distribution μ^* is globally stable under the action of T^* and the rate of convergence to such stationary solution is linear. We now present a formal version of this convergence result which is needed in the sequel. A real-valued function f on S is called Lipschitz with constant L if $|f(s) - f(s')| \leq L \|s - s'\|$ for all pairs s and s'. Let $d = diam(S) = max_{s,s' \in S} \|s - s'\|$.

Theorem 5 [cf. Stenflo (2001)] *Let f be a Lipschitz function with constant L. Assume that φ satisfies Condition C. Then under Assumptions 1-2 there exists a unique $\mu^* = T^*\mu^*$ such that for every sequence $\{\mu_n\}$, with $\mu_{n+1} = T^*\mu_n$ for all n,*

$$|\int f(s)\mu_n(ds) - \int f(s)\mu^*(ds)| \leq \frac{Ld\gamma^n}{1-\gamma}. \tag{4.1}$$

Now, let $\widehat{\varphi}$ be an approximate solution. Then, under Assumptions 1-2 this function defines an operator $\widehat{T}^*$ that has a fixed-point solution $\widehat{\mu}^* = \widehat{T}^*\widehat{\mu}^*$. The following theorem bounds the distance between the expected values of f over distributions μ^* and $\widehat{\mu}^*$. This result applies to any numerical approximation, and hence it does not require a sequence of approximate policy functions converging to the true solution.

Theorem 6 *Let f be a Lipschitz function with constant L. Let $d(\varphi, \widehat{\varphi}) \leq \delta$ for some $\delta > 0$. Assume that φ satisfies Condition C. Then under Assumptions 1-2,*

$$|\int f(s)\mu^*(ds) - \int f(s)\widehat{\mu}^*(ds)| \leq \frac{L\delta}{1-\gamma} \tag{4.2}$$

where μ^* is the unique invariant distribution of φ, and $\widehat{\mu}^*$ is any invariant distribution of $\widehat{\varphi}$.

Note that function $\widehat{\varphi}$ may not satisfy Condition C, and hence it may contain multiple Markovian invariant distributions $\widehat{\mu}^*$. Stenflo (2001) proves a related result in which the distance $d(\varphi, \widehat{\varphi})$ is in the sup norm and the approximate function $\widehat{\varphi}$ is required to satisfy Condition C. But this condition is not necessary and it may be hard to verify in applications.

Indeed, Condition C may not be preserved under polynomial interpolations and related high-order approximation schemes under which function $\widehat{\varphi}$ may have been calculated.

Using Theorem 6, we can now establish a sharper version of Theorem 4 that provides error bounds for the statistics from simulations of approximate solutions. Again, this is a new result that holds even if the approximated solution has multiple invariant distributions. For given s_0 and $\omega = \{\varepsilon_n\}$, let $\widehat{s}_{n+1}(s_0, \omega) = \widehat{\varphi}(\widehat{s}_n(s_0, \omega), \varepsilon_{n+1})$ for all $n \geq 0$.

Corollary 5 *Under the conditions of Theorem 6, for every* $\eta > 0$ *there exists a function* $\widehat{N}(\omega)$ *such that for all* $N \geq \widehat{N}(\omega)$,

$$|\frac{1}{N}\sum_{n=1}^{N} f(\widehat{s}_n(s_0, \omega)) - E(f)| \leq \frac{L\delta}{1-\gamma} + \eta \tag{4.3}$$

for all s_0 *and* λ*-almost all* ω.

All these bounds have been constructed on a worst-case scenario for the approximation error, and so they may not be operative in particular applications. Some approximations work well along a "typical" sample path, but perform quite poorly for points outside the support of an invariant distribution μ^*. This latter property may actually be a virtue of the algorithm since points outside the ergodic sets may have a negligible weight on the simulated statistics. Linear-quadratic approximations or more refined algorithms such as the perturbation method of Judd (1998), Chapter 13, or the PEA method of den Haan and Marcet (1994) are often quite effective, even though these methods may not deliver good approximations for extreme points of the state space. Hence, it is usually very useful to focus on the distribution of simulated sample paths $\{\widehat{s}_n\}$ of the approximate solution $\widehat{\varphi}$. Reiter (2000) has suggested some procedures to bound the distance between an (unknown) policy rule φ and a numerical approximation $\widehat{\varphi}$ over the simulated paths $\{\widehat{s}_n\}$. For situations in which those bounds are available, we can offer tighter error estimates in terms of an expected approximation error $\widehat{\delta}_N$ defined over sequences of sample paths $\{\widehat{s}_n\}_{n=1}^{N}$ of finite length. As before, for given s_0 and $\omega = \{\varepsilon_n\}$, let $s_{n+1}(s_0, \omega) = \varphi(s_n(s_0, \omega), \varepsilon_{n+1})$ and $\widehat{s}_{n+1}(s_0, \omega) = \widehat{\varphi}(\widehat{s}_n(s_0, \omega), \varepsilon_{n+1})$, for all $n \geq 0$.

Theorem 7 *Under the conditions of Theorems 5-6, for every* N,

$$|E(f) - E(f(\widehat{s}_N(s_0, \omega)))| \leq \frac{Ld\gamma^N}{1-\gamma} + \frac{L\widehat{\delta}_N}{1-\gamma} \tag{4.4}$$

where

$$\widehat{\delta}_N = max_{n=1,2,...N} E(\|\varphi(\widehat{s}_{n-1}(s_0,\omega),\omega) - \widehat{\varphi}(\widehat{s}_{n-1}(s_0,\omega),\omega)\|). \tag{4.5}$$

Let us explain this result which combines arguments from Theorems 5 and 6. Our objective is to estimate the expected value $E(f) = \int f(s)\mu^*(ds)$ from sequences of simulated sample paths $\{\widehat{s}_n\}$. Consequently, it should be understood that the expected error bound $\widehat{\delta}_N$ in (4.5) is calculated over sequences of shocks $\{\varepsilon_1, ..., \varepsilon_N\}$ for points $\widehat{s}_n$ generated by $\widehat{\varphi}$. We pick an initial state s_0 and a time horizon N. A good guess for s_0 is useful to sharpen the bound $\widehat{\delta}_N$ in (4.5). We decompose the upper error bound in (4.4) as follows. Chopping the horizon to time N leads to an approximation error $|E(f) - E(f(s_N(s_0,\omega)))|$, where $E(f(s_N(s_0,\omega)))$ is the expected value of function f over the distribution of random vector $s_N(s_0,\omega)$. Then, Theorem 5 implies that $|E(f(s)) - E(f(s_N(s_0,\omega)))| \leq \frac{Ld\gamma^N}{1-\gamma}$, which is the first component of the error in (4.4). The second component in (4.4) comes from the approximation of $E(f(s_N(s_0,\omega)))$ by $E(f(\widehat{s}_N(s_0,\omega)))$. The difference between these two values can be bounded along the lines of Theorem 6 (see the Appendix). Moreover, by the strong law of large numbers applied to sequences of length N for realizations of variable ε we can compute $E(f(\widehat{s}_N(s_0,\omega)))$ from a sufficiently large number M of sample paths $\{\{\widehat{s}_n^m\}_{n=1}^N\}_{m=1}^M$ of length N. Hence, for given s_0 and N,

$$E(f(\widehat{s}_N(s_0,\omega))) = \lim_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} f(\widehat{s}_N^m) \tag{4.6}$$

almost surely. Finally, the bound $\widehat{\delta}_N$ can be estimated by analytical arguments [cf. Reiter (2000) and Santos (2000)] together with the law of large numbers. Therefore, Theorem 7 provides an estimate of $E(f)$ from $E(f(\widehat{s}_N(s_0,\omega)))$ under the bound $\widehat{\delta}_N$ in (4.5), and both $E(f(\widehat{s}_N(s_0,\omega)))$ and $\widehat{\delta}_N$ can be computed from a sufficiently large number M of simulated sample paths $\{\{\widehat{s}_n^m\}_{n=1}^N\}_{m=1}^M$ of length N.

5 Economic Applications

In this section we present several illustrative examples on the role of our main assumptions and the applicability of our results to economic models. Assumption 1 could be weakened, since some basic results in probability theory have suitable extensions to non-compact domains [e.g., Futia (1982)]. This extension will burden the presentation significantly, and it will not be pursued here. Assumption 2 and Condition C seem much harder to replace.

Assumption 2 is required to guarantee the existence of an invariant distribution, and some further regularity conditions are needed to establish error bounds. There are however some important classes of economic models that are not covered by the present results, and an open issue addressed later is whether these models can be simulated by reliable numerical methods.

5.1 The Feller property

As already discussed, the weak continuity condition (2.3) – known as the Feller property – allows for jumps in mapping $\varphi(s, \varepsilon)$ that can be flattened out when integrating over ε. Discontinuities in the policy rule can arise in economic decisions because of fixed costs, indivisibilities, or a switch to a new policy that may be optimal after reaching a certain threshold. The importance of condition (2.3) can be appreciated in the following two examples. Our first example shows the necessity of this condition for the existence of an invariant distribution (Theorem 1) and for the law of large numbers (Theorem 3). The second example introduces a model of technology adoption with fixed costs.

EXAMPLE 5.1: Consider a dynamical system that is conformed by the random application of two real-valued functions g and $\widehat{g}$. These two mappings are depicted in Figure 2. We assume that g and $\widehat{g}$ are discontinuous from the left at point $s = 1$. Moreover, $g(s) > \widehat{g}(s) > s$ for all $s < 1$, and $\lim_{s \to 1^-} g(s) = \lim_{s \to 1^-} \widehat{g}(s) = 1$; also, $s > g(s) > \widehat{g}(s)$ for all $s \geq 1$. The random iteration of these two mappings follows an *iid* process in which at each date $n = 0, 1, 2, \cdots$, the system moves by function g with probability $1 > q > 0$, and by function $\widehat{g}$ with probability $1 - q$.

The resulting stochastic Markov chain fails to satisfy condition (2.3) at point $s = 1$, since both functions g and $\widehat{g}$ have a discontinuity at $s = 1$. Moreover, there does not exist an invariant distribution[7] and the above law of large numbers does not hold. Indeed, for every sample path $\{s_n\}$ generated by this random dynamical system the partial sums $\{\frac{1}{N}\sum_{n=1}^{N} s_n\}$ converge to 1. This limit point cannot be the mean of an invariant distribution, since no invariant distribution does exist.

EXAMPLE 5.2: We present a simple version of a model of technology adoption analyzed in Alpanda and Peralta-Alva (2004). This model is motivated by the energy price shocks of 1973-74 that occurred together with a 50 percent drop in the value of US corporations and a subsequent progressive decline in the ratio of consumption of US corporate energy to

[7]This example could be slightly modified so that the set of invariant distributions of the random dynamical system is not compact and hence the correspondence of invariant distributions is not upper semicontinuous (*viz.* Theorem 2).

output. The production process allows for a switch to an energy-saving technology, but past investments are irreversible. Hence, at the time of the switch the value of capital installed in the old production process can fall sharply.

A social planner has preferences represented by the discounted objective

$$\max E \sum_{t=0}^{\infty} \beta^t \log c_t$$

where E is the expectations operator conditioning on information at time $t = 0$ over the stream of consumption values $\{c_t\}$, for $t = 0, 1, 2, \cdots$, and $0 < \beta < 1$ is the discount factor.

There are two available technologies, $i = 1, 2$, to produce the single good. Both technologies use capital, k, energy, e, and labor, n. Technology 1 is already operative at time $t = 0$. Technology 2 is energy saving but a one-time fixed cost $I > 0$ must be incurred at the starting date of its operation. This initial cost for technology adoption could be modelled in more refined ways [Boldrin and Levine (2001)] but our results seem to be unaffected. The total quantity of output $y = y_1 + y_2$ is the sum of production under the two technologies

$$y_1 = F(k_1, e_1, n) \tag{5.1a}$$

$$y_2 = F(k_2, \xi e_2, 1 - n) \tag{5.1b}$$

where F is linearly homogeneous, monotone, concave and continuous. Technology 2 is in all aspects equal to technology 1 except that it requires less energy per unit of output. Hence, we set $\xi > 1$. The law of motion for each type of capital is given by

$$k_{it+1} = (1 - \pi)k_{it} + x_{it+1} \tag{5.2}$$

where $t = 0, 1, 2, \cdots$, $x_i \geq 0$ is investment in capital of type $i = 1, 2$, and $0 \leq \pi \leq 1$ is the depreciation rate. Hence, investment is irreversible: A given type of capital cannot be consumed or transformed into another type of capital. Energy is imported from abroad. The price of energy p_t is determined by a first-order autoregressive process

$$p_{t+1} = \rho p_t + \varepsilon_{t+1} \tag{5.3}$$

where $0 \leq \rho < 1$, and $\{\varepsilon_t\}$ is a sequence of *iid* random variables with continuous density and support in some compact interval $[\varepsilon_l, \varepsilon_h]$.

Production takes place at the beginning of each period t, and the realization of the shock ε_{t+1} is known at the end of period t. After this information is revealed the planner can choose the optimal quantities of consumption c_t and investment x_{it+1} for $i = 1, 2$. Aggregate feasibility implies

$$y_t = c_t + p_t(e_{1t} + e_{2t}) + x_{1t+1} + x_{2t+1} + I_t \tag{5.4}$$

where $I_t = I$ in the first date t such that $x_{2t+1} > 0$, and $I_t = 0$ at every other date. At time 0 the planner owns k_0 units of capital of type 1, and 0 units of capital of type 2. Then, for given p_0 the decision problem is to choose a contingent plan $\{c_t, n_t, e_{1t}, e_{2t}, x_{1t+1}, x_{2t+1}\}$ that solves $(P0)$ subject to constraints (5.1) – (5.4).

Alpanda and Peralta-Alva (2004) consider various functional forms for function F. They compute the optimal solution numerically and in their parameterizations technology 2 is eventually active with probability 1 for all positive initial conditions (k_0, p_0). We follow a similar calibration procedure, and under the same solution methods for all our parameterizations we find that the optimal policy $\varphi(k_0, p_0, \varepsilon)$ satisfies the following property: For each positive vector (k_0, p_0) there exists at most one $\tilde{\varepsilon}$ such that for $\varepsilon \neq \tilde{\varepsilon}$ the optimal choice $\varphi(k_0, p_0, \varepsilon)$ is unique, and $x_{11} = 0$ for $\varepsilon > \tilde{\varepsilon}$ and $x_{12} = 0$ for $\varepsilon < \tilde{\varepsilon}$. That is, for given (k_0, p_0) the planner will only invest in technology 2 if the energy shock ε is sufficiently high, and will at most invest in technology 1 if the energy shock ε is sufficiently low; for any given (k_0, p_0) there is at most one value $\tilde{\varepsilon}$ for the shock in which the planner is indifferent between investing in either of the two technologies.

It is now easy to check that such policy $\varphi(k_0, p_0, \varepsilon)$ satisfies condition (2.3). This is a simple consequence of the following two properties: (*i*) By the upper semicontinuity of the correspondence of maximizers for any given initial condition (k_0, p_0) the optimal policy $\varphi(k_0, p_0, \varepsilon)$ must be a continuous function at every (k_0, p_0, ε) for all $\varepsilon \neq \tilde{\varepsilon}$ since the optimal solution is unique; moreover, at the point of discontinuity $(k_0, p_0, \tilde{\varepsilon})$ the value $\tilde{\varepsilon}$ has measure zero; and (*ii*) it is not profitable to invest in technology 1 after technology 2 begins operating; hence, after the switch the planner is actually solving a concave optimization problem in which the optimal policy is a continuous function.

There are many other dynamic models in economics with discontinuous policy rules that may nevertheless satisfy condition (2.3). These models may involve some discrete choices, and the decision rule may not be monotone. Typical examples include (S, s) policies in inventory theory [Hall and Rust (2001)], patents [Pakes (1986)], durable goods replacement [Rust (1987)], fertility decisions [Wolpin (1984)], and occupational choice [Ljungqvist and Sargent (2004), Chapter 6]. Our asymptotic results demonstrate that these models can be simulated by standard numerical methods.

5.2 Condition C

We now analyze two additional examples that satisfy Condition C. Both examples focus on a simple version of the one-sector growth model, but the methods developed here are of general interest. In the first example the policy function has an analytical representation, and our main objective is to evaluate the performance of our theoretical error bounds for two numerical approximations. In contrast, in the second example the solution is not analytic, and we provide some methods for the estimation of the modulus of contraction $0 < \gamma < 1$ of Condition C.

EXAMPLE 5.3: This is a simple version of the stochastic one-sector growth model of Brock and Mirman (1972):

$$\max E \sum_{t=0}^{\infty} \beta^t \log c_t$$

$$\text{subject to } k_{t+1} = A\varepsilon_t k_t^\alpha + (1-\pi)k_t - c_t$$

$$k_0 \text{ fixed, } k_t \geq 0,\ t = 0, 1, \cdots$$

$$0 < \beta < 1,\ A > 0,\ 0 < \alpha < 1,\ 0 < \pi \leq 1,$$

where E is the expectations operator, and $\{\varepsilon_t\}$ is an *iid* process drawn from a truncated log-normal distribution over the domain $[-0.032, 0.032]$ with $E(\log \varepsilon) = 0$ and $Var(\log \varepsilon) = \sigma_\varepsilon^2$.

The decision problem is to allocate at each date $t = 0, 1, \cdots$, the amounts of consumption c_t and capital for the next period k_{t+1} so as to satisfy feasibility constraint (5.5). As is well known, for $\pi = 1$ the optimal policy function is given by $k_{t+1} = \alpha\beta A\varepsilon_t k_t^\alpha$, for all t. As $\{\varepsilon_t\}$ is an *iid* process, total output produced $y = A\varepsilon k^\alpha$ is actually a state variable. Then, taking logs the equilibrium dynamics can be characterized by the following equation

$$\log y_{t+1} = \log A + \alpha \log(\alpha\beta) + \alpha \log y_t + \log \varepsilon_{t+1}. \tag{5.6}$$

Clearly, this transformed system satisfies Condition C with modulus of contraction $\gamma = \alpha$.

Moreover, it follows from this equation that asymptotically

$$E(\log y) = \frac{\log A + \alpha \log(\alpha\beta)}{1-\alpha} \tag{5.7a}$$

$$Var(\log y) = \frac{\sigma_\varepsilon^2}{1-\alpha^2}. \tag{5.7b}$$

As is typical in quantitative economic analysis, let us suppose that the researcher does not have access to the true solution of the model, but only to a numerically generated policy. For concreteness we check the performance of our error bounds for the following two approximations:

(a) A multiplicative perturbation φ_δ of the exact policy, $\varphi_\delta(y_t, \varepsilon_{t+1}) = \delta\varphi(y_t, \varepsilon_{t+1})$ with $\delta \neq 1$; hence, $y_{t+1} = \varphi_\delta(y_t, \varepsilon_{t+1}) = \delta A \varepsilon_{t+1}(\alpha\beta y_t)^\alpha$.

(b) A linear approximation φ_L of the exact policy φ around the deterministic steady state $y^* = \varphi(y^*, 1)$; hence, $y_{t+1} = \varphi_L(y_t, \varepsilon_{t+1}) = y^* + \alpha^2\beta A(\alpha\beta y^*)^{\alpha-1}(y_t - y^*) + y^*(\varepsilon_{t+1} - 1)$ for $y^* = \sqrt[1-\alpha]{A(\alpha\beta)^\alpha}$.

For the numerical experiments below we consider the following baseline parameterization,

$$\beta = 0.95, A = 1.4685, \alpha = 0.34, \pi = 1, \sigma_\varepsilon = 0.008. \tag{5.8}$$

Whenever some of these parameter values are changed, normalizing constant A will be adjusted so that the deterministic steady state $y^* = \varphi(y^*, 1)$ always remains at the benchmark value $y^* = 1.00$. Hence, some statistics reported below may also be read as percentage deviations from the steady state value.

Error Bounds

Let y_δ refer to the random variable generated by function φ_δ, and y_L refer to the random variable generated by function φ_L. One readily sees that asymptotically

$$E(\log y_\delta) = \frac{\log \delta + \log A + \alpha \log(\alpha\beta)}{1-\alpha} \tag{5.9a}$$

$$Var(\log y) = \frac{\sigma_\varepsilon^2}{1-\alpha^2}. \tag{5.9b}$$

From (5.7a) and (5.9a) we get that $|E(logy) - E(logy_\delta)| = \frac{\log \delta}{1-\alpha}$. Since for $f(\log y) = \log y$ the Lipschitz constant L of Theorem 6 equals 1, it follows that for the mean the bound obtained in Theorem 6 is the best possible; for the above parameterization this bound $\frac{\log \delta}{1-\alpha}$ equals $1.5 \log \delta$. Also, from (5.7b) and (5.9b) we get that $|Var(\log y) - Var(\log y_\delta)| = 0$. Now, for the computation of constant L there is an additional approximation error, since the derivative of $f(\log y) = (\log y - E(\log y))^2$ is not constant in $\log y$. In this case we make L equal to the maximum value of the derivative of function $(\log y - E(\log y))^2$ with respect to $\log y$ over the support of the invariant distribution since this domain can be computed analytically. The bound in Theorem 6 is now equal to $0.0745 \log \delta$. Hence, this error bound is a small percentage of the approximation error of the numerical policy.

For the linear approximation φ_L the invariant distribution of y_L does not have an analytical representation. We then appeal to Theorem 3 and compute the first and second order moments of this distribution from a generic path of arbitrarily large length.[8] To evaluate the accuracy of the moments generated by the linear approximation φ_L, we consider variations in parameter α over the values 0.17, 0.34, and 0.68. Our numerical exercises are reported in Tables I and II. The second column of Table I reports the difference between the means of the invariant distributions of variables $\log y$ and $\log y_L$. (The expression $xE - y$ means x times 10^{-y}.) The third column reports the error bound for the mean as given by Theorem 6 (EBMTH6) and the last column reports the error bound for the mean as given by Theorem 7 (EBMTH7). Regarding the computation of (EBMTH6), the distance $d(\log \varphi, \log \varphi_L)$ has been calculated over the support of the invariant distribution.[9] As we can see in Table I, our theoretical error bound (EBMTH6) seems reasonably good, as it is around 10^2 times bigger than the observed error. This bound gets worse as we increase α. Indeed, as the modulus of contraction goes up the ergodic set is expanded, and hence the distance between the the true policy function and the quadratic approximation increases over the ergodic set. Thus, for $\alpha = 0.68$ the bound (EBMTH7) becomes much tighter. Table II reports corresponding computations for the variance with very similar results. The better performance of (EBVTH7) is mainly observed for high values of α.

In summary, this example illustrates that for some approximations the upper bound of Theorem 6 is the best possible, but this bound cannot be expected to be tight in all circumstances. For sensible calibrations of the model this bound was quite effective, i.e

[8] In all our numerical exercises for the computation of all our estimates we use a sequence of pseudo-random numbers for ε of length $N = 150,000$; moreover, to minimize the transitory effects of an arbitrary choice of the initial condition y_0 we delete the first 1000 observations. The length of this path seems larger than necessary in all cases, since the absolute difference of the simulated statistics with those computed using much larger paths was always smaller than 10^{-12}.

[9] Although our state variable is y, in this example we compute all the statistics for the transformed variable $\log y$ since from (5.6) the transformed variable satisfies Condition C with modulus of contraction α.

around 10^2 times bigger than the observed error. In general, this bound will perform well if the distance between the numerical approximation and the true solution is uniform over all values of state variable y and if the Lipschitz constant of the moment function f is a tight upper bound for its derivative over these values. Sharpening these error bounds along the lines of Theorem 7 should prove useful in applications.

EXAMPLE 5.4: We now extend the analysis of the previous example for depreciation factors $0 < \pi < 1$. In this case the model does not possess an analytical solution, and hence we propose a simple operational way to estimate numerically the modulus of the random contraction of Condition C. This numerical procedure computes the derivative of the policy function and the invariant distribution of the model. Theorems 2 and 3 play a fundamental role in this analysis.

As shown in Brock and Mirman (1972) and Hopenhayn and Prescott (1994) the one-sector growth model with *iid* shocks has a unique invariant distribution μ^*; moreover it is easy to show that in our case the support of this distribution $\sigma(\mu^*)$ is a non-degenerate interval of values $[y_l, y_h]$. We are not aware of any theoretical work for this model to bound the modulus of contraction of Condition C. Our approach will be numerical. That is, for a given parameterization of the model we assume that the researcher knows an approximate policy function $\widehat{\varphi}$, and an upper bound δ for the distance $d(\varphi, \widehat{\varphi})$ between the approximate function $\widehat{\varphi}$ and the true policy function φ, which may be unknown.

Let μ^* be the unique Markovian invariant distribution of φ, and $\mu^*_{\widehat{\varphi}}$ an invariant distribution of $\widehat{\varphi}$. By Theorem 2 above, every sequence $\{\mu^*_{\widehat{\varphi}_n}\}$ converges to μ^* as $\widehat{\varphi}_n \to_n \varphi$. Furthermore, as the support $\sigma(\mu^*)$ is a non-degenerate interval of values $[y_l, y_h]$, weak convergence in the space of probability measures [Billingsley (1968), Theorem 2.1] implies that $\mu^*_{\widehat{\varphi}}$ must be unique for $\widehat{\varphi}$ sufficiently close to φ. For if there were at least two ergodic invariant distributions for a sufficiently close approximation $\widehat{\varphi}$, then their supports would overlap; but under mild regularity conditions satisfied in the present model, these ergodic sets must be disjoint [cf. Futia (1982), pages 387 and 397].

Next, we compute the support $\sigma(\mu^*_{\widehat{\varphi}})$ from a generic, arbitrarily large sample path generated by $\widehat{\varphi}$. For given y_0 and $\omega = \{\varepsilon_n\}$, let $\widehat{y}_{t+1}(y_0, \omega) = \widehat{\varphi}(\widehat{y}_t(y_0, \omega), \varepsilon_{t+1})$ for all $t \geq 0$. For any sample of points $\{\widehat{y}_t(y_0, \omega)\}_{t=1}^N$ of length N we define the empirical distribution $\widehat{\mu}^N_{(y_0,\omega)}$ with support on $\{\widehat{y}_t(y_0, \omega)\}_{t=1}^N$ as $\widehat{\mu}^N_{(y_0,\omega)}(\widehat{y}_t(y_0, \omega)) = \frac{1}{N}$ for $t = 1, 2, \cdots, N$. Then, by a standard argument [cf. Walters (1982), Chapter 6, Theorem 6.4] it follows from our law of large numbers (Theorem 3) that for all y_0 and λ-almost all ω the sequence $\{\widehat{\mu}^N_{(y_0,\omega)}\}$ converges weakly to the invariant distribution $\mu^*_{\widehat{\varphi}}$, as $N \to \infty$. Therefore, Theorems 2 and 3 entail that for a good approximation $\widehat{\varphi}$ and for large N the support $\sigma(\widehat{\mu}^N_{(y_0,\omega)})$ yields a good approximation for the ergodic set $\sigma(\mu^*)$. As shown below, φ is a stochastic contraction and so we can appeal to Theorem 6 to provide error estimates.

Finally, to estimate constant γ in Condition C we compute the derivative of the *true* policy function over our approximated ergodic set $\sigma(\widehat{\mu}^N_{(y_0,\omega)})$. For the computation of this derivative we have considered a quadratic optimization problem derived in Santos (1999), page 678. This optimization problem yields an *exact* estimate for the derivative of the policy function. Such an optimization problem is not computable and must be suitably discretized. Moreover, the problem requires knowledge of the true policy function φ so that the quadratic objective is evaluated over all optimal sample paths of the state variable. We generate these sample paths by the approximate policy function $\widehat{\varphi}$. In our estimation below of constant γ we report the maximum value that the derivative of the policy function attains over the computed ergodic set $\sigma(\widehat{\mu}^N_{(y_0,\omega)})$.

We now apply this computation procedure for the estimation of constant γ and the model's invariant distribution to the above baseline parameterization (5.8) with the added value $\pi = 0.1$. As before the scale parameter A is adjusted so as to keep the stationary solution $y^* = 1$ at the deterministic steady state in which $\varepsilon_t = 1$ for all $t \geq 0$. This value can be calculated from the first-order conditions. Since the model does not have an explicit solution, for the computation of the policy function we use a PEA algorithm with Chebyshev polynomial interpolation and collocation along the lines of Christiano and Fisher (2000). Our finest grid uses 8 collocation points for y over our estimated domain $[\widehat{y}_l, \widehat{y}_h]$ and 5 collocation points for ε over $[-0.032, 0.032]$. The Euler equation residuals generated by this policy function are of order at most 10^{-9}, which seems a very good approximation to the true solution [cf. Santos (2000)]. We let $\widehat{y}_{FIN}$ denote the random variable generated by the computed policy function with the finest grid. Under the above computational procedures we obtain that the ergodic set of $\widehat{y}_{FIN}$ is the interval of values $[0.9527, 1.0497]$. Over this set the maximum value for the computed derivative of the true policy function $D_1\varphi(y_0, \varepsilon_{t+1})$ is 0.875; hence, this latter value is an upper estimate for γ.

Tables III and IV replicate the above numerical experiments for the first and second order moments of $\widehat{y}_{FIN}$ as compared to those generated by two other PEA solutions obtained from coarser grids and by the linear approximation. The only difference here is that we consider that $E(\widehat{y}_{FIN})$ is an accurate estimate for $E(y)$, where y is the random variable generated by the true policy function. In Table III we can see the absolute value of the difference between $E(\widehat{y}_{FIN})$ and $E(\widehat{y}_{(m,n)})$, where $\widehat{y}_{(m,n)}$ is the random variable generated from a computed policy function by the PEA algorithm with m interpolation points for variable y and n interpolation points for variable ε. Note that the approximation error $|E(\widehat{y}_{FIN}) - E(y)|$ should be relatively small, since $|E(\widehat{y}_{FIN}) - E(\widehat{y}_{(m,n)})| \leq |E(\widehat{y}_{FIN}) - E(y)| + |E(y) - E(y_{(m,n)})|$, and by Theorem 6 the first term $|E(\widehat{y}_{FIN}) - E(y)|$ would be almost negligible if the approximate policy function generating $\widehat{y}_{FIN}$ is relatively close to the true solution. For the computation of the bounds (EBMTH6) and (EBMTH7) we let $\gamma = 0.875$, which is the estimate obtained in our above calculations.

The performance of our error bounds is quite similar to the previous case with a closed-form solution, which provides indirect evidence that our methods for the estimation of the true moments are quite accurate. There are two salient features in both Tables III and IV which are worth noticing. First, from the second column of these tables we can see that the linear approximation yields a better approximation for the mean and the variance than the two numerical PEA policy functions, even though the linear approximation is the most distant one to the policy function in the metric (3.1). Second, for the two PEA algorithms the error bound from Theorem 7 yields roughly the same values as the one from Theorem 6, whereas for the linear approximation the error bound from Theorem 7 is much tighter. As already discussed, this is because the linear policy function is a local approximation whereas the PEA algorithm yields a better global approximation with much less variation in the approximation error.

5.3 Non-optimal economies

Although the preceding examples illustrate that our results can be applied to a wide variety of economic models, the Feller property may not be satisfied in dynamic games and in competitive economies with financial frictions, taxes, externalities, and other market distortions. Therefore, an open issue is whether these models can be simulated by reliable numerical procedures. Kubler and Schmedders (2003) present a computational algorithm for models with heterogeneous agents and financial frictions, but some technical difficulties – mentioned below – are not fully addressed in this algorithm. This latter paper is a useful reference for computational work in this area [see also Rios-Rull (1999) for an early introduction to this literature].

One major problem for the computation of equilibria for non-optimal dynamic economies is that a Markov equilibrium solution may not exist or may be non-continuous [Kubler and Polemarchakis (2004) and Santos (2002)]. Following an approach pioneered by Kydland and Prescott (1980) and Duffie *et al*. (1994) one can nevertheless show for these economies the existence of a Markovian equilibrium solution in an enlarged state space that includes additional endogenous variables such as individual consumptions and asset prices. Duffie *et al*. (1994) establish the existence of a Markovian ergodic distribution under a randomization of the equilibrium correspondence. But this latter fundamental result does not appear to be very useful for the simulation of these economies. The randomization device used by these authors to prove the existence of an ergodic invariant distribution is not amenable to computation. Moreover, no general conditions are known to insure that Markovian equilibrium solutions of non-optimal dynamic economies satisfy the Feller property. Therefore, the above results on the continuity of invariant distributions and laws of large numbers may not hold for these economies.[10]

Another strand of the literature has been concerned with the existence of Markov equilibria over a minimal state space for monotone equilibrium dynamics [e.g., Bizer and Judd (1989), Coleman (1991), Hopenhayn and Prescott (1994) and Datta, Mirman and Reffett (2002)]. In general, the simulation of economies with monotone equilibrium dynamics seems to be less problematic – either the Markov equilibrium is continuous and so the Feller property is satisfied or it should be possible to establish suitable extensions of Theorems 2 and 3. In economic applications, however, monotonicity of the equilibrium dynamics may become quite restrictive. For the canonical one-sector growth model with taxes, externalities and other market distortions, monotone equilibrium dynamics follow from fairly mild restrictions on the primitives, but monotonicity is much harder to obtain in multisector models with heterogeneous agents and incomplete financial markets.

6 Concluding Remarks

Lucas (1980) argues that theoretical economics should provide fully articulated, artificial economic systems that can be simulated and contrasted with available data sets. For reasons of mathematical tractability and further considerations, most economic models are not conceived as computer programs that can be realized in a finite number of instructions. These models are characterized by non-linear utility and production functions, and are solved by numerical methods. Therefore, typically a researcher simulates a numerical approximation of the equilibrium solution in order to learn about the behavior of an economic model. But relatively little is known about the accuracy properties of numerical simulations; and some common practices may not be justifiable on theoretical grounds.

In this paper we have delved into the foundations of numerical simulation for stochastic dynamic models. Our analysis combines a continuity property of the correspondence of invariant distributions with a generalized law of large numbers to show that statistical, long-run average properties of a good enough numerical approximation will be close to those of the exact equilibrium solution. Moreover, under a certain contractivity condition we derive error

[10] In the existence result of Duffie *et al.* there is also the added problem that to compute the moments of the invariant distribution from numerical simulations one would need to start with an initial condition s_0 in the ergodic set; this may be an unattainable task if such set has Lebesque measure zero. One important aspect of Theorem 3 of practical relevance in applications is that it holds for every initial condition s_0.

bounds for these statistics. One issue of major concern in this investigation is that several important assumptions underlying existing results are hard to verify in practice. Existing results on continuity properties of the correspondence of invariant distributions are not directly applicable to numerical approximations as they are formulated under conditions on the transition probability P. Laws of large numbers build along the lines of Hypothesis D or the assumption of a unique invariant distribution. And error bounds concerning perturbations of stochastic contractions require that the approximating function $\widehat{\varphi}$ must also be a stochastic contraction and the distance of the approximation to the true equilibrium function must be in the sup norm. All these conditions are difficult to verify in economic models, and may not be preserved for numerical approximations or under further stochastic perturbations of the original system. Our analysis builds on the assumptions of a compact domain and a mild continuity condition on the equilibrium solution known as the Feller property. These hypotheses are usually validated from primitive conditions of economic models, and hold true for most numerical approximation schemes. There are various dynamic economic models that satisfy our assumptions, and hence it seems adequate to simulate these models by standard numerical methods. But as discussed in the preceding section, it not clear from existing theoretical results that Markovian equilibrium solutions of non-optimal dynamic economies would generally satisfy our continuity requirement. For these models our two basic results – the continuity property of invariant distributions and the generalized law of large numbers – may not hold, and hence the convergence of statistical, long-run average properties of sample paths of numerical approximations may break down.

The assumption of a compact domain can be weakened [e.g., see Billingsley (1968) and Futia (1982)]. The Feller property is a more delicate assumption, and can only be dispensed at the cost of imposing some other specific conditions [e.g., Hopenhayn and Prescott (1992)]. As discussed in example 5.1 above, the Feller property plays an essential role for the existence of an invariant distribution, the upper semicontinuity of the correspondence of invariant distributions, and our generalized version of the law of large numbers. Therefore, our analysis has singled out a mild continuity condition that seems indispensable for all our asymptotic results.

Our results are of further interest for the comparative analysis of stationary solutions, for the derivation of error bounds for invariant distributions of approximate solutions, and for the estimation of structural dynamic models. Error bounds for these invariant distributions were obtained under the assumption of a random contraction. In example 5.4 we developed some methods to identify this contractive property in models that do not admit analytical solutions. For the estimation of structural dynamic models, our results are essential to derive asymptotic properties of simulation-based estimators. Standard proofs for the consistency of these estimators [cf. Duffie and Singleton (1993)] assume the continuity of the expectations operator with respect to a vector of parameters and rely on a uniform law of large numbers over the whole parameter space. The continuity property of the expectations operator follows from Theorem 2 above. The uniform law of large numbers is a much stronger result than Theorem 3 and it is known to hold under further regularity assumptions such as Condition C or a monotonicity property on the equilibrium law of motion [Santos (2004)].

Finally, we would like to conclude with a brief discussion of an important and controversial issue among practitioners regarding the simulation of stochastic dynamic models. Macroeconomists often compute the distribution of the simulated moments from a large number of sample paths of length equal to that of the data sample. Thus, if the length of the data sample is N the simulation exercise would proceed by producing a large number of the model's sample paths all of the same length N. Then, the simulated moments computed over each of these sample paths are compared with those of the data sample. Our analysis, however, seems to suggest that a proper way to simulate a dynamic model would be produce one single sample path of arbitrarily large length. This is because by the law of large numbers the simulated moments would be approaching generically the moments of some invariant distribution of the model as the length of the sample path gets large. Therefore, our work offers no justification for replicating a large number of the model's sample paths – each of the same length N as that of the data sample – since they may be heavily influenced by the choice of the initial values s_0. We should also note that under some specific conditions on the dynamics of the model it is possible to bound the influence of an arbitrary initial value on the distribution of the vector of state variables. The case appears in Theorem 7 where for any given initial condition s_0 we can bound the distance between the N-step transition function and the invariant distribution of the original model. Then, in those situations it may be more operative to compute the moments of such N-step transition probability from a large number of sample paths of length N.

Dept. of Economics, Arizona State University, P.O. Box 873806, Tempe, AZ 85287-3806, U.S.A.; Manuel.Santos@asu.edu

and

Dept. of Economics, University of Miami, 5250 University Dr., Coral Gables, FL 33124-6550, U.S.A.; a.peraltaalva@miami.edu

7 Appendix

In addition to the proof of our main results, at the end of this Appendix we present a simplified version of Crauel (2002), Proposition 6.21, page 95.

Proof of Theorem 2: For an associated pair (φ, T^*) and a probability μ, let $\varphi \cdot \mu$ stand for $T^*\mu$. Then, following Dubins and Freedman (1966), page 839, the theorem will be established if we can show the continuity of the evaluation map $ev(\varphi, \mu) = \varphi \cdot \mu$. Recall that the space of probability measures is endowed with the topology of weak convergence, and $d(\varphi, \widehat{\varphi})$ in (3.1) is the distance function in the space of mappings. As is well known [e.g., see Shiryaev (1996)], the topology of weak convergence can be defined by the following metric:

$$d(\mu, \nu) = \sup_{f \in \mathcal{A}} \{ | \int f(s)\mu(ds) - \int f(s)\nu(ds) | \} \tag{7.1}$$

where $\mathcal{A}$ is the space of Lipschitz functions on S with constant $L \leq 1$ and such that $-1 \leq f \leq 1$.

Let f belong to $\mathcal{A}$. Then, for any two mappings φ and $\widehat{\varphi}$, and any two measures μ and ν, we have

$$\begin{aligned} & | \int f(s)[\varphi \cdot \mu(ds)] - \int f(s)[\widehat{\varphi} \cdot \nu(ds)] | \\ = & | \int [\int f(\varphi(s,\varepsilon))Q(d\varepsilon)]\mu(ds) - \int [\int f(\widehat{\varphi}(s,\varepsilon))Q(d\varepsilon)]\nu(ds) | \\ \leq & | \int [\int f(\varphi(s,\varepsilon))Q(d\varepsilon)]\mu(ds) - \int [\int f(\varphi(s,\varepsilon))Q(d\varepsilon)]\nu(ds) | \\ & + | \int [\int f(\varphi(s,\varepsilon))Q(d\varepsilon)]\nu(ds) - \int [\int f(\widehat{\varphi}(s,\varepsilon))Q(d\varepsilon)]\nu(ds) | \\ & \leq | \int [\int f(\varphi(s,\varepsilon))Q(d\varepsilon)][\mu(ds) - \nu(ds)] + d(\varphi, \widehat{\varphi}). \end{aligned}$$

The first inequality comes from the triangle inequality, and the second inequality follows directly from the definition of $d(\varphi, \widehat{\varphi})$ for f in $\mathcal{A}$.

Then, by (7.1) the theorem will be established if we can show that for every arbitrary $\eta > 0$ there exists a weak neighborhood $V(\mu)$ of μ such that for all ν in $V(\mu)$ and all f in $\mathcal{A}$,

$$| \int [\int f(\varphi(s,\varepsilon))Q(d\varepsilon)][\mu(ds) - \nu(ds)] | < \eta. \tag{7.2}$$

By the Arzela-Ascoli theorem, the set $\mathcal{A}$ is compact. Hence, we can find a finite set of elements $\{f^j\}$ such that for every f in $\mathcal{A}$ there exists an element f^j so that in the sup norm $\|f - f^j\| < \frac{\eta}{3}$. Also, by Assumption 2 the mapping $\int f(\varphi(s,\varepsilon))Q(d\varepsilon)$ is continuous in s. Hence, for every f^j there exists a weak neighborhood $V_j(\mu)$ such that for all ν in $V_j(\mu)$,

$$| \int [\int f^j(\varphi(s,\varepsilon))Q(d\varepsilon)][\mu(ds) - \nu(ds)]| < \frac{\eta}{3}.$$

Therefore, (7.2) must hold for all f with $\|f - f^j\| < \frac{\eta}{3}$. Finally, let $V(\mu) = \bigcap_j V_j(\mu)$. Then, (7.2) must hold for every ν in $V(\mu)$ and all f in $\mathcal{A}$. *Q.E.D.*

Proof of Theorem 3: For the proof of this theorem, it is convenient to let time n range from $-\infty$ to ∞ so as to consider sequences of the form $(\dots, \varepsilon_{-n}, \dots, \varepsilon_{-1}, \varepsilon_0, \varepsilon_1, \dots, \varepsilon_n, \dots)$. Using the construction in (3.5), we can then define a probability space $(\widehat{\Omega}, \widehat{\mathbb{F}}, \widehat{\lambda})$ over these doubly infinite sequences. Also, we shall view $\mathbb{F}$ as σ-subfield of $\widehat{\mathbb{F}}$. For each integer J we define the J-shift operator ϑ_J: $\widehat{\Omega} \rightarrow \widehat{\Omega}$, as $\vartheta_J(\dots, \varepsilon_{-n}, \dots, \varepsilon_{-1}, \varepsilon_0, \varepsilon_1, \dots, \varepsilon_n, \dots) = (\dots, \varepsilon_{-n+J}, \dots, \varepsilon_{-1+J}, \varepsilon_J, \varepsilon_{1+J}, \dots, \varepsilon_{n+J}, \dots)$. Note that the mapping ϑ_J is bijective and measurable. Hence, $(\widehat{\Omega}, \widehat{F}, \widehat{\lambda}, \vartheta_J)$ is a stationary ergodic system.

Let

$$G_N(\omega) = sup_{s_0 \in S}[\sum_{n=1}^{N} f(s_n(s_0, \omega))].$$

Then, by the measurable selection theorem [Hildenbrand (1974), Proposition 3, page 60], function G_N is measurable. Moreover, the following inequality must be satisfied:

$$G_{N+J}(\omega) \leq G_N(\vartheta_J(\omega)) + G_J(\omega)$$

for all positive integers N and J. Hence, by the subadditive ergodic theorem of Kingman (1968) there exists a constant H such that for λ-almost all ω,

$$\lim_{N \to \infty} \frac{G_N(\omega)}{N} = H. \tag{7.3}$$

Moreover, by the ergodic theorem, it is easy to see that

$$\int f(s)\mu^*(ds) \leq H \tag{7.4}$$

for any invariant distribution μ^*. Therefore, by (7.3) equality (3.6b) will be established if we can show that (7.4) holds in fact with equality.

For given $\epsilon > 0$, let $H_N^\epsilon = \{(\omega, s_0) : \frac{1}{N} \sum_{n=1}^{N} f(s_n(s_0, \omega)) \geq H - \epsilon\}$ for each N. Then, by Assumption 2 the set H_N^ϵ is measurable. Hence, by a standard argument (cf. *op. cit.*, page 54) there exists a measurable function h_N on $Proj_{\widehat{\Omega}}(H_N^\epsilon)$ such that $(\omega, h_N(\omega)) \in H_N^\epsilon$. Moreover, we can trivially extend this measurable function h_N over the whole domain $\widehat{\Omega}$.

Now, for any sample path $\{s_n(s_0,\omega)\}_{n=1}^N$ we define the empirical distribution $\mu_{(s_0,\omega)}^N$ on $\{s_n(s_0,\omega)\}_{n=1}^N$ as $\mu_{(s_0,\omega)}^N(s_n(s_0,\omega)) = \frac{1}{N}$ for every $s_n(s_0,\omega)$, for $1 \leq n \leq N$. Then, let

$$\mu^N = \int \mu_{(h_N(\omega),\omega)}^N \lambda(d\omega).$$

It follows from (7.3) and the construction of H_N^ϵ that if μ^* is a limit point of $\{\mu^N\}$ then for $f \in C(S)$,

$$\int f(s)\mu^*(ds) \geq H - \epsilon. \tag{7.5}$$

We now claim that every weak limit point μ^* of the sequence $\{\mu^N\}$ is an invariant distribution $\mu^* = T^*\mu^*$. Indeed, for every $f \in C(S)$ we have

$$|\int [\int f(\varphi(s,\varepsilon))Q(d\varepsilon)]\mu^*(ds) - \int f(s)\mu^*(ds)| =$$
$$\lim_{N\to\infty} |\int [\int f(\varphi(s,\varepsilon))Q(d\varepsilon)]\mu^N(ds) - \int f(s)\mu^N(ds)| = 0$$

where the first equality follows from Assumption 2 and the second equality follows from the construction of μ^N [cf. Crauel (2002), Theorem 6.12, page 87].

Therefore, by (7.5) there exists an invariant distribution μ^* such that $\int f(s)\mu^*(ds) = H$. This proves that (7.4) holds with equality for some invariant distribution μ^*, and completes the proof of (3.6b). The proof of (3.6a) proceeds in an analogous way. *Q.E.D.*

Proof of Theorem 6: For an initial point s_0, let $Z^n(s_0)$ denote the random vector

$\varphi(\varphi\cdots(\varphi(s_0,\varepsilon_1),\varepsilon_2)\cdots\varepsilon_n)$ and let $\widehat{Z}^n(s_0)$ denote $\widehat{\varphi}(\widehat{\varphi}\cdots(\widehat{\varphi}(s_0,\varepsilon_1),\varepsilon_2)\cdots\varepsilon_n)$. Then,

$$\begin{gathered}
|E[f(Z^n(s_0))]-E[f(\widehat{Z}^n(s_0))]| = \\
|E[f(\varphi(Z^{n-1}(s_0),\varepsilon_n))]-E[f(\widehat{\varphi}(\widehat{Z}^{n-1}(s_0),\varepsilon_n))]| \le \\
|E[f(\varphi(Z^{n-1}(s_0),\varepsilon_n))]-E[f(\varphi(\widehat{Z}^{n-1}(s_0),\varepsilon_n))]| + \\
|E[f(\varphi(\widehat{Z}^{n-1}(s_0),\varepsilon_n))]-E[f(\widehat{\varphi}(\widehat{Z}^{n-1}(s_0),\varepsilon_n))]| = \\
|E[E[f(\varphi(Z^{n-1}(s_0),\varepsilon_n))-f(\varphi(\widehat{Z}^{n-1}(s_0),\varepsilon_n))]|\varepsilon_n]| + \\
|E[f(\varphi(\widehat{Z}^{n-1}(s_0),\varepsilon_n))]-E[f(\widehat{\varphi}(\widehat{Z}^{n-1}(s_0),\varepsilon_n))]| \le \\
L\gamma E\|Z^{n-1}(s_0)-\widehat{Z}^{n-1}(s_0)\| + L\delta.
\end{gathered}$$

Observe that the first inequality comes from the triangle inequality. The second equality results by first conditioning on ε_n, and then by an application of the law of iterated expectations. And the last inequality follows from the assumptions of the theorem. Now, by a similar argument we get

$$\begin{gathered}
L\gamma E\|Z^{n-1}(s_0)-\widehat{Z}^{n-1}(s_0)\| \le \\
L\gamma^2 E\|Z^{n-2}(s_0)-\widehat{Z}^{n-2}(s_0)\| + L\gamma\delta.
\end{gathered}$$

Combining these inequalities and proceeding inductively it follows that

$$|E[f(Z^n(s_0))]-E[f(\widehat{Z}^n(s_0))]| \le \frac{L\delta}{1-\gamma} \text{ for all } n\ge 1. \tag{7.6}$$

Now, consider an invariant distribution $\widehat{\mu}^*$ of mapping $\widehat{\varphi}$. Then,

$$|\int E[f(Z^n(s_0))]\widehat{\mu}^*(ds_0)-\int f(s)\widehat{\mu}^*(ds)| =$$

$$|\int E[f(Z^n(s_0))]\widehat{\mu}^*(ds_0)-\int E[f(\widehat{Z}^n(s_0))]\widehat{\mu}^*(ds_0)| \le \frac{L\delta}{1-\gamma} \text{ for all } n\ge 1. \tag{7.7}$$

Observe that the equality comes from the fact that $\widehat{\mu}^*$ is an invariant distribution of $\widehat{\varphi}$; and the inequality in (7.7) is a consequence of (7.6). This inequality holds true for all $n\ge 1$; furthermore, by Theorem 5 for every s_0 the term $Ef(Z^n(s_0))$ converges uniformly to $\int f(s)\mu^*(ds)$. Therefore, by (7.7) we get

$$|\int f(s)\mu^*(ds)-\int f(s)\widehat{\mu}^*(ds)| \le \frac{L\delta}{1-\gamma}. \tag{7.8}$$

Q.E.D.

Proof of Theorem 7: Let $\left\{\varepsilon_1^j, ..., \varepsilon_N^j\right\}_{j=1}^M$ be an M-collection of sample paths of ε of length N. For fixed s_0, let $\left\{s_1^j, ..., s_N^j\right\}_{j=1}^M$ be the corresponding sample path generated by function φ, and $\left\{\widehat{s}_1^j, ..., \widehat{s}_N^j\right\}_{j=1}^M$ be the sample path generated by $\widehat{\varphi}$.

For given ε_1^i, we have

$$\begin{aligned}\left\|s_2^j - \widehat{s}_2^j\right\| = \left\|\varphi(s_1^i, \varepsilon_2^j) - \widehat{\varphi}(\widehat{s}_1^i, \varepsilon_2^j)\right\| \leq \\ \left\|\varphi(s_1^i, \varepsilon_2^j) - \varphi(\widehat{s}_1^i, \varepsilon_2^j)\right\| + \\ \left\|\varphi(\widehat{s}_1^i, \varepsilon_2^j) - \widehat{\varphi}(\widehat{s}_1^i, \varepsilon_2^j)\right\|.\end{aligned}$$

Summing up over j, for $j = 1, 2, \cdots, M$, we then obtain

$$\begin{aligned}\frac{1}{M}\sum_{j=1}^{M}\left\|s_2^j - \widehat{s}_2^j\right\| \leq \frac{1}{M}\sum_{j=1}^{M}\left\|\varphi(s_1^i, \varepsilon_2^j) - \varphi(\widehat{s}_1^i, \varepsilon_2^j)\right\| \\ +\frac{1}{M}\sum_{j=1}^{M}\left\|\varphi(\widehat{s}_1^i, \varepsilon_2^j) - \widehat{\varphi}(\widehat{s}_1^i, \varepsilon_2^j)\right\|.\end{aligned} \tag{7.9}$$

Since $\{\varepsilon_2^j\}$ is an *iid* process, the strong law of large numbers implies that for almost all sequences $\{\varepsilon_2^j\}$ as $M \to \infty$,

$$E[\left\|s_2^j - \widehat{s}_2^j\right\| |\varepsilon_1^i] \leq E[\left\|\varphi(s_1^j, \varepsilon_2^j) - \varphi(\widehat{s}_1^j, \varepsilon_2^j)\right\| |\varepsilon_1^i] + E[\left\|\varphi(\widehat{s}_1^i, \varepsilon_2^j) - \widehat{\varphi}(\widehat{s}_1^i, \varepsilon_2^j)\right\| |\varepsilon_1^i]. \tag{7.10}$$

Hence, by Condition C and the law of iterated expectations applied to (7.10) we get

$$E[\left\|s_2^j - \widehat{s}_2^j\right\|] \leq \gamma\widehat{\delta}_1 + \widehat{\delta}_2 \leq (1+\gamma)\widehat{\delta}_2.$$

The same argument applied to function f yields

$$E[\|f(s_2) - f(\widehat{s}_2)\|] \leq L(1+\gamma)\widehat{\delta}_2.$$

Moreover, proceeding by induction it follows that

$$E[\|f(s_n) - f(\widehat{s}_n)\|] \leq (L\sum_{i=1}^{n}\gamma^i)\widehat{\delta}_n \tag{7.11}$$

for $n = 1, \ldots, N$.

Note that the second term in (4.4) bounds (7.11). Moreover, the first term in (4.4) follows directly from Theorem 5 applied to the distribution of random vector $s_N(s_0,\omega)$. *Q.E.D.*

Finally, we present a simplified version of Crauel (2002), Proposition 6.21, page 95. To state this result, we draw on several terms defined in footnote 5 and the proof of Theorem 3. This theorem applies to moment functions $F(s,\omega)$ whereas Theorem 3 applies to moment functions $f(s)$ that only depend on s.

THEOREM 8 *Assume that $\{\varepsilon_n\}$ is a stationary and ergodic process. Assume that $F : S\times\widehat{\Omega} \to R$ is a bounded and measurable function such that $F(\cdot,\omega)$ is continuous for each $\omega \in \widehat{\Omega}$. Let S be a compact set. Assume that $\varphi : S \times \Omega \to R$ is a bounded and measurable function such that $\varphi(\cdot,\omega)$ is continuous for each $\omega \in \widehat{\Omega}$. Then, for $\widehat{\lambda}$-almost all ω,*

(7.12a) (*i*)

$$\lim_{N\to\infty}(min_{s_0\in S}[\frac{1}{N}\sum_{n=1}^{N}F(s_n(s_0,\omega),\vartheta_n(\omega))]) = min_{\nu\in I_\varphi}[\int_{S\times\widehat{\Omega}}F(s,\omega)\nu(ds,d\omega)]$$

(7.12b)

(*ii*)

$$\lim_{N\to\infty}(max_{s_0\in S}[\frac{1}{N}\sum_{n=1}^{N}F(s_n(s_0,\omega),\vartheta_n(\omega))]) = max_{\nu\in I_\varphi}[\int_{S\times\widehat{\Omega}}F(s,\omega)\nu(ds,d\omega)].$$

REFERENCES

Alpanda, S. and A. Peralta-Alva (2004): "The Oil Crisis, Energy-Saving Technological Change and the Stock Market Collapse of the 70's," manuscript, University of Miami.

Arellano, C. and E. G. Mendoza (2005): "Credit Frictions and 'Sudden Stops' in Small Open Economies: An Equilibrium Business Cycle Framework for Emerging Market Crises," NBER Working Paper 8880.

Arnold, L. (1998): *Random Dynamical Systems*. Berlin: Springer-Verlag.

Aruoba, S. B., J. Fernandez-Villaverde and J. F. Rubio-Ramirez (2003): "Comparing Solution Methods for Dynamic Equilibrium Economies," manuscript, Federal Reserve Bank of Atlanta.

Bajari, P., C. L. Benkard and J. Levin (2004): "Estimating Dynamic Models of Imperfect Competition," manuscript, Stanford University.

Bhattacharya, R. and M. Majumdar (2004): "Random Dynamical Systems: A Review," *Economic Theory*, 23, 13-38.

Billingsley, P. (1968): *Convergence of Probability Measures*. New York: John Wiley & Sons.

Bizer, K., and K. L. Judd (1989): "Taxation under Uncertainty," *American Economic Review*, 79, 331-336.

Boldrin, M., L. J. Christiano and J. D. M. Fisher (2001): "Habit Persistence, Assets Returns and the Business Cycle," *American Economic Review*, 91, 149-166.

Boldrin, M. and D. K. Levine (2001): "Perfectly Competitive Innovation," manuscript, UCLA.

Breiman, L. (1960): "The Strong Law of Large Numbers for a Class of Markov Chains," *Annals of Mathematical Statistics*, 31, 801-803.

Brock, W. A. and L. J. Mirman (1972): "Optimal Economic Growth and Uncertainty: The Discounted Case," *Journal of Economic Theory*, 4, 479-513.

Castaneda, A., J. Diaz-Gimenez and J.-V. Rios-Rull (2005): "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111, 818-857.

Chari, V. V., P. J. Kehoe and E. R. McGrattan (2002): "Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?," *Review of Economic Studies*, 69, 533-563.

Christiano, L. J., M. Eichembaum and C. L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113, 1-45.

Christiano, L. J. and J. D. M. Fisher (2000): "Algorithms for Solving Dynamic Models with Occasionally Binding Constraints," *Journal of Economic Dynamics and Control*, 24, 1179-1232.

Coleman, W. J. (1991): "Equilibrium in a Production Economy with an Income Tax," *Econometrica*, 59, 1091-1114.

Cooley, T. F. and V. Quadrini (2004a): "Financial Markets and Firm Dynamics," manuscript, Stern School of Business, NYU.

——— (2004b): "Optimal Monetary Policy in a Phillips-Curve World," *Journal of Economic Theory* , 118, 174-208.

Crauel, H. (1991): "Lyapunov Exponents of Random Dynamical Systems on Grassmanians," in *Lyapunov Exponents. Proceedings, Oberwolfach 1990, Springer Lecture Notes in Mathematics, vol. 1486,* ed. by L. Arnold, H. Crauel and J.-P. Eckmann. New York: Springer-Verlag.

——— (2002): *Random Probability Measures on Polish Spaces.* New York: Taylor & Francis.

Crawford, G. S. and M. Shum (2005): "Uncertainty and Learning in Pharmaceutical Demand," *Forthcoming in Econometrica.*

Datta, M., L. Mirman and K. Reffett (2002): "Existence and Uniqueness of Equilibrium in Distorted Dynamic Economies with Capital and Labor," *Journal of Economic Theory*, 103, 377-410.

den Haan, W. and A. Marcet (1994): "Accuracy in Simulations," *Review of Economic Studies*, 61, 3-17.

De Nardi, M. (2004): "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, 71, 743-768.

Doob, J. L. (1953): *Stochastic Processes.* New York: John Wiley & Sons.

Dubins, L. E. and D. Freedman (1966): "Invariant Probability Measures for Certain Markov Processes," *Annals of Mathematical Statistics*, 37, 837-848.

Dunford, N. and J. T. Schwartz (1958): *Linear Operators, Part I: General Theory.* New York: Interscience.

Duffie, D., J. Geanakoplos, A. Mas-Colell and A. McLennan (1994): "Stationary Markov Equilibria," *Econometrica*, 62, 745-781.

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

Ellison, G. and D. Fudenberg (1993): "Rules of Thumb for Social Learning," *Journal of Political Economy*, 101, 612-643.

Fernandez-Villaverde, J. and J. F. Rubio-Ramirez (2004): "Estimating Non-linear Dynamic Equilibrium Economies: A Likelihood Approach," manuscript, University of Pennsylvania.

Fershtman, C. and A. Pakes (2005): "Finite State Dynamic Games with Asymmetric Information: A Framework for Applied Work," manuscript, Harvard University.

Foley, D. and M. Hellwig (1975): "Asset Management with Trading Uncertainty," *Review of Economic Studies*, 42, 327-346.

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

Gallant, A. R. and G. Tauchen (1996): "Which Moments to Match?," *Econometric Theory*, 12, 657-681.

Gourieroux, C. and A. Monfort (1996): *Simulation-Based Econometric Methods.* Oxford, U.K.: Oxford University Press.

Hall, G. and J. Rust (2001): "The (S, s) Rule is an Optimal Trading Strategy in a Class of Commodity Price Speculation Problems," manuscript, University of Maryland.

Hildenbrand, W. (1974): *Core and Equilibria of a Large Economy.* Princeton, N.J.: Princeton University Press.

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

Jain, N. and B. Jamison (1967): "Contributions to Doeblin's Theory of Markov Processes," *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 8, 19-40.

Judd, K. L. (1998): *Numerical Methods in Economics*. Cambridge, MA.: The MIT Press.

Kingman, J. F. C. (1968): "The Ergodic Theory of Subadditive Stochastic Processes," *Journal of the Royal Statistical Society: Series B*, 30, 499-510.

Krengel, U. (1985): *Ergodic Theorems*. Berlin: Walter de Gruyter.

Kubler, F. and H. Polemarchakis (2004): "Stationary Markov Equilibria for Overlapping Generations," *Economic Theory*, 24, 623-643.

Kubler, F. and K. Schmedders (2003): "Stationary Equilibria in Asset-Pricing Models with Incomplete Markets and Collateral," *Econometrica*, 71, 1767-1793.

Kydland, F. E. and E. C. Prescott (1980): "Dynamic Optimal Taxation, Rational Expectations and Optimal Control," *Journal of Economic Dynamics and Control*, 2, 79-91.

——— (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50, 1345-1370.

Ljungqvist, L. and T. J. Sargent (2004): *Recursive Macroeconomic Theory*. Cambridge, MA.: The MIT Press.

Lucas, R. E. (1980): "Methods and Problems in Business Cycle Theory," *Journal of Money, Credit and Banking*, 12, 696-715.

Manuelli, R. (1985): "A Note on the Behavior of Solutions to Dynamic Stochastic Models," manuscript, University of Wisconsin, Madison.

Meyn, S. P. and R. L. Tweedie (1993): *Markov Chains and Stochastic Stability*. New York: Springer.

Norman, F. M. (1972): *Markov Processes and Learning Models*. New York: Academic Press.

Pakes, A. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patents Stocks," *Econometrica*, 54, 755-785.

Reiter, M. (2000): "Estimating the Accuracy of Numerical Solutions to Dynamic Optimization Problems," manuscript, Universitat Pompeu Fabra.

Restuccia, D. and C. Urrutia (2004): "Intergenerational Persistence of Earnings: The Role of Early and College Education," *American Economic Review*, 94, 1354-1378.

Rios-Rull, J.-V. (1999): "Computation of Equilibria in Heterogeneous Agent Models," in *Computational Methods for the Study of Dynamic Economies: An Introduction,* ed. by R. Marimon and A. Scott. Oxford, U.K.: Oxford University Press.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999-1033.

——— (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, vol. 4, ed. by R. F. Engle and D. L. McFadden. Amsterdam: Elsevier.

Sanghvi, A. P. and M. J. Sobel (1976): "Bayesian Games as Stochastic Processes," *International Journal of Game Theory*, 5, 1-22.

Santos, M. S. (1999): "Numerical Solution of Dynamic Economic Models," in *Handbook of Macroeconomics*, vol. 1, ed. by J. B. Taylor and M. Woodford. Amsterdam: Elsevier.

——— (2000): "Accuracy of Numerical Solutions Using the Euler Equation Residuals," *Econometrica*, 68, 1377-1402.

——— (2002): "On Non-existence of Markov Equilibria in Competitive-Market Economies," *Journal of Economic Theory*, 105, 73-98.

——— (2004): "Simulation-Based Estimation of Dynamic Models with Continuous Equilibrium Solutions," *Journal of Mathematical Economics*, 40, 465-491.

Schenk-Hoppe, K. R. and B. Schmalfuss (2001): "Random Fixed Points in a Stochastic Solow Growth Model," *Journal of Mathematical Economics*, 36, 19-30.

Schmalensee, R. (1975): "Alternative Models of Bandit Selection," *Journal of Economic Theory*, 10, 333-342.

Schmalfuss, B. (1996): "A Random Fixed-Point Theorem Based upon Lyapunov Exponents," *Random and Computational Dynamics*, 4, 257-268.

Shiryaev, A. N. (1996): *Probability.* New York: Springer-Verlag.

Stenflo, O. (2001): "Ergodic Theorems for Markov Chains Represented by Iterated Function Systems," *Bulletin of the Polish Academy of Sciences: Mathematics*, 49, 27-43.

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

Tan, W. (2004): "A Dynamic Analysis of the U.S. Cigarette Market and Antismoking Policies," manuscript, Johns Hopkins University.

Taylor, J. B. and H. Uhlig (1990): "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods," *Journal of Business and Economic Statistics*, 8, 1-18.

Walters, P. (1982): *An Introduction to Ergodic Theory.* New York: Springer-Verlag.

Wolpin, K. I. (1984): "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," *Journal of Political Economy*, 92, 852-874.

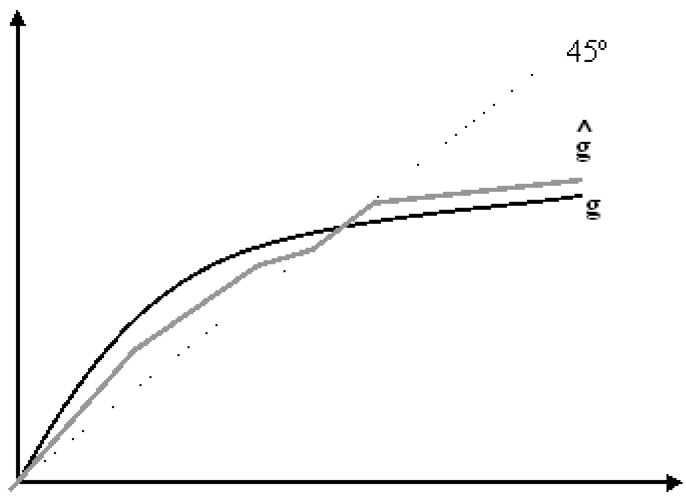

FIGURE 1-Multiplicity of stationary solutions for the numerical approximation

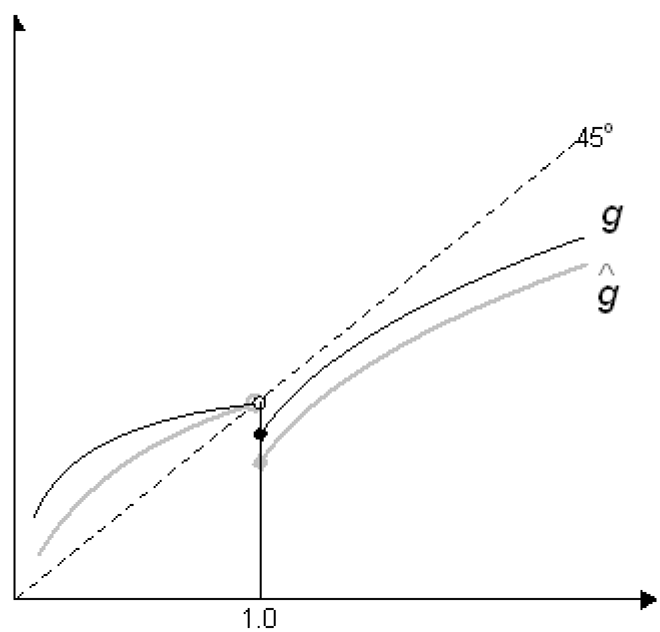

FIGURE 2-Non-existence of an invariant distribution

TABLE I

Error Bounds for the Mean

	$\lvert\mathrm{E}(\log y)\text{-}\mathrm{E}(\log y_L)\rvert$	EBMTH6	EBMTH7
α=0.34	1.4402E-4	1.7128E-2	9.2438E-3
α=0.17	1.1277E-4	2.7286E-3	2.7101E-3
α=0.68	5.1889E-4	2.9783E-1	2.7808E-2

TABLE II

Error Bounds for the Variance

	$\lvert\mathrm{Var}(\log y)\text{-}\mathrm{Var}(\log y_L)\rvert$	EBVTH6	EBVTH7
α=0.34	1.0216E-5	8.3046E-4	4.4819E-4
α=0.17	2.7856E-6	1.5198E-4	1.0448E-4
α=0.68	4.6712E-5	2.9783E-2	2.7808E-3

TABLE III

Error Bounds for the Mean

	$\lvert\mathrm{E}(\hat{y}_{FIN})\text{-}\mathrm{E}(y)\rvert$	EBMTH6	EBMTH7
$y = y_{2,2}$	7.0895E-2	1.5354E-1	1.3620E-1
$y = y_{3,3}$	1.6176E-3	1.8882E-2	1.8303E-2
$y = y_L$	8.2885E-5	1.6827E-2	1.2555E-3

TABLE IV

Error Bounds for the Variance

	$\lvert\mathrm{Var}(\hat{y}_{FIN})\text{-}\mathrm{Var}(y)\rvert$	EBVTH6	EBVTH7
$y = y_{2,2}$	6.9253E-5	7.4444E-2	6.6037E-2
$y = y_{3,3}$	6.9051E-5	9.1550E-4	8.8743E-4
$y = y_L$	5.9681E-5	8.1586E-4	6.0873E-5