This paper studies the asymptotic convergence of computed dynamic models when the shock is unbounded. Most dynamic economic models lack a closed-form solution. As such, approximate solutions by numerical methods are utilized. Since the researcher cannot directly evaluate the exact policy function and the associated exact likelihood, it is imperative that the approximate likelihood asymptotically converges— as well as to know the conditions of convergence— to the exact likelihood, in order to justify and validate its usage. In this regard, Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006) show convergence of the likelihood, when the shock has compact support. However, compact support implies that the shock is bounded, which is not an assumption met in most dynamic economic models, e.g., with normally distributed shocks. This paper provides theoretical justification for most dynamic models used in the literature by showing the conditions for convergence of the approximate invariant measure obtained from numerical simulations to the exact invariant measure, thus providing the conditions for convergence of the likelihood.

KEYWORDS: Dynamic economic models, Convergence, Computation.

1. INTRODUCTION

THIS PAPER STUDIES the convergence of dynamic economic models. While dynamic economic models have become a central tool for research and policy, most do not have a closed-form solution. Due to this, the policy function of these economic models are approximated by numerical methods. This approximation means that the researcher can only evaluate the approximated transition function associated with the approximated invariant measure, rather than the exact invariant measure implied by the exact transition function. Given that the researcher cannot evaluate the exact measure, it is natural to ask whether the approximate measure converges to the exact measure, at least asymptotically. If, for example, it does not converge, or the conditions for convergence are not met in practice, then the validity of the estimated economic model and its output comes into question. It is, therefore, critical that there is a theoretical foundation that provides the conditions for convergence to justify the usage of these dynamic economic models.

As a response, much econometric analysis has been done to provide this theoretical foundation. For example, Santos (2004) and Santos and Peralta-Alva (2005) provide the foundations of simulations of approximate solutions for stochastic dynamic models by studying its accuracy properties, showing that the computed moments from the numerically approximated policy converge to the exact moments as the approximation errors of the computed solutions go to zero. Further, and more relevant to this paper, Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006) extends the results of Santos and Peralta-Alva (2005) to the convergence of the likelihood of computed economic models, providing conditions for which the approximated likelihood functions converges to the exact likelihood. While the convergence results in
Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006) provide some justification for dynamic economic models, one assumption it employs to obtain their result is rarely met. This assumption is the compactness of the state space, which implies that the support of the shock of a dynamical system is bounded. Although this assumption is standard in the numerical literature, it excludes—among others—dynamical models with normally distributed shocks. As assuming a normally distributed shock is standard in empirical studies, in which the evaluation of the likelihood is done by the usage of the Kalman filter (Smets and Wouters, 2007), it is simply vital that the results in Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006) extend to non-compact support, i.e., unbounded shock. For example, recent works by Stachurski (2002), Nishimura and Stachurski (2005), Kamihigashi (2007) and Kamihigashi and Stachurski (2016) study the asymptotic invariant measure of the stochastic neoclassical growth model without compactness of the shocks and states. The purpose of this paper is to relax the compactness assumption for the convergence of the approximated invariant measure, providing the theoretical foundation and justification for these models.

The rest of this paper is organized as follows. Section 2 gives the set-up of dynamic economic models and preliminary of the Markov operator. Section 3 presents our result on the convergence of the invariant measure. In Section 4, we derive error bounds for these approximations. Section 5 presents our main result on the convergence of computed likelihoods.

2. MODEL SET-UP AND PRELIMINARIES

We follow the set of notations and models in Santos and Peralta-Alva (2005). The equilibrium law of motion of the state variables is specified by a dynamical system of the form

\[ s_{n+1} = \varphi (s_n, \varepsilon_{n+1}), \quad n = 0, 1, 2, \ldots, \]

where \( s_n \) is a vector of state variables that characterize the evolution of the system. The vector \( s_n \) belongs to a measurable state space \( (S, \mathcal{S}) \). We endow \( S \) with its relative Borel \( \sigma \)-algebra, which we denote by \( S \). The variable \( \varepsilon \) is an independent and identically distributed shock, which is defined on the sample space \( (E, \mathcal{E}) \). The distribution of the shock \( \varepsilon \) is given by a stochastic kernel \( Q : S \times E \rightarrow [0, 1] \), where \( Q(s, A) \) is the probability of realizing the event \( A \in \mathcal{E} \), given that the current state is \( s \in S \).

Given a random dynamical system, one can define a transition probability on the state space in the following way. Define the transition probability function as

\[ P(s, A) = Q(\{ \varepsilon \mid \varphi(s, \varepsilon) \in A \}). \]

The transition function, \( P : S \times S \rightarrow [0, 1] \), is defined by

\[ P(s, A) = Q(s, \varphi^{-1}(A)). \]

Let \( B(S) \) be the set of all bounded \( S \)-measurable real valued functions on \( S \), with sup norm \( |f| = \sup_{s} |f(s)| \). The Markov operator associated with \( P \) is defined as

\[ T f(s) \triangleq \int f(t) P(s, dt) \]

\[ = \int f(\varphi(s, \varepsilon)) Q(s, d\varepsilon) \quad (1) \]

\[ = \int f(s) Q(s, d\varepsilon) \]

\[ = \int f(s) (Q(s, \varepsilon)) \]

\[ = \int f(s) \]
For any given initial condition $\mu_0$ on $S$, the evolution of future probabilities, $\{\mu_n\}$, can be specified by the following operator $T^*$ that takes the space

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

for all $A$ in $S$ and $n \geq 0$. The adjoint $T^*$ of $T$ is defined by the formula

$$(T^* \mu) (A) = \int P(t, A) \mu (dt).$$

We maintain the following basic assumptions.

**Assumption 2.1.** The space of sets $S$ and $E$ are both locally compact and $\sigma$-compact.

Locally compact means that for each point $x \in S$, there is some compact subspace $C$ of $S$ that contains a neighborhood of $x \in S$. Further, $\sigma$-compact is a countable union of compact spaces. Note that the space $\mathbb{R}^d$ is both locally compact and $\sigma$-compact. A space that is both locally compact and $\sigma$-compact can be written as an increasing union of countably many open sets, each of which is compact and closed. In Santos and Peralta-Alva (2005), they impose the compactness assumption on both states, $S$ and $E$, which, again, is not an assumption met in most dynamic economic models used in empirical studies. In an important distinction, we relax this restriction to the non-compact case, which allows us to use the whole Euclidean state, $S = \mathbb{R}^d$, and unbounded distributions, such as the normal distribution.

Recall that the probability measure $P$ is called tight if for all $\epsilon > 0$ there is a compact set $K \subset S$ such that $P(K) \geq 1 - \epsilon$. Any probability measure on the complete separable metric space is tight.

**Assumption 2.2.** The Markov operator $T^*$ has a unique fixed point $\mu_0$: $T^* \mu_0 = \mu_0$.

A sufficient condition for Assumption 2.2 is that there exists a point, $s_0 \in S$, such that, for any point $s \in S$, any neighborhood $U$ of $s_0$ and any integer $k \geq 1$, we have $P^{nk}(s, U) > 0$ (see Futia, 1982, Section 3.2).

**Assumption 2.3.** Function $\phi: S \times E \rightarrow S$ is jointly measurable. Moreover, for every continuous function $f: S \rightarrow \mathbb{R}$,

$$\int f(\phi(s, \epsilon)) Q(d\epsilon) \rightarrow \int f(\phi(s, \epsilon)) Q(d\epsilon) \text{ as } s_j \rightarrow s.$$

Assumption 2.3 is the same as Assumption 2 in Santos and Peralta-Alva (2005).

In most cases, the researcher does not know the exact form of the transition equation $\varphi$, and only has access to the numerical approximation of the transition equation, $\varphi_j$, with index $j$. The index $j$ indicates the approximation and implies that, as $j$ goes to infinity, the approximation, $\varphi_j$, converges to the exact value (the metric of convergence is defined later). Every numerical approximation $\varphi_j$ defines the transition probability $P_j$ on $(S, S)$. Given an approximation $\varphi_j$, we define the corresponding approximation of the transition probability as

$$P_j(s, A) = Q(\{\epsilon \mid \varphi_j(s, \epsilon) \in A\}),$$

and define the approximated transition function, $P_j: S \times S \rightarrow [0, 1]$, as

$$P_j(s, A) = Q(\{s, \varphi_j^{-1}(A)\}).$$

The Markov operator associated with $P_j$ is defined as

$$T_j f(s) \triangleq \int f(t) P_j(s, dt) \quad (2)$$
\begin{align*}
&= \int f(\varphi_j(s, \varepsilon)) Q(s, d\varepsilon).
\end{align*}

The evolution of future probabilities, \( \{\mu^j_n\} \), can be specified by the following operator \( T^*_j \) that takes the space

\[
\mu^j_{n+1} = (T^*_j \mu^j_n)(A) = \int P_j(s, A) \mu^j_n(ds),
\]

for all \( A \) in \( S \) and \( n \geq 0 \). The adjoint \( T^*_j \) of \( T_j \) is defined by

\[
(T^*_j \mu)(A) = \int P_j(t, A) \mu(dt).
\]

Every numerical approximation \( \varphi_j \) satisfies a structure parallel to that of the above Assumptions 2.2 and 2.3. We further assume:

**Assumption 2.4.** The Markov operator \( T^*_j \) has a unique fixed point \( \mu^j_0 \):

\[
T^*_j \mu^j_0 = \mu^j_0, \quad \text{for all} \quad j.
\]

**Assumption 2.5.** For each \( j \), the function \( \varphi_j : S \times E \to S \) is jointly measurable. Moreover, for every continuous function \( f : S \to \mathbb{R} \),

\[
\hat{f}(\varphi_j(s_j, \varepsilon)) Q(d\varepsilon) \to \hat{f}(\varphi(s, \varepsilon)) Q(d\varepsilon) \quad \text{as} \quad s_j \to s.
\]

### 3. CONVERGENCE OF THE INVARIANT DISTRIBUTION

Now, recall the convergence of probability measures on \( S \). When the state space \( S \) is separable, we can introduce a metric \( D \) in the space of probability measures on \( S \), such that

\[
\lim_{n} D(\mu_n, \mu) = 0 \quad \text{if and only if} \quad \mu_n \text{ converges in law to} \quad \mu.
\]

Specifically, the metric we use is the Fortet-Mourier metric (Dudley, 2002, Section 11.3):

\[
D(\mu_n, \mu) = \sup_{f \in BL(S)} \left| \int_{S} f(s) d\mu_n - \int_{S} f(s) d\mu \right|,
\]

where the supremum \( \sup_{f \in BL(S)} \) is taken over all bounded Lipschitz continuous functions defined on \( S : BL(S) \).

The main question we answer in this paper is the following: How strong of a topology is sufficient for the approximate transition equation, \( \varphi_j \), to converge to the true transition equation, \( \varphi \), in order for the approximate invariant measure, \( \mu_n \), to converge to the convergence in law distance eq. (3). Santos and Peralta-Alva (2005), assuming that the state-space is compact, showed that convergence under the following topology is sufficient to prove the convergence of the invariant measure. Endow the metric in the space of functions \( \varphi \) and \( \hat{\varphi} \) as

\[
\max_{s \in S} \left[ \int \| \varphi(s, \varepsilon) - \hat{\varphi}(s, \varepsilon) \| Q(d\varepsilon) \right]
\]

where \( \| \cdot \| \) is the max norm in \( \mathbb{R}^d \). This metric only works under the compactness assumption on \( S \). To consider the functional approximation of the transition equation, \( \varphi \), under non-compactness, this uniform topology is not practical. In the following, we extend the state-space, \( S \), to non-compactness and weaken the uniform convergence topology of the functional approximation to a local uniform topology.

First, note that \( BL(S) \) can be relaxed to infinitely continuously differentiable functions on \( S : C^\infty(S) \) by using the mollifier method. Then, we have the following lemma:
Lemma 3.1. The limit, \( \lim_{n \to \infty} D(\mu_n, \mu) = 0 \), holds if and only if

\[
\lim_{n \to \infty} \sup_{f \in C_c(S)} \left| \int_S f(s) d\mu_n - \int_S f(s) d\mu \right| = 0.
\]

Note that by Assumptions 2.2 and 2.3, each \( \varphi_n \) defines the associated pair \( (P_j, T_j) \). The adjoint \( T_j^* \) of \( T_j \) is

\[
\langle T_j f, \mu_j \rangle = \int \int f(\varphi_j(s, \varepsilon)) Q(s, d\varepsilon) d\mu_j(s)
\]

\[
\langle f, T_j^* \mu_j \rangle = \int \int f(\varphi_j(s, \varepsilon)) Q(s, d\varepsilon) d\mu_j(s).
\]

Moreover, there always exists an invariant distribution \( \mu_j^* = T_j^* \mu_j^* \).

Given Lemma 3.1, we have the following result:

**Proposition 3.1.** Suppose Assumptions 2.1, 2.2, and 2.3 are satisfied for each approximated model \( \varphi_i \) and \( \varphi_0 \). Then, a sufficient condition for a sequence of the measure \( \mu_j \), associated with \( T_j \), to converge to \( \mu_0 \), associated with \( T_0 \), in the sense of eq. (3), is the strong convergence of \( T_j \) to \( T_0 \).

Since \( S \) is a completely regular space, it has Stone-Čech compactification \( \beta(S) \):

**Theorem 3.1.** (Stone-Čech compactification: Munkres, 2000, Theorem 38.2). Let \( S \) be a completely regular space. Then, there exists a compactification \( \beta(S) \) of \( S \) having the property that every bounded continuous function \( f: S \to \mathbb{R} \) extends uniquely to a continuous function of \( \beta(S) \) into \( \mathbb{R} \).

The Stone-Čech compactification, \( \beta(S) \), of which \( S \) is a dense subspace, satisfies the property that each bounded continuous function, \( f: S \to \mathbb{R} \), has a continuous extension, \( g: \beta(S) \to \mathbb{R} \). We endow the metric in the space of functions defined on the locally compact and \( \sigma \)-compact space, \( S \). For any two vector-value functions \( \varphi \) and \( \hat{\varphi} \), let \( d(\cdot, \cdot) \) be

\[
d(\varphi, \hat{\varphi}) = \max_{f \in C_c(\beta(S))} \max_{s \in \beta(S)} \left[ \int |f(\varphi(s, \varepsilon)) - f(\hat{\varphi}(s, \varepsilon))| Q(d\varepsilon) \right]. \tag{4}
\]

The metric in eq. (4) is weaker than the metric of Santos and Peralta-Alva (2005), and extends to the non-compact state space. In this section, convergence of the sequence of functions \( \{\varphi_j\} \) is in this distance, as this metric can accommodate the noncontinuous functions \( \varphi \) and \( \hat{\varphi} \). Although we will impose continuous differentiability on \( \varphi \) for the convergence of the approximate likelihood studied in Section 5, the metric \( d(\cdot, \cdot) \) is sufficient to guarantee the convergence of the invariant distribution.

Then, we have the following theorem:

**Theorem 3.2.** Let \( \{\varphi_j\} \) be a sequence of functions that converge to \( \varphi \), in the sense of \( d(\cdot, \cdot) \) in eq. (4). Let \( \{\mu_j^*\} \) be a sequence of probabilities on \( S \) associated with \( \{\varphi_j\} \), such that \( \mu_j^* = T_j^* \mu_j^* \), for each \( j \). Under Assumptions 2.1 and 2.2, if \( \mu^* \) is a weak limit point of \( \{\mu_j^*\} \), then \( \mu^* = T^* \mu^* \).

This theorem asserts the bilinear convergence of \( T_j^* \mu_j^* \) to \( T^* \mu^* \) in the weak topology.

4. ERROR BOUNDS

In this section, we study the error bounds of these approximations under non-compactness. The error bounds are important for two reasons. First, in numerical applications, it is often desirable to bound the size of the approximation error in order to know the theoretical limit of the
approximation. Second, computations cannot go on forever and must stop in finite time. Hence, knowing the error bounds can dictate an efficient stopping criteria to minimize computational cost while ensuring convergence. As such, Santos and Peralta-Alva (2005) give a bound on the size of the approximation error under the compactness assumption.

To begin, we introduce the notion of compactness for the Markov operator. The Markov operator $T$ is compact if the image $T(bX)$ has compact closure in $X$, where $bX = \{x \in X \mid \|x\| \leq 1\}$. The Markov operator $T$ is quasi-compact if there is a unique compact operator $L$ and an integer $n$ such that

$$\sup_{x \in bX} \|T^n x - Lx\| < 1.$$ 

If the above quasi-compactness is satisfied, one can obtain the convergence of the sequence of operators, $\{T^n\}$, to the invariant probability at a geometric rate. The following theorem gives this result.

**Theorem 4.1.** (Yosida and Kakutani, 1941). Let $T$ be a stable, quasi-compact Markov operator defined by eq. (1) satisfying Assumptions 2.1, 2.2, and 2.3. Then, there exist constants, $C$, with $\varepsilon > 0$, such that

$$\sup_{s \in \beta(S)} \|T^n f(s) - T^* f(s)\| \leq \frac{C}{(1 + \varepsilon)^n}.$$ 

The following theorem bounds the approximation error between the expected values of $f$ over the true invariant measure $\mu^*$ and the approximate invariant measure $\hat{\mu}^*$ of $\hat{\varphi}$.

**Proposition 4.1.** Let $f$ be a Lipschitz function with constant $L$. Suppose we have a numerical approximation $\hat{\varphi}$ with $d(\hat{\varphi}, \varphi) \leq \delta$, for some $\delta > 0$. Then,

$$\left| \int f(s) \mu^* (ds) - \int f(s) \hat{\mu}^* (ds) \right| \leq \frac{Ld(\hat{\varphi}, \varphi)}{\varepsilon},$$

where $\mu^*$ is the unique invariant measure of the exact $\varphi$, and $\hat{\mu}^*$ is a unique invariant measure of $\hat{\varphi}$.

Quasi-compact operators enjoy a very useful property in Theorem 4.1. Furthermore, quasi-compact operators are easily recognizable. In fact, we find that most operators are quasi-compact (see Futia, 1982). In our circumstance, Assumption 2.2 guarantees the quasi-compactness of the Markov operator.

5. APPLICATION: CONVERGENCE OF COMPUTED LIKELIHOOD

Given the convergence of the invariant measure in the previous section, we prove the convergence of the approximate likelihood in Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006). We relax the compactness assumption in the state-space and shock, and prove equivalent results to Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006), justifying the construction of the likelihood via the Kalman filter, among others.

5.1. Likelihood Induced by Random Dynamical Systems

The equilibrium law of motion of the state space system can be specified as

$$s_t = \varphi(s_{t-1}, \varepsilon_t; \theta),$$

where

$$s_t = \varphi(s_{t-1}, \varepsilon_t; \theta),$$

and

$$s_t = \varphi(s_{t-1}, \varepsilon_t; \theta),$$

are the state variables at time $t$. The state vector $s_t$ is a function of the previous state $s_{t-1}$, a shock $\varepsilon_t$, and a parameter $\theta$. The function $\varphi$ is a nonlinear mapping that describes the evolution of the state variables over time.
where eq. (5) is the transition equation, and eq. (6) is the measurement equation. Here, the variables $\varepsilon_t$ and $\eta_t$ are tight random elements and are independent and identically distributed shocks with values in some Euclidean space, with bounded and continuous densities. Their distribution is given by the probability measure, $Q$, defined on a measurable space, $(E, \mathcal{E}) \subset (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. We do not impose the compactness on the support of $Q$ but impose tightness, in order to deal with unbounded shocks, such as normally distributed shocks. The parameter, $\theta \in \Theta \subset \mathbb{R}^n$, is a vector of structural parameters and $\Theta$ is on a compact set. The vector, $y_t$, is the observables in each period, $t$. Let $Y_T = \{y_t\}_{t=1}^T$ with $Y^0 = \{\emptyset\}$. To avoid singularity, we impose $\dim(\varepsilon_t) + \dim(\eta_t) \geq \dim(Y_t)$. And we partition $\{\varepsilon_t\}$ into $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})$, such that $\dim(\varepsilon_{2,t}) + \dim(\eta_t) = \dim(y_t)$. As in the previous section, we index the approximations by $j$, the numerical approximation to the transition equations is $\varphi_j$, and the measurement equations is $g_j$.

As with the previous section, we assume that each state-space system has an invariant measure and that invariance measure is absolute continuous with regard to a Lebesgue measure:

**Assumption 5.1.** For all $\theta$ and all $j$, there exists a unique invariant distribution for $S$, $\mu(S; \theta)$, and $\mu_j(S; \theta)$, that has a Radon-Nikodym derivative with respect to the Lebesgue measure.

The exact likelihood is constructed, using the change of variables formula, as follows. First we assume that the system can solve the error term, exactly.

**Assumption 5.2.** For all $\theta$ and $t$, the system of equations eq. (5) and eq. (6) have a unique solution,

$$
\begin{align*}
\eta_t &= \eta^t(\varepsilon^t, s_0, y^t; \theta), \\
\epsilon_{1,t} &= s^t(\varepsilon^t, s_0, y^t; \theta), \\
\epsilon_{2,t} &= \epsilon_{2}^t(\varepsilon^t, s_0, y^t; \theta),
\end{align*}
$$

and we can evaluate $p(v^t(W^t_1, S_0, y^t; \theta))$ and $p(w^t_2(W^t_1, S_0, y^t; \theta))$ for all $S_0$, $W^t_1$, and $t$.

We further assume that the observation equation (6) is continuously differentiable.

**Assumption 5.3.** For all $\theta$, function $g(\cdot, \cdot, \theta)$ is continuously differentiable, with bounded partial derivatives.

From Assumptions 5.3, 5.1, and 5.2, we construct the likelihood function by the change of variable formula:

$$
p(y_t \mid W^t_1, S_0, y^{t-1}; \theta) = p(v_t; \theta)p(w_{2,t}; \theta) \cdot dy(v_t, w_{2,t}; \theta),
$$

where

$$
|dy(v_t, w_{2,t}; \theta)| = \det \left[ \nabla \frac{\partial g}{\partial v_t} \nabla \frac{\partial g}{\partial w_{2,t}} \right].
$$

Further, we have the following assumption (Fernandez-Villaverde, Rubio-Ramirez, and Santos, 2006, Assumption 4).

**Assumption 5.4.** For all $\theta$ and $t$, the model gives some positive probability to the data $y^T$, that is, $p(y_t \mid W^t_1, S_0, y^{t-1}; \theta) > \xi \geq 0$ for all $S_0$ and $W^t_1$.

From Assumption 5.4, the likelihood is as follows,

$$
L(y^T; \gamma) = \prod_{t=1}^{T} p(y_t \mid y^{t-1}; \theta)
$$
\[
= \prod_{t=1}^{T} \int \int p(y_t \mid W_1^t, S_0, y_t^{t-1}; \theta) p(W_1^t, S_0 \mid y_t^{t-1}; \theta) dW_1^t dS_0
\]

\[
= \left( \int \left( \prod_{t=1}^{T} p(W_1^t; \theta) p(y_t \mid W_1^t, S_0, y_t^{t-1}; \theta) dW_1^t \right) \right) \mu^*(dS_0; \theta).
\]

Next, we also assume that, also for the approximate state-space functions, \( \{ \varphi_j \} \), and measurement function, \( \{ g_j \} \), the system can solve the error term, exactly.

**Assumption 5.5.** For all \( j \), the system of equations

\[
S_1 = \varphi_j(S_0, (W_{1,1}, W_{2,1}); \theta),
\]

\[
y_m = g_j(S_m, V_m; \theta) \quad \text{for} \quad m = 1, 2, \ldots, t,
\]

\[
S_m = \varphi_j(S_{m-1}, (W_{1,m}, W_{2,m}); \theta) \quad \text{for} \quad m = 2, 3, \ldots, t,
\]

has a unique solution,

\[
V_{j,t} = v_j^t(W_1^t, S_0, y^t; \theta),
\]

\[
S_{j,t} = s_j^t(W_1^t, S_0, y^t; \theta),
\]

\[
W_{j,2,t} = w_{j,2}^t(W_1^t, S_0, y^t; \theta),
\]

and we can evaluate \( p(v_j^t(W_1^t, S_0, y^t; \theta); \theta) \) and \( p(w_{j,2}^t(W_1^t, S_0, y^t; \theta); \theta) \) for all \( S_0, W_1^t, \) and \( t \).

We also assume that the measurement function, \( \{ g_j \} \), is continuously differentiable.

**Assumption 5.6.** For all \( j \), functions \( g_j(\cdot, \cdot, \theta) \) are continuously differentiable at all points except at a finite number of points. At the points of differentiability, all partial derivatives are bounded, and the bounds are independent of \( j \).

Then, \( dy_j(v_{j,t}, w_{j,2}; \theta) \) exists for all but a finite set of \( S_0 \) and \( W_1^t \), we have, for all \( j, \theta, \) and \( t \),

\[
p_j(y_t \mid W_1^t, S_0, y_t^{t-1}; \theta) = p(v_{j,t}; \theta) p(w_{j,2,t}; \theta) \mid dy_j(v_{j,t}, w_{j,2,t}; \theta) \mid,
\]

where

\[
\mid dy_j(v_{j,t}, w_{j,2,t}; \theta) \mid = \det \left[ \nabla_{v_{j,t}} \frac{\partial g_j}{\partial v_{j,t}} \nabla_{w_{j,2,t}} \frac{\partial g_j}{\partial w_{j,2,t}} \right],
\]

for all \( S_0 \) and \( W_1^t \), but a finite number of points.

As \( j \) goes to infinity, \( \varphi_j \) and \( g_j \) converge to their exact values. Unlike the previous section, convergence of the sequence of functions \( \{ \varphi_j \} \) and \( \{ g_j \} \) need to be a stronger topology, which is defined in the following way. For any two vector-valued functions \( \varphi \) and \( \tilde{\varphi} \), let

\[
d_{C^1}(\varphi, \tilde{\varphi}) = \max_{s \in I} \sup_{\varepsilon \in \mathbb{S}_I} \left[ \int \| \varphi(s, \varepsilon) - \tilde{\varphi}(s, \varepsilon) \| Q(\varepsilon) d\varepsilon + \int \| \nabla \varphi(s) - \nabla \tilde{\varphi}(s) \| Q(\varepsilon) d\varepsilon \right],
\]

where \( \{ S_i \}_{i \in I} \) is an exhaustive sequence of compact sets of \( S \). In this section, convergence of a sequence of functions, \( \{ \varphi_j, g_j \} \), should be understood in this norm.

Assumption 5.5 is required for the change of variable formula in eq. (7). Convergence in \( C^1 \) implies the convergence of the solutions, \( v_{j,t}^t \), \( s_{j,t}^t \), and \( w_{j,2,t}^t \), from the same argument in the proof of consistency of Z-estimators. More importantly, the convergence of the Jacobian, \( \mid dy_j(v_{j,t}, w_{j,2,t}; \theta) \mid \), can be derived from the Ascoli-Arzela theorem.
Theorem 5.1. (Ascoli-Arzela). Let $S$ be a separable $\sigma$-compact metric space and $S_i$, $i \in I$, be an exhausting sequence of compact sets. Further, let $C^1(S)$ be the Banach space of complex-valued continuous functions, $f(s)$, normed by

$$\| f \| = \max_{i \in I} \left[ \sup_{s \in S_i} |f(s)| + \sup_{s \in S_i} |f'(s)| \right].$$

If the following two conditions are satisfied:

$$\sup_{n \geq 1} \sup_{s \in S} |f_n(s)| < \infty, \quad (\text{equi-bounded})$$

$$\lim_{\delta \downarrow 0} \sup_{n \geq 1, \delta(s', s'') \leq \delta} |f_n(s') - f_n(s'')| = 0, \quad (\text{equi-continuous})$$

then the sequence, $\{f_n(s)\} \subseteq C(S)$, is relatively compact in $C(S)$.

Given this, we have the following proposition that proves the convergence of the approximate likelihood to the exact likelihood with unbounded shock.

Proposition 5.1. Suppose Assumptions 5.3-5.5, and suppose $d_{C^1}(\varphi_j, \varphi) \to 0$ and $d_{C^1}(g_j, g) \to 0$ as $j \to \infty$. Then, it holds that

$$\prod_{t=1}^{T} p_j(y_t | y_{t-1}; \gamma) \xrightarrow{\text{loc}} \prod_{t=1}^{T} p(y_t | y_{t-1}; \gamma).$$

The result shows that, as the researcher gets better approximations of the policy function in a dynamic economic model, the computed likelihood converges to the exact likelihood, even if the shock is unbounded. This result goes beyond the result in Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006) and is particularly relevant to researchers using dynamic economic models with unbounded shocks, such as normally distributed shocks—a standard specification in the literature—as it guarantees, asymptotically, that the likelihood function implied by the model is the correct object of interest.

REFERENCES

DUDLEY, R. M. (2002): “Real analysis and probability, volume 74 of Cambridge Studies in Advanced Mathematics,”

FERNANDEZ-VILLaverDE, J., J. RUBIo-RAMirez, AND M. SANTos (2006): “Convergence Properties of the Likelihood of Computed Dynamic Models,” Econométrica, 74(1), 93–119.

FUTIA, C. (1982): “Invariant distributions and the limiting behavior of Markovian economic models,” Econometrica, 50(2), 377–408.

KAMihGASHI, T. (2007): “Stochastic optimal growth with bounded or unbounded utility and with bounded or unbounded shocks,” Journal of Mathematical Economics, 43(3), 477–500.

KAMihGASHI, T. AND J. STAChurstKI(2016): “Seeking ergodicity in dynamic economies,” Journal of Economic Theory, 163, 900–924.

MUNKRES, J. R. (2000): Topology, Prentice Hall.

NISHIMURA, K., AND J. STAChurstKI (2005): “Stability of Optimal growth models: a new approach,” Journal of Economic Theory, 122, 100–118.

SANTos, M., AND A. PERALTA-ALVA (2005): “Accuracy of simulations for stochastic dynamic models,” Econométrica, 73(6), 1939–1976.

SANTos, M. S. (2004): “Simulation-based estimation of dynamic models with continuous equilibrium solutions,” Journal of Mathematical Economics, 40(3-4), 465–491.

SMETS, F., AND R. WOUTERS (2007): “Shocks and frictions in US business cycles: A Bayesian DSGE approach,” American Economic Review, 97(3), 586–606.

STAChurstKI, J. (2002): “Stochastic Optimal growth with unbounded shock,” Journal of Economic Theory, 106, 40–65.

Yosida, K., AND S. KAKUTANI (1941): “Operator-theoretical treatment of Markoff’s process and mean ergodic theorem,” Annals of Mathematics, pp. 188–228.
APPENDIX: APPENDIX

PROOF OF LEMMA 3.1: The “Only if” part is derived from \( C^\infty \subset BL \). In this proof we will show the “if” part. For any \( f \in BL(S) \) and \( \varepsilon > 0 \), there exists \( u \in C^\infty(S) \), such that

\[
\sup_{s \in S} |f(s) - u(s)| < \varepsilon
\]

by the mollifier method. By triangular inequality, we have

\[
\sup_{f \in BL(S)} \left| \int_S f(s) d\mu_n - \int_S f(s) d\mu \right| \\
\leq \sup_{f \in BL(S)} \left| \int_S f(s) d\mu_n - \int_S u(s) d\mu_n \right| + \sup_{u \in C^\infty(S)} \left| \int_S u(s) d\mu_n - \int_S u(s) d\mu \right| \\
+ \sup_{f \in BL(S)} \left| \int_S u(s) d\mu - \int_S f(s) d\mu \right|
\]

The first and third term of the right hand side of the inequality are smaller than \( \varepsilon \). The second term is given in the definition. \( \Box \).

PROOF OF PROPOSITION 3.1: The strong convergence of the sequence of operators is

\[
\sup_{s \in S} \|Tf(s) - T_jf(s)\| = \sup_{s \in S} \left| \int f(\varphi(s,\varepsilon)) Q(s,d\varepsilon) - \int f(\varphi_j(s,\varepsilon)) Q(s,d\varepsilon) \right| \\
= \sup_{s \in S} \left| \int \{f(\varphi(s,\varepsilon)) - f(\varphi_j(s,\varepsilon))\} Q(s,d\varepsilon) \right| \\
\to 0
\]

for all \( f \in C^2(S) \). This simply means that

\[
\lim_{j \to \infty} \mathbb{E}[f(\varphi_j(s))] = \mathbb{E}[f(\varphi(s))], \quad \forall f \in C^2(S).
\]

Let \( f \) belong to \( A \). Then, for any two \( T \) and \( T_j \), and corresponding invariant measures \( \mu \) and \( \mu_j \), we have

\[
|\langle Tf, \hat{\mu} \rangle - \langle T_jf, \mu_j \rangle| \leq |\langle Tf, \mu \rangle - \langle T_jf, \mu_j \rangle| + |\langle Tf, \mu_j \rangle - \langle T_jf, \mu_j \rangle|.
\]  (8)

It follows from Proholov’s Theorem that \( \{\mu_j\} \) has a weakly convergent subsequence. Let \( \{\mu_{j_k}\} \) be such a subsequence, and let \( \hat{\mu} \) be its limit. Then for the first term in eq. (8), we have

\[
|\langle Tf, \hat{\mu} \rangle - \langle T_jf, \mu_{j_k} \rangle| \to 0.
\]

Next, we show the equality: \( \hat{\mu} = \mu_0 \). We have

\[
|\langle f, \hat{\mu} \rangle - \langle T_0f, \hat{\mu} \rangle| \leq |\langle f, \hat{\mu} \rangle - \langle f, \mu_{j_k} \rangle| + |\langle f, \mu_{j_k} \rangle - \langle T_0f, \mu_{j_k} \rangle| \\
+ |\langle T_0f, \mu_{j_k} \rangle - \langle T_0f, \hat{\mu} \rangle|.
\]

Since \( Tf \) and \( T_0f \) are continuous and \( \{\mu_{j_k}\} \) convergence weakly to \( \hat{\mu} \), the first and third terms on the right hand side approaches zero as \( j \) goes to infinity. For the second term, we have

\[
|\langle f, \mu_{j_k} \rangle - \langle T_0f, \mu_{j_k} \rangle| = |\langle f, T_0^* \mu_{j_k} \rangle - \langle T_0f, \mu_{j_k} \rangle| \\
= |\langle T_0 f, \mu_{j_k} \rangle - \langle T_0f, \mu_{j_k} \rangle| \\
\leq \|T_0 f - T_0f\|
\]
Proof of Theorem 3.2: The topology of weak convergence can be defined by the metric on the probability measure space,

\[ d(\mu, \nu) = \sup_{f \in \mathcal{A}} \left\{ \left| \int f(s) \mu(ds) - \int f(s) \nu(ds) \right| \right\}, \]

where \( \mathcal{A} \) is the space of Lipschitz functions on \( S \), with constant \( L \leq 1 \) and \(-1 \leq f \leq 1 \). Then, we have

\[ \|Tf(s) - T_j f(s)\| = \left| \int f(\varphi(s, \varepsilon)) Q(\varepsilon) - \int f(\varphi_j(s, \varepsilon)) Q(\varepsilon) \right| \]

\[ = \left| \int f(\varphi(s, \varepsilon)) - f(\varphi_j(s, \varepsilon)) Q(\varepsilon) \right|. \]

Since \( f \in C^\infty \), there exists a constant \( K \), such that

\[ \left| \int [f(\varphi(s, \varepsilon)) - f(\varphi_j(s, \varepsilon))] Q(\varepsilon) \right| \leq K d(\varphi, \varphi_j). \]

\[ Q.E.D. \]

Proof of Proposition 4.1: Denote \( s_n(s_0) \) as

\[ \varphi(\underbrace{\varphi(\cdots(\varphi(s_0, \varepsilon_1), \varepsilon_2))}_{n \text{ times}}). \]

And denote \( \hat{s}_n(s_0) \) as

\[ \hat{\varphi}(\underbrace{\hat{\varphi}(\cdots(\hat{\varphi}(s_0, \varepsilon_1), \varepsilon_2))}_{n \text{ times}}). \]

Since \( f \) is a Lipschitz function with constant \( L \), we have

\[ \left| \mathbb{E}[f(s_n(s_0))] - \mathbb{E}[f(s_n(s_0))] \right| \]

\[ = \left| \mathbb{E}[f(\varphi(s_{n-1}(s_0), \varepsilon_n))] - \mathbb{E}[f(\hat{\varphi}(s_{n-1}(s_0), \varepsilon_n))] \right| \]

\[ \leq \left| \mathbb{E}[f(\varphi(s_{n-1}(s_0), \varepsilon_n))] - \mathbb{E}[f(\varphi(s_{n-1}(s_0), \varepsilon_n))] \right| \]

\[ + \left| \mathbb{E}[f(\varphi(s_{n-1}(s_0), \varepsilon_n))] - \mathbb{E}[f(\hat{\varphi}(s_{n-1}(s_0), \varepsilon_n))] \right| \]

\[ \leq L(1 - \varepsilon) \mathbb{E}[s_{n-1}(s_0) - \hat{s}_{n-1}(s_0)] + Ld(\hat{\varphi}, \varphi). \]

By the same argument above, we have

\[ L(1 - \varepsilon) \mathbb{E}[s_{n-1}(s_0) - \hat{s}_{n-1}(s_0)] \]

\[ \leq L(1 - \varepsilon)^2 \mathbb{E}[s_{n-2}(s_0) - \hat{s}_{n-2}(s_0)] + L(1 - \varepsilon)^2 d(\hat{\varphi}, \varphi). \]

Iterating this, we obtain

\[ \left| \mathbb{E}[f(s_n(s_0))] - \mathbb{E}[f(s_n(s_0))] \right| \leq \frac{Ld(\hat{\varphi}, \varphi)}{\varepsilon}. \]
Integrating this by an invariant measure $\hat{\mu}^*$ of $\hat{\varphi}$, we have

$$\left| \int \mathbb{E} \left[ f \left( s_n \left( s_0 \right) \right) \right] \hat{\mu}^* \left( ds_0 \right) - \int \mathbb{E} \left[ f \left( \hat{s}_n \left( s_0 \right) \right) \right] \hat{\mu}^* \left( ds_0 \right) \right| \leq \frac{Ld \left( \hat{\varphi}, \varphi \right)}{\varepsilon}.$$ 

Note that the second term on the left-hand side is equal to $\int f \left( s \right) \hat{\mu}^* \left( ds \right)$. From Theorem 4.1, for every $s_0$, $\mathbb{E} \left[ f \left( s_n \left( s_0 \right) \right) \right]$ converges uniformly on $\beta \left( S \right)$ to $\int f \left( s \right) \mu^* \left( ds \right)$. Finally, we obtain

$$\left| \int f \left( s \right) \mu^* \left( ds \right) - \int f \left( s \right) \hat{\mu}^* \left( ds \right) \right| \leq \frac{Ld \left( \hat{\varphi}, \varphi \right)}{\varepsilon}.$$ 

Q.E.D.

**Theorem A.2.** (Change of variable formula). Let $g \colon \mathbb{R}^n \rightarrow \mathbb{R}^m$ be Lipschitz continuous, $n \leq m$. Then, for each measurable function $p \colon \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\int_{\mathbb{R}^n} p \left( x \right) Jg \left( x \right) m \left( dx \right) = \int_{\mathbb{R}^n} \left[ \sum_{x \in g^{-1} \left( y \right)} p \left( x \right) \right] m \left( dy \right).$$

**PROOF OF PROPOSITION 5.1:** Let

$$\prod_{t=1}^{T} p_j \left( y_t \mid y_{t-1}; \theta \right) = \int \left( \prod_{t=1}^{T} p \left( \varepsilon^t_1; \theta \right) p_j \left( y_t \mid \varepsilon^t_1, s_0, y_{t-1}; \theta \right) d\varepsilon^t_1 \right) \mu^*_j \left( ds_0; \theta \right),$$

$$\prod_{t=1}^{T} \tilde{p}_j \left( y_t \mid y_{t-1}; \theta \right) = \int \left( \prod_{t=1}^{T} p \left( \varepsilon^t_1; \theta \right) \tilde{p}_j \left( y_t \mid \varepsilon^t_1, S_0, y_{t-1}; \theta \right) d\varepsilon^t_1 \right) \tilde{\mu}^*_j \left( ds_0; \theta \right),$$

$$\prod_{t=1}^{T} p \left( y_t \mid y_{t-1}; \theta \right) = \int \left( \prod_{t=1}^{T} p \left( \varepsilon^t_1; \theta \right) p \left( y_t \mid \varepsilon^t_1, S_0, y_{t-1}; \theta \right) d\varepsilon^t_1 \right) \mu^* \left( ds_0; \theta \right).$$

Consider the convergence of

$$p_j \left( y_t \mid \varepsilon^t_1, s_0, y_{t-1}; \theta \right) = p \left( \eta_j, t; \theta \right) p \left( \varepsilon_j, t; \theta \right) \left| dy_j \left( \eta_j, t, \varepsilon_j, t; \theta \right) \right|,$$

where

$$\left| dy_j \left( v_t, w_{t, \gamma}; \psi \right) \right| = \left| \det \left[ \frac{\partial q_j}{\partial w_{t, \gamma}} \right] \right|.$$

Since the change of variable formula eq. (A.2) holds for $Jg_j \left( \varepsilon, \eta; \theta \right) = \left| dy_j \left( v_{j, t}, w_{j, t, \gamma}; \psi \right) \right|$, we have

$$\int p_j \left( y_t \mid W^t_1, S_0, y_{t-1}; \theta \right) Q \left( d\varepsilon, d\eta \right) = \int p \left( \varepsilon; \theta \right) p \left( \eta; \theta \right) Jg_j \left( \varepsilon, \eta; \theta \right) Q \left( d\varepsilon, d\eta \right).$$

By $d_{C^1} \left( \varphi_j, \varphi \right) \rightarrow 0$ and $d_{C^1} \left( g_j, g \right) \rightarrow 0$, we have

$$\int p \left( \varepsilon; \theta \right) p \left( \eta; \theta \right) Jg_j \left( \varepsilon, \eta; \theta \right) Q \left( d\varepsilon, d\eta \right) \rightarrow \int p \left( \varepsilon; \theta \right) p \left( \eta; \theta \right) Jg \left( \varepsilon, \eta; \theta \right) Q \left( d\varepsilon, d\eta \right).$$
Then, we have

\[
\prod_{t=1}^{T} p_j(y_t \mid y^{t-1}; \theta) \rightarrow \prod_{t=1}^{T} \tilde{p}_j(y_t \mid y^{t-1}; \theta).
\]

By Lemma 1 of Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006), \((\int \prod_{t=1}^{T} p(W_t^1; \theta)p(y_t \mid W_t^1)\,dW_t^1)\) is a continuous function. Further, given Proposition 1 of Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006), we have \(\mu_j^*(dS_0; \theta) \rightarrow \mu^*(dS_0; \theta)\). Therefore,

\[
\prod_{t=1}^{T} \tilde{p}_j(y_t \mid y^{t-1}; \theta) \rightarrow \prod_{t=1}^{T} p(y_t \mid y^{t-1}; \theta).
\]

Q.E.D.