Accuracy Properties of the Statistics from Numerical Simulations

Manuel S. Santos
Department of Economics
Arizona State University

Adrian Peralta-Alva
Department of Economics
University of Minnesota
and Federal Reserve Bank of Minneapolis

Preliminary and Incomplete Version
October 18, 2001

Abstract

This paper is concerned with the approximation errors involved in the statistics from the sample paths generated by an approximate or computed policy function. Under fairly general assumptions, we establish convergence of these statistics to their true expected values as the approximation error of the policy function converges to zero. Furthermore, under additional regularity conditions we also show that the approximation errors of these statistics are of the same order of magnitude as that of the computed policy function. The constants bounding these orders of convergence are explicitly computed, and are shown to depend on primitive parameters.

KEYWORDS: Statistic, Random Dynamical System, Invariant Probability Measure, Convergence, Approximation Error.

1 Introduction

One common procedure for the calibration and testing of an economic model is to lay out a notion of distance in which the model's performance is compared against a selected set of basic observations. In the real business cycle literature [e.g. Cooley and Prescott (1995)], this notion of distance is determined as the difference between the second-order moments generated by numerical simulations and the sample moments. In financial economics, measures of skewness and kurtosis are often invoked. These are just some simple examples of statistics that can be considered.

As emphasized in Christiano and Eichenbaum (1992), the moments from time series data are subject to sampling error. This error can generally be estimated by well established statistical techniques, which usually yield confidence intervals for the true moments of the postulated data generating process. Much less attention, however, has been devoted to the approximation error of the moments generated by numerical simulations. This approximation error may be hard to appraise for stochastic models with pronounced non-linearities.

A widely used method for checking the accuracy properties of a given numerical algorithm [e.g., Taylor and Uhlig (1990)] is to compare the moments generated by the putative solution with those of the available exact solution or with those generated from a more reliable algorithm. Of course, it should be clear that these simple accuracy checks may only be justified under certain stability properties of the stationary or invariant distribution of the original economic model. Indeed, a simple example below illustrates that an arbitrarily small perturbation of the transition probability defining a Markov process may change drastically the set of invariant distributions. Correspondingly, a very small error in the specification of the law of motion of a dynamic equilibrium system may propagate over time, and alter substantially the simulated moments.

In this paper we establish several convergence results for the statistics of sample paths generated by approximate stochastic equilibrium laws of motion or decision rules. If the original law of motion contains a unique invariant distribution, then we show that the statistics from numerical simulations converge to their true expected values as the approximate policy function converges to the original one. In addition, if the original law of motion satisfies a certain contractive property, then we show that the approximation errors of the simulated statistical values are of the same order of magnitude as the error of the approximate decision rule. As is well understood, this contractive property can usually be verified for solutions involving small shocks around a stationary point, and it guarantees uniqueness of the invariant distribution. In the absence of this condition, we cannot insure that both approximation errors are of the same order of magnitude.

There are several available results on the accuracy properties of the value and policy functions based upon the size of the Euler equation residuals or on the grid size of some discretizations [cf. Santos (1999)]. Given the impossibility of sorting out sample paths in stochastic modelizations, however, error bounds on the policy and value functions are not operative for the calibration and testing of a model. For such purposes, one usually needs to define a notion of distance for the model's performance based upon a selected set of statistics as compared to their data counterparts, and where proper account can be taken of approximation and sampling errors.

In our study we make repetitive use of standard results from the theory of Markov processes concerning the existence and uniqueness of invariant distributions and laws of large numbers. [For most of purposes, these results are adequately covered in Futia (1982) and Stokey, Lucas and Prescott (1989); see, of course, the further sources quoted therein.] As part of our analysis, we also establish certain basic convergence properties of invariant distributions for random dynamic systems; these results seem to be important in their own right. Moreover, contrary to some common practices, our analysis suggests a specific way for the simulation and testing of dynamic models. We take up this latter issue in the final section.

The paper is structured as follows. In Section 2 we set out our basic framework of analysis. Our main analytic results are presented in Section 3. The assumptions underlying these results are further discussed in Section 4 along with some illustrative examples. We then conclude in Section 5 with some final remarks.

2 Random Dynamical Systems

In several economic models, the equilibrium law of motion of the state variables can be specified by a dynamical system of the following form

$$\begin{aligned} z_{n+1} &= \xi(z_n, \varepsilon_n) \\ k_{n+1} &= g(k_n, z_n), \qquad n = 0, 1, 2, \ldots \end{aligned} \tag{RDS}$$

Here, z may represent a vector made up of stochastic exogenous variables such as some indices of factor productivity or market prices. This random vector is always contained in a space Z, and is subject to an i.i.d. shock, ε, in a set of "events" E, governed by a probability law Q. Then, k is a vector of endogenous state variables, and 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 on a set K. Hence, $s = (k, z)$ is the vector of state variables belonging to the set $S = K \times Z$. Let $(S, \mathbb{S})$ denote a measurable space.

Assumption 1 *The set $S = K \times Z \subset R^l \times R^m$ is compact, and $\mathbb{S}$ is the Borel $\sigma-$field. $(E, \mathbb{E}, Q)$ is a probability space.*

Assumption 2 *Function $g : K \times Z$ is continuous. For each z in Z, $\xi(z, \cdot)$ is measurable, and for each ε in E, $\xi(\cdot, \varepsilon)$ is continuous.*

For expository purposes, let us express (RDS) in the more compact form

$$s_{n+1} = \psi(s_n, \varepsilon_n) \quad n = 0, 1, 2, \ldots \tag{2.1}$$

Under the foregoing regularity conditions, the evolution of state variable s may be described by a transition probability on the state space. For a Borel set $A \in \mathbb{S}$, let $A_s^{\psi^{-1}} = \{\varepsilon \text{ in } E | \psi(s, \varepsilon) \in A\}$. Then, one can define a function $P : S \times \mathbb{S}$ by the rule

$$P(s, A) = Q(A_s^{\psi^{-1}}) \tag{2.2}$$

for every s in S and every Borel set A in $\mathbb{S}$. As is well known [cf., Futia (1982, Th. 5.2)] P is a transition function on $(S, \mathbb{S})$. That is, for each s in S, $P(s, \cdot)$ is a probability on $\mathbb{S}$, and for each $A \in \mathbb{S}$, $P(\cdot, A)$ is measurable. Moreover, it follows from a standard result in measure theory that function ψ is measurable on the product space $(E \times S, \mathbb{E} \times \mathbb{S})$.

The transition probability P completely characterizes the Markov process. For any initial probability measure μ_0 on $(S, \mathbb{S})$, the evolution of future distributions is determined by the law

$$\mu_{n+1}(A) = \int P(s, A)\mu_n(ds) \tag{2.3}$$

for all Borel sets $A \in \mathbb{S}$ and $n \geq 0$. An invariant distribution or invariant probability measure is a fixed point of (2.3). The invariant distribution is some sort of stochastic steady state on which the system may settle down. Then the analysis of invariant distributions seems a very first step to understand the dynamics of the system. Further, uniqueness of the invariant distribution is a highly desirable property, since the model has sharper predictions. In this paper we shall confine the analysis to dynamical systems with a unique, globally stable invariant distribution, but some convergence results extend to more general settings.

In many economic applications, the most one can hope for is to get an approximate function $\widehat{\psi}$, which may have been calculated by some numerical procedure. Even though function ψ remains unknown, with the aid of some accuracy checks [cf. Santos (2000)] we may be able to bound the distance between functions ψ and $\widehat{\psi}$. Of course, function $\widehat{\psi}$ may give rise to a transition probability $\widehat{P}$ on $(S, \mathbb{S})$. But even if function $\widehat{\psi}$ is arbitrarily close to ψ, the following questions may come to the fore:

How different are the dynamics generated by function $\widehat{P}$?

More specifically, how different is the set of invariant distributions under $\widehat{P}$?

And, what can be said about the accuracy of the statistics from numerical simulations?

As the present well known example illustrates, without further specific assumptions we cannot expect good stability results.

EXAMPLE 2.1: The state space S is a discrete set with three possible states, s_1, s_2, s_3. The transition probability P on S is defined by the following Markov matrix

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

Here, the rows of Π describe the corresponding functions $P(s_i, \cdot)$, for $i = 1, 2, 3$. That is, each row i specifies the probabilities of moving from state i to any state in S. Note that

$\Pi^n = \Pi$ for all $n \geq 1$. Hence, probabilities $p = (1, 0, 0)$, and $p = (0, 1/2, 1/2)$ are invariant distributions under Π. All other invariant distributions are convex combinations of these two probabilities. Therefore, $\{s_1\}$ and $\{s_2, s_3\}$ are the only ergodic sets.

Let us now perturb Π slightly so that the new stochastic matrix is the following

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

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 invariant distribution under $\widehat{\Pi}$. Moreover, $\{s_1\}$ is a transient state, and $\{s_2, s_3\}$ is the only ergodic set. Consequently, a small perturbation on a transition probability P may lead to a pronounced change in the set of invariant distributions. Indeed, small errors may propagate over time and alter the existing ergodic sets.

If we consider the correspondence from Π to the set of invariant distributions under Π, then the example shows that this correspondence fails to be lower semicontinuous. A key step below is to establish that this correspondence is upper semicontinuous for Markov processes generated by random dynamical systems. A much simpler reasoning is sufficient to validate upper semicontinuity for the present case with a finite number of states. Hence, the lack of continuity occurs from sudden implosions in the set of invariant distributions.

Even more troublesome may be the fact that the matrix sequence $\{\Pi^n\}_{n\geq 1}$ does not always converge, and hence for certain paths the system may not settle down to a long-run distribution or invariant probability. There could be situations in which results along lines of laws of large numbers fail to exist, and hence statistical inference may lose its full strength. As is well known [e.g. Stokey, Lucas and Prescott (1989, Ch. 11)], for some important purposes what it is needed is a weak notion of convergence of the mapping $A_N(\Pi) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Pi^n$. For discrete state spaces, $A_N(\Pi)$ always converges to an invariant stochastic matrix. But for more general state spaces, further regularity conditions are required. A typical case is if the Markov process satisfies *Doeblin's Condition* below.

3 Main Results

This section contains our main analytical results. Our first goal is to establish convergence of the statistics generated by numerical approximations to their expected values under the invariant distribution associated with ψ. Then we impose a contractive property on ψ and show that the approximation error involved in these statistical measures is of the same order of magnitude as the error of the numerical approximation. As a step toward this analysis, we first review some fundamental results in the theory of Markov processes.

3.1 Mathematical Preliminaries

Let $B(S)$ be the set of all bounded, $\mathbb{S}-$measurable, real-valued functions on S. Then, $B(S)$ is a Banach space when endowed with the norm $\|f\| = \sup_{s\in S} |f(s)|$. A transition probability P on $(S, \mathbb{S})$ defines a Markov operator T from $B(S)$ to itself via the following integration law

$$(Tf)(s) = \int f(s')P(s, ds') \quad \text{all } s \in S. \tag{3.1}$$

We can associate with every operator T its adjoint T^*. Thus, if we write $\langle f, \mu\rangle = \int f(s)\mu(ds)$ then $\langle Tf, \mu\rangle = \langle f, T^*\mu\rangle$ for all f and μ. Moreover, T^* can be specified by the corresponding integration law

$$(T^*\mu)(A) = \int P(s, A)\mu(ds) \quad \text{all } A \in \mathbb{S} \tag{3.2}$$

mapping the space of probability measures on $(S, \mathbb{S})$ to itself. We shall also refer to operators T^n and T^{*n}, for all n. Let $P^n(s, A) = (T^{*n}\delta_s)(A)$, where δ_s is the Dirac probability measure assigning unit value to point s.

A sequence $\{\mu_j\}$ of probability measures on $(S, \mathbb{S})$ 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 the space $C(S)$ of continuous functions. 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. It should be remembered that every weakly convergent sequence $\{\mu_j\}$ of probability measures has a unique limit point; further, in the weak topology the set of all probability measures on $(S, \mathbb{S})$ is compact.

THEOREM 3.1 *There exists a probability measure μ^* such that $\mu^* = T^*\mu^*$.*

The existence of an invariant probability measure μ^* can be established as follows. First, one can show [e.g., see Futia (1982, Prop. 5.6)] that T maps the space $C(S)$

of continuous functions into itself. Then, operator T^* must be weakly continuous. Moreover, in the weak topology the set of all probability measures on $(S, \mathbb{S})$ is compact, and it is obviously a convex set. Therefore, T^* must have a fixed-point solution, $\mu^* = T^*\mu^*$.

Define the linear operator $A_N(T) = \frac{1}{N}\sum_{n=1}^{N} T^n$. Then, $A_N(T^*) = \frac{1}{N}\sum_{n=1}^{N} T^{*n}$ is the adjoint of operator $A_N(T)$. By a compactness argument, for every probability measure μ on $(S, \mathbb{S})$, the sequence $\{A_N(T^*)\mu\}$ must have a weak limit point and so there must be a convergent subsequence $\{A_{N'}(T^*)\mu\}$. Every limit point μ^* of $\{A_N(T^*)\mu\}$ is an invariant probability measure on $(S, \mathbb{S})$. Proper convergence of the full sequence is guaranteed if there exists a unique solution $\mu^* = T^*\mu^*$, or if the transition probability P satisfies Doeblin's Condition.

CONDITION D: There exists a probability measure λ, an integer n, and a δ with $0 < \delta < 1$, such that if $A \in \mathbb{S}$ and $\lambda(A) \leq \delta$, then $P^n(s, A) \leq 1 - \delta$ for all s.

The importance of this condition is exemplified in the following result [cf. Stokey, Lucas and Prescott (1989)]

THEOREM 3.2 *Assume that P satisfies Condition D. Then*

(a) S can be partitioned into a transient set and M ergodic sets, where $1 \leq M \leq \frac{\lambda(S)}{\delta}$.

(b)For every probability measure μ, the sequence $\{A_N(T^*)\mu\}$ converges to an invariant probability, $\mu^* = T^*\mu^*$.

(c) Each of the M ergodic sets is the support[1] of a unique probability measure. Every other invariant probability measure μ^* is a convex combination of these M probabilities.

Finally, we shall also invoke the following strong law of large numbers due to Breiman (1960). Let $\Omega = S \times S \times S \times \ldots$ be the infinite product of the state space. Let $\mathbb{G} = \mathbb{S} \times \mathbb{S} \times \mathbb{S} \times \ldots$ be the product $\sigma-$field. It is well known that for any s_0 in S the transition function P induces a corresponding product probability measure $\mu(s_0, \cdot)$ on $(\Omega, \mathbb{G})$. For a fixed s_0 in S and $\omega = \{\varepsilon_n\}_{n\geq 0}$ let $s(\omega) = \{s_n(\omega)\}_{n\geq 1}$ be the infinite sequence of state variables generated by (2.1). That is, $s_n(\omega) = \psi(s_{n-1}(\omega), \varepsilon_{n-1})$ for all $n \geq 1$. Recall that if a property holds for all ω in a set of probability one, then it is said to hold almost surely (a. s.).

THEOREM 3.3 [cf., Stokey, Lucas and Prescott (1989, Th. 14.7)]: *Assume that $\{A_N(T^*)\mu\}$ converges weakly to μ^*. Then, for every s_0 in S and every function f in $C(S)$,*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(s_n(\omega)) = \int f(s)\mu^*(ds), \qquad \mu(s_0, \cdot) - a.\ s.$$

[1]The support $\sigma(\mu^*)$ of a probability measure μ^* is the smallest close set F such that $\mu^*(F) = 1$.

3.2 Convergence of the Statistics from Numerical Approximations

For every f in $C(S)$ and "almost all" sample paths $\{\widehat{s}_n(\omega)\}_{n\geq 1}$ generated under an approximate function $\widehat{\psi}$, we now establish the convergence of $\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{\infty} f(\widehat{s}(\omega))$ to $\int f(s)\mu(ds)$, as function $\widehat{\psi}$ approaches ψ. To this end, we assume that system (2.1) contains a unique invariant distribution. A key result in this part is Proposition 3.4, where an upper semicontinuity property of the set of invariant probability measures is established. This convergence property holds for every Markov process induced by a random dynamical system.

CONDITION U:

(i)There exists a unique probability measure μ^* such that $\mu^* = T^*\mu^*$.

(ii) $\sigma(\mu^*)$ has non-empty interior.

There are many economic applications in which Condition U is satisfied, as for instance in the stochastic growth model of Brock and Mirman (1972), or in the industry investment model of Lucas and Prescott (1971). As discussed below, with some minor qualifications our main results can be extended to systems with multiple invariant distributions.

Even though in many economic models the invariant distribution μ^* is unique, it is rarely the case that the equilibrium law of motion corresponding to ψ has an analytical representation. But via a numerical procedure we may be able to obtain an approximation $\widehat{\psi}$ of the true function ψ. We shall always suppose that this approximation $\widehat{\psi}$ can be specified by a pair of functions $(\widehat{\xi},\widehat{g})$ satisfying Assumptions 1-2. Then, for each $\widehat{\psi}$ there is associated a transitional probability $\widehat{P}$ and a corresponding pair of operators $\widehat{T}$ and $\widehat{T}^*$. In some cases we assume that the transition probability $\widehat{P}$ generated by a random dynamical system $\widehat{\psi}$ satisfies Condition D. As discussed below, Condition D is a fairly mild restriction.

For a vector-valued function $\psi = (\ldots,\psi^i,\ldots)$, let $\|\psi\| = \max_{1\leq i\leq l+m}\left\|\psi^i\right\|$. In the following result, convergence of the sequence $\{\widehat{\psi}_j\}$ should be understood as uniform convergence, that is, in the metric induced by this norm. Also, every approximate function $\widehat{\psi}_j$ defines the associated triple $(\widehat{P}_j,\widehat{T}_j,\widehat{T}_j^*)$. By Theorem 3.1, there always exists a fixed-point solution, $\widehat{\mu}_j^* = \widehat{T}_j^*\widehat{\mu}_j^*$. There is no guarantee, however, that the invariant probability measure $\widehat{\mu}_j^*$ is unique, since $\widehat{\psi}_j$ may have been derived from some numerical procedure.

PROPOSITION 3.4 *Assume that $\widehat{\psi}_j \to_j \psi$. Let $\{\widehat{\mu}_j^*\}$ be a corresponding sequence of invariant probability measures, $\widehat{\mu}_j^* = \widehat{T}^*\widehat{\mu}_j^*$, for every j. Then, every weak limit point μ^* of $\{\widehat{\mu}_j^*\}$ is an invariant probability measure, $\mu^* = T^*\mu^*$.*

Observe that the proposition asserts the bilinear convergence of $\langle \widehat{T}_j^*, widehat\mu_j^*\rangle =$

$\widehat{T}_j^* \widehat{\mu}_j^*$ to $\langle T^*, \mu^* \rangle = T^* \mu^*$. This property is usually much stronger than the standard notion of convergence in the weak topology. Stokey, Lucas and Prescott (1989, Theorem 12.3) present a related result [see also Manuelli (1985)], but their condition (b) seems to be hard to check in applications. The practical importance of Proposition 3.4 is that it applies to all Markov processes generated by random dynamical systems satisfying Assumptions 1-2.

The following result is a simple consequence of Proposition 3.4.

Corollary 3.5 *For fixed ψ, assume that P satisfies Condition U. For all $\widehat{\psi}_j$, assume that $\widehat{P}_j$ satisfies Condition D. Then, there exists $\delta > 0$ such that for every $\widehat{\psi}_j$ with $\left\| \widehat{\psi}_j - \psi \right\| \leq \delta$, there is a unique invariant distribution, $\widehat{\mu}_j^* = \widehat{T}_j^* \widehat{\mu}_j^*$.*

As before, for a fixed s_0 and a sequence of realizations, $\omega = \{\varepsilon_n\}_{n \geq 0}$, let $s(\omega) = \{s_n(\omega)\}_{n \geq 1}$ be the sample path generated by ψ. Also, let $\widehat{s}_j(\omega) = \{\widehat{s}_{jn}(\omega)\}$ be the corresponding sample path generated by $\widehat{\psi}_j$.

Theorem 3.6 *For a given ψ, assume that P satisfies Condition U(i). For all $\widehat{\psi}_j$, assume that $\widehat{P}_j$ satisfies Condition D. Let $\widehat{\psi}_j \rightarrow_j \psi$. Then, for every s_0 in S and every function f in $C(S)$,*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(\widehat{s}_{jn}) \longrightarrow_j \int f(s) \mu^*(ds), \quad \widehat{\mu}_j^*(s_0, \cdot) - \textit{ and } \mu^*(s_0, \cdot) - \textit{a. s.}$$

Proof: Observe that

$$| \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(\widehat{s}_{jn}(\omega)) - \int f(s) \mu^*(ds) | \leq$$

$$| \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f(\widehat{s}_{jn}(\omega)) - \int f(s) \widehat{\mu}_j^*(ds) | + | \int f(s) \widehat{\mu}_j^*(ds) - \int f(s) \mu^*(ds) |.$$

By Theorems 3.2 and 3.3, the first term is equal to zero. Regarding the second term, note that under Condition U(i) the invariant distribution μ^* is unique. Moreover, by Proposition 3.4 the sequence $\{\widehat{\mu}_j^*\}$ must weakly converge to μ^*. Therefore, $\int f(s) \widehat{\mu}_j^*(ds) \longrightarrow_j \int f(s) \mu^*(ds)$. These two facts together prove the theorem.

It is readily seen that Theorem 3.6 implies the convergence of the first- and second-order moments from numerical simulations. Let $\widehat{E}_j(\widehat{s}_j) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \widehat{s}_{jn}(\omega)$, and $E(s) = \int s \mu^*(ds)$.

Corollary 3.7 *Under the conditions of Theorem 3.6,*

$$\widehat{E}_j(s) \longrightarrow_j E(s), \quad \mu_j^*(s_0,\cdot) - \text{ and } \mu^*(s_0,\cdot) - a.\ s.$$

Let $s = (\ldots, s^i, \ldots, s^r, \ldots)$, for s in S. Then $\widehat{Var}_j(\widehat{s}_j{}^i) = \lim_{N\to\infty} \sum_{n=1}^N [\widehat{s}^i_{jn}(\omega) - \widehat{E}_j(\widehat{s}^i_j)]^2$, and $\widehat{Cov}_j(\widehat{s}_j{}^i, \widehat{s}_j{}^r) = \lim_{N\to\infty} \sum_{n=1}^N [\widehat{s}^i_{jn}(\omega) - \widehat{E}_j(\widehat{s}^i_j)][\widehat{s}^r_{jn}(\omega) - \widehat{E}_j(\widehat{s}^r_j)]$. Also, define $Var(s^i) = \int (s^i - E(s^i))^2 \mu^*(ds)$, and $Cov(s^i, s^r) = \int ((s^i - E(s^i))(s^r - E(s^r))\mu^*(ds)$.

Corollary 3.8 *Under the conditions of Theorem 3.6,*

$$\widehat{Var}_j(\widehat{s}_j{}^i) \longrightarrow_j Var(s^i), \quad \widehat{\mu}_j^*(s_0,\cdot) - \text{ and } \mu^*(s_0,\cdot) - a.\ s.$$

Corollary 3.9 *Under the conditions of Theorem 3.6,*

$$\widehat{Cov}_j(\widehat{s}_j{}^i, \widehat{s}_j{}^r) \longrightarrow Cov(s^i, s^r), \quad \widehat{\mu}_j^*(s_0,\cdot) - \text{ and } \mu^*(s_0,\cdot) - a.\ s.$$

Proof of Proposition 3.4: Without loss of generality, assume that $\{\widehat{\mu}_j^*\}$ converges weakly to μ^*, where $\widehat{\mu}_j^* = \widehat{T}_j^* \widehat{\mu}_j^*$, for every j. We have to show that $\mu^* = T^*\mu^*$.

For every f in $C(S)$ we have

$$(\widehat{T}_j f)(s) = \int f(s')\widehat{P}_j(s, ds') = \int f(\widehat{\psi}_j(s,\varepsilon))Q(d\varepsilon).$$

Since S is a compact set, function f is uniformly continuous. Hence, for every $\eta > 0$ there exists $\delta > 0$ such that

$$|f(\widehat{\psi}(s,\varepsilon)) - f(\psi(s,\varepsilon))| < \eta$$

whenever

$$\left\|\widehat{\psi}_i(\cdot,\varepsilon) - \psi(\cdot,\varepsilon)\right\| < \delta \text{ for all } \varepsilon.$$

Therefore, for every $\eta > 0$ there exists J such that

$$\left\|\widehat{T}_j f - Tf\right\| \leq \eta \text{ for all } j \geq J.$$

We then have,

$$\begin{aligned} |\langle f, \mu^*\rangle - \langle f, T^*\mu^*\rangle| &\leq \\ |\langle f, \mu^*\rangle - \langle f, \widehat{\mu}_j^*\rangle| + |\langle f, \widehat{T}_j^*\widehat{\mu}_j^*\rangle - \langle f, T^*\widehat{\mu}_j^*\rangle| &+ \\ |\langle f, T^*\widehat{\mu}_j^*\rangle - \langle f, T^*\mu^*\rangle| & \end{aligned}$$

By assumption, $|\langle f, \mu^* \rangle - \langle f, \widehat{\mu}_j^* \rangle| \longrightarrow_j 0$. Regarding the second term of this inequuality,

$$|\langle f, \widehat{T}_j^* \widehat{\mu}_j^* \rangle - \langle f, T^* \widehat{\mu}_j^* \rangle| = |\langle \widehat{T}_j f, \widehat{\mu}_j^* \rangle - \langle T f, \widehat{\mu}_j^* \rangle|$$

and $|\langle \widehat{T}_j f, \widehat{\mu}_j^* \rangle - \langle T f, \widehat{\mu}_j^* \rangle| \longrightarrow_j 0$, since $\widehat{T}_j f \rightarrow_j T f$ uniformly. Finally, regarding the third term of the above inequality,

$$|\langle f, T^* \widehat{\mu}_j^* \rangle - \langle f, T^* \mu^* \rangle| = |\langle T^* f, \widehat{\mu}_j^* \rangle - \langle T^* f, \mu^* \rangle|.$$

By assumption, $|\langle T^* f, \widehat{\mu}_j^* \rangle - \langle T^* f, \widehat{\mu}^* \rangle| \longrightarrow_j 0$. Therefore, $|\langle f, \mu^* \rangle - \langle f, T^* \mu^* \rangle| = 0$. As this latter equality holds for every f in $C(S)$, we get that $\mu^* = T^* \mu^*$. The result is thus established.

Proof of Corollary 3.5: It suffices to show that there is no sequence $\widehat{\psi}_j \rightarrow_j \psi$ such that every $\widehat{\psi}_j$ contains two distinct measures $\widehat{\mu}_j^*$ and $\widehat{v}_j^*$. By Theorem 3.2, these measures can be chosen with non-overlapping supports, $\sigma(\widehat{\mu}_j^*) \bigcap \sigma(\widehat{v}_j^*) = \phi$ for every j. Furthermore, by Proposition 3.4 we may assume that each of these sequences $\{\widehat{\mu}_j^*\}$ and $\{\widehat{v}_j^*\}$ converges weakly to the invariant probability measure μ^*, where $\mu^* = T^* \mu^*$. Hence, the corresponding sequences of sets $\{\sigma(\widehat{\mu}_j^*)\}$ and $\{\sigma(\widehat{v}_j^*)\}$ must both converge to $\sigma(\mu^*)$ in the topology induced by the Hausdorff distance [cf. Hildenbrand (1974, p.17)]. But by Condition U(ii) the set $\sigma(\mu^*)$ has non-empty interior, and hence $\sigma(\widehat{\mu}_j^*) \bigcap \sigma(\widehat{v}_j^*) \neq \phi$ for j large enough. We therefore arrive to a contradiction, since measures $\widehat{\mu}_j^*$ and $\widehat{v}_j^*$ were assumed to have non-overlapping supports. This contradiction proves the result.

3.3 Error Bounds

The preceding convergence results are usually not adequate for computational work. Computations must stop in finite time; hence, one generally needs fairly specific guidelines on the admissable size of the error. Under the following simple contradictive property on ψ, we establish that the approximation error of the statistic from a numerical simulation is of the same order of magnitude as the error of the approximate function $\widehat{\psi}$.

CONDITION C: There is γ with $0 < \gamma < 1$ such that $\|\psi(s, \varepsilon) - \psi(s', \varepsilon)\| \leq \gamma \|s - s'\|$ for all ε in E and all pairs of vectors s and s' in S.

As discussed below, Condition C seems to be indispensable for some of our main results. Of course, certain familiar extensions can be accommodated without substantial changes. All that is needed in our proofs is that the mapping ψ be an N-contraction on s over the ergodic set. Condition C holds in familiar stochastic growth models [e.g. Kydland and Prescott (1982)] in which the i.i.d. component has a small variance, and eventually all sample paths fluctuate around the corresponding deterministic steady-state solution.

THEOREM 3.10 *Assume that ψ satisfies condition C. Then, there exists a unique μ^* such that $\mu^* = T^*\mu^*$.*

Proof: By Theorem 3.1 there exists a solution μ^* such that $\mu^* = T^*\mu^*$. We next show that μ^* is unique.

Accordingly, assume that T^* contains two fixed-point solutions, μ^* and υ^*. Let $\sigma(\mu^*)$ and $\sigma(\upsilon^*)$ be the corresponding supports of these invariant probability measures. Then, by Condition C the distance between the two sets $\sigma(\mu^*)$ and $\sigma(\upsilon^*)$ must converge to zero. Therefore, $\sigma(\mu^*) = \sigma(\upsilon^*)$. That is, all invariant distributions must have a common support.

Finally, to establish that there is a unique invariant distribution, we shall show that the following uniqueness criterion [cf. Futia (1982, p. 385)] is satisfied: There is a point $\widetilde{s}$ in S such that for every neighborhood U of $\widetilde{s}$, and any initial point s in S, one can find an integer n such that $P^n(s, U) > 0$.

Pick any point $\widetilde{s}$ in the support $\sigma(\mu^*)$. Let s be an arbitrary point in S, and let δ_s be the Dirac measure. Then, a subsequence of $\{A_N(T^*)\delta_s\}_{N\geq 1}$ must converge weakly to an invariant distribution, μ^*. Hence, there is some $\widehat{n}$ such that $A_{\widehat{n}}(T^*)\delta_s(U) > 0$; for, if not, $\widetilde{s}$ does not belong to the common support of all the invariant distributions. But $A_n(T^*)\delta_s(U) > 0$ implies that $P^n(s, U) > 0$ for some n. Therefore, there exists a unique μ^* such that $\mu^* = T^*\mu^*$.

As above, we now pick an initial condition s_0 in S, and let $\omega = \{\varepsilon_n\}_{n\geq 0}$ denote a sequence of stochastic pertubations. Let $s(\omega) = \{s_n(\omega)\}_{\geq 0}$ be the corresponding ψ-sample path; let $\widehat{s}(\omega) = \{\widehat{s}_n(\omega)\}_{n\geq 1}$ be the $\widehat{\psi}$-sample path.

PROPOSITION 3.11 *Assume that ψ satisfies Condition C. Assume that $\left\|\widehat{\psi} - \psi\right\| \leq \delta$ for $\delta > 0$. Then, for any initial condition s_0, and any sequence $\omega = \{\varepsilon_n\}_{n\geq 0}$,*

$$\|\widehat{s}_n(\omega) - s_n(\omega)\| \leq \frac{\delta}{1-\gamma} \qquad \textit{for all } n \geq 1.$$

Proof: By assumption,

$$\|\widehat{s}_1(\omega) - s_1(\omega)\| = \left\|\widehat{\psi}(s_0, \varepsilon_0) - \psi(s_0, \varepsilon_0)\right\| \leq \delta.$$

Now, the proof procedes by a recursive argument. First, note that

$$\begin{aligned} \|\widehat{s}_2(\omega) - s_2(\omega)\| = \left\|\widehat{\psi}(\widehat{s}_1(\omega), \varepsilon_1) - \psi(s_1(\omega), \varepsilon_1)\right\| \leq \\ \left\|\widehat{\psi}(\widehat{s}_1(\omega), \varepsilon_1) - \psi(\widehat{s}_1(\omega), \varepsilon_1)\right\| + \|\psi(\widehat{s}_1(\omega), \varepsilon_1) - \psi(s_1(\omega), \varepsilon_1)\| \leq \\ \delta + \gamma\delta, \end{aligned}$$

where the first inequality comes from the triangle inequality of the norm, and the second inequality is true by Condition C and our induction argument.

Then, for any other $n > 2$,

$$\|\widehat{s}_n(\omega) - s_n(\omega)\| = \left\|\widehat{\psi}(\widehat{s}_{n-1}(\omega), \varepsilon_{n-1}) - \psi(s_{n-1}(\omega), \varepsilon_{n-1})\right\| \leq$$

$$\left\|\widehat{\psi}(\widehat{s}_{n-1}(\omega), \varepsilon_{n-1}) - \psi(\widehat{s}_{n-1}(\omega), \varepsilon_{n-1})\right\| + \|\psi(\widehat{s}_{n-1}(\omega), \varepsilon_{n-1}) - \psi(s_{n-1}(\omega), \varepsilon_{n-1})\| \leq$$

$$\delta + \gamma(\sum_{i=0}^{n-2} \gamma^i \delta) = \sum_{i=0}^{n-1} \gamma^i \delta,$$

where again the first inequality comes from the triangle inequality, and the second inequality is true by Condition C and our induction argument. Therefore,

$$\|\widehat{s}_n(\omega) - s_n(\omega)\| \leq \frac{\delta}{1-\gamma} \qquad \text{for all } n \geq 1.$$

We say that f in $C(S)$ is a Lipschitz function with constant L if $|f(s) - f(s')| \leq L\|s - s'\|$ for all s, s' in S. Note that in the next convergence result, we are not assuming that $\widehat{\psi}$ satisfies Condition D. Hence, the sequence $\{\frac{1}{N}\sum_{n\geq 1}^{N} f(\widehat{s}_n(\omega))\}$ may not have a unique limit point.

Theorem 3.12 *Let f be a Lipschitz function with constant L. Assume that ψ satisfies Condition C. Assume that $\left\|\widehat{\psi} - \psi\right\| \leq \delta$, for $\delta > 0$. Then, for a fixed s_0, every limit point $\widehat{E}(f)$ of the sequence $\{\frac{1}{N}\sum_{n\geq 1}^{N} f(\widehat{s}_n(\omega))\}_{N\geq 1}$ has the property*

$$|\widehat{E}(f) - \int f(s)\mu(ds)| \leq \frac{L}{1-\gamma}\delta, \qquad \mu^*(s_0, \cdot) - a.\ s.$$

Proof: By Theorem 3.3,

$$|\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(s_n(\omega)) - \int f(s)\mu^*(ds)| = 0, \qquad \mu^*(s_0, \cdot) - \text{a. s.}$$

Now, for present purposes there is no restriction of generality to assume that $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(\widehat{s}_n(\omega))$

is well defined. Then,

$$
\begin{aligned}
&|\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} f(\widehat{s}_n(\omega)) - \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} f(s_n(\omega))| = \\
&|\lim_{N\to\infty}\frac{\sum_{n=1}^{\infty}(f(\widehat{s}_n(\omega)) - f(s_n(\omega)))}{N}| \leq \\
&|\lim_{N\to\infty}\frac{\sum_{n=1}^{N}|f(\widehat{s}_n(\omega)) - f(s_n(\omega))|}{N}| \leq \\
&|\lim_{N\to\infty}\frac{\sum_{n=1}^{N} L\,\|\widehat{s}_n(\omega) - s_n(\omega)\|}{N}| \leq \\
&L\delta|\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\sum_{i=0}^{n-1}\gamma^i|
\end{aligned}
$$

Observe that the first weak inequality comes from the triangle inequality. The other two inequalities follow from the Lipschitz property of f and from the upper bounds derived in Proposition 3.11.

Now,

$$
\begin{aligned}
L\delta|\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\sum_{i=0}^{n-1}\gamma^i| &= L\delta|\lim_{N\to\infty}\sum_{n=1}^{N}\frac{1}{N}\frac{1-\gamma^{n+1}}{1-\gamma}| \\
&= \frac{L\delta}{1-\gamma}|1 - \lim_{N\to\infty}\frac{1}{(1-\gamma)N}| = \frac{L\delta}{1-\gamma}.
\end{aligned}
$$

The proof is complete.

Corollary 3.13 *Under the conditions of Theorem 3.12, every limit point $\widehat{E}(\widehat{s})$ of the sequence $\{\frac{1}{N}\sum_{n=1}^{N}\widehat{s}_n(\omega)\}_{N\geq 0}$ has the property*

$$
|\widehat{E}(\widehat{s}) - E(s)| \leq \frac{\delta}{1-\gamma}, \qquad \mu^*(s_0,\cdot) - a.\ s.
$$

Corollary 3.14 *Under the Conditions of Theorem 3.12, there exists a constant $\kappa > 0$ such that every limit point $\widehat{Var}(\widehat{s}^i)$ of the sequence $\{\frac{1}{N}\sum_{n=1}^{N}[\widehat{s}_n^i(\omega) - \widehat{E}(\widehat{s}^i)]^2\}$ has the property*

$$
|\widehat{Var}(\widehat{s}^i) - Var(s^i)| \leq \kappa\delta \quad \mu^*(s_0,\cdot) - a.\ s.
$$

COROLLARY 3.15 *Under the Conditions of Theorem 3.12, there exists a constant* $\zeta > 0$ *such that every limit point* $\widehat{Cov}(\widehat{s}^i, \widehat{s}^r)$ *of the sequence* $\{\frac{1}{N}\sum_{n=1}^{N}[\widehat{s}_n^i(\omega) - \widehat{E}(\widehat{s}^i)][\widehat{s}_n^r(\omega) - \widehat{E}(\widehat{s}^r)]\}$ *has the property*

$$|\widehat{Cov}(\widehat{s}^i, \widehat{s}^r) - Cov(s^i, s^r)| \leq \zeta\delta \quad \mu^*(s_0, \cdot) - a.\ s.$$

In the Appendix below, we provide upper bound estimates for constants κ and ζ.

4 Further Discussion of the Assumptions

(1) *Uniqueness of the Invariant Distribution*: Proposition 3.4 proves the upper semicontinuity of the set of invariant distributions. Lower semicontinuity holds if there is a unique invariant distribution. Example 2.1 illustrates that lower semicontinuity may fail. Some important economic models contain a unique invariant distribution.

(2) *Doeblin's Condition*: Condition D can be derived from primitive assumptions. Following Futia (1982, Th. 4.9 and Prop. 5.7) we can show that Condition D is satisfied for some basic stochastic processes.

(3) *Error Bounds*: The contraction property is indispensable. Examples can be constructed to illustrate that our error bounds are fairly tight.

References

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.

Christiano, L.J. and M. Eichenbaum (1992), "Current real business-cycle theories and aggregate labor-market fluctuations," *American Economic Review*, 82: 430-450.

Cooley, T.F. and E.C. Prescott (1995), "Economic growth and business cycles," in Cooley, ed., *Frontiers of Business Cycles Research*, Princeton University Press, Princeton, NJ.

Futia, C. (1992), "Invariant distributions and the limiting behavior of Markovian economic models," *Econometrica*, 50: 377-408.

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

Kydland, F. and E.C. Prescott (1982), "Time to build and aggregate fluctuations," *Econometrica*, 50:1345-1370.

Lucas, R.E. and E.C. Prescott (1971), "Investment under uncertainty," *Econometrica*, 39:659-681.

Manuelli, R. (1985), "A note on the behavior of solutions to dynamic stochastic models," mimeo, University of Wisconsin, Madison.

Santos, M.S. (1999), "Numerical solutions of dynamic economic models," in: J.B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics*, vol. 1: 312-386, Elsevier, Amsterdam.

Santos, M.S. (2000), "Accuracy of numerical solutions using the Euler equation residuals," *Econometrica*, 68: 1377-1402.

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

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.