
# Fitted Value Function Iteration With Probability One Contractions

Jenö Pál and John Stachurski

ANU Working Papers in Economics and Econometrics
No. 560

October, 2011

ISBN: 086831 560 5

# FITTED VALUE FUNCTION ITERATION WITH PROBABILITY ONE CONTRACTIONS

JENÖ PÁL AND JOHN STACHURSKI

ABSTRACT. This paper studies a value function iteration algorithm that can be applied to almost all stationary dynamic programming problems. Using nonexpansive function approximation and Monte Carlo integration, we develop a randomized fitted Bellman operator and a corresponding algorithm that is globally convergent with probability one. When additional restrictions are imposed, an $O_P(n^{-1/2})$ rate of convergence for Monte Carlo error is obtained. *Keywords:* Dynamic programming, value iteration, Monte Carlo

## 1. INTRODUCTION

Many economic models contain stochastic dynamic programs (SDPs), either as representations of competitive equilibria, or, more commonly, as sub-problems defining the behavior of firms, households, or other individual agents. When solving these SDPs, computational constraints remain a major bottleneck. The difficulty is particularly acute in settings where the SDP must be solved at a large number of different parameterizations, either to compute equilibria (as in Bewley models and dynamic games), or to estimate structural econometric models with unknown parameters in the primitives of the SDP.

In recent years, many specialist algorithms have been proposed. These algorithms take advantages of certain features of a given application in order to obtain fast convergence rates. In most of these studies, global (or even

*Date*: September 30, 2011.

The authors are indebted to Toni Braun, Chiaki Hara, Hidehiko Ichimura, Timothy Kam, Takashi Kamihigashi, Pedro Gomis-Porqueras, Kevin Reffett, Juan Pablo Rincón-Zapatero, Yiannis Vailakis and seminar participants at Australian National University, Tokyo University, and the 10th SAET conference in Singapore for many helpful comments; and also to the Institute of Economic Research at Kyoto University for hosting some of our joint work.

local) convergence of the algorithm to the optimum is not proved. Instead, the authors test the algorithm against a special case possessing analytical solutions, and then compare their convergence rates against competing algorithms. This leaves open the question of convergence for more realistic problems without analytical solutions. It is for precisely these problems that numerical methods are actually needed.

There are good reasons to be concerned about whether common dynamic programming routines do in fact converge to optimal value functions and policies. This is particularly the case in continuous state space settings, where iterative techniques involve some form of function approximation. The interplay between function approximation and dynamic programming routines appears to be relatively delicate. For example, in value function iteration, many function approximation techniques fail to preserve the global contraction properties of the Bellman operator, and several authors have demonstrated how adding very standard function approximation steps can lead to cycles and a failure of convergence (Boyan and Moore, 1995; Baird, 1995; Tsitsiklis and Van Roy, 1996).

While specialist algorithms that are known to converge quickly in particular settings certainly have their place, in this paper our aim is to study a simulation-based value iteration algorithm that has guaranteed convergence properties across a very wide variety of applications. Our set up includes a function approximation step, admits continuous action spaces, and makes no use of densities. We provide general conditions under which the fixed point of our random fitted Bellman operator converges uniformly to the value function with probability one. Under additional regularity conditions, we show that the sup norm deviation is $O_P(n^{-1/2})$.

The technique described here is a natural alternative to discretized value function iteration, which also has very broad applicability, and remains a popular benchmark in economic applications. In discretized value function iteration, a continuous state/action problem is replaced by a "nearby" discrete problem. Relative to the method we study here, discretization has several disadvantages. First, while the discretized algorithm always locates the solution to the discrete problem, the deviation between this discrete solution and the solution to the original problem is not easily bounded. To the

best of our knowledge, no global convergence results are available in a setting as general as the one that we treat here.[1] Second, in terms of finite time properties, discrete representation of continuous curves is costly relative to continuous parametric representations[2] and inherently subject to the curse of dimensionality.[3]

1.1. **Related Literature.** The emphasis in the paper is on global convergence in a general setting. Numerical methods for more specific models with additional structure can be found in many papers. The literature is too large too survey here. Useful introductions can be found in Marimon and Scott (2001) or Aruoba et al. (2006). The latter includes discussion of methods for solving smooth SDPs where optimal policies satisfy Euler equations. (We assume no such smoothness here.) An iterative method for concave SDPs is analyzed in the recent paper of Fukushima and Waki (2011). Some well-known algorithms based on Monte Carlo include those found in Keane and Wolpin (1994), Rust (1997), Pakes and McGuire (2001) and Longstaff and Schwartz (2001).

Recently, several authors have published studies on finite-time bounds for fitted value function iteration. These include the seminal work of Rust (1997), who proposes an ingenious function approximation step that can be implemented when the one-step transition probabilities for the dynamic programming problem are absolutely continuous with respect to Lebesgue measure (i.e., the distribution for the next period state given current state and action can be represented by a density). Rust proved that for decision

[1] One issue is that discretization errors for continuous curves tend to be bounded in terms of derivatives, which fail to exist or cannot be bounded in many economic settings. Optimal growth models often have unbounded derivatives as a result of Inada conditions. Derivatives can fail to exist when models included discrete choices, binding constraints, non-convexities and so on.

[2] An analogy can be made with raster graphics (i.e., bitmap) files, which are less efficient than vector graphics files for storing regular curves.

[3] The number of data points needed to represent function in $\mathbb{R}^d$ parametrically may be polynomial in $d$, while discrete representations are always exponential. The intractability of discrete representations in moderate to high dimensions has led practitioners in fields such as engineering and computer science to *reverse* the discretization process, replacing discrete dynamic programs with continuous ones. This idea dates back to Bellman (Bellman et al., 1963).

problems satisfying certain Lipschitz conditions, his algorithm breaks the curse of dimensionality, in the sense that worst-case computational complexity is polynomial in the dimension of the state space. Further important developments for models satisfying similar restrictions were reported in Munos and Szepesvári (2008).

Unlike these contributions, our paper studies only consistency, rather than attempting to derive finite time bounds. However, our setting is considerably less restrictive. For example, we do not assume that the one-step transition probabilities are absolutely continuous, an assumption that was central to Rust's algorithm. This additional generality is important in economics, since many applications have one-step transition probabilities that fail to be absolutely continuous. To give an example, consider a benchmark macroeconomic model, where next period capital stock is given by

$$k_{t+1} = (1 - \delta)k_t + f(k_t, z_t) - c_t$$

Here $\delta$ is a depreciation rate, $f$ is a production function, $c_t$ is consumption and $(z_t)$ is an exogenous shock process, typically Markovian. Observe that as soon as the current state $(k_t, z_t)$ and the current action $c_t$ is given, next period capital is deterministic. As a result, the one-step transition probability fails to be absolutely continuous, and cannot be represented by a density.

In this example, the problem is caused by stochastic rank deficiency—the shock space has lower dimension that the state space. While the example is simplistic, it is also representative of the growth and macroeconomic literature—see for example the standard formulation of Stokey and Lucas (1989, p. 290)—and illustrates the fact that a great many models in these fields cannot be treated with density-based approaches.

Failure of absolute continuity can also be caused by discrete shocks (e.g., labor productivity shocks following discrete Markov chains), occasionally binding constraints and other common features. Representative dynamic programming problems where the transition probability fails to be absolutely continuous include those found in Kydland and Prescott (1982), Cooley and Hansen (1989), Imrohoroglu (1989), Aiyagari (1994), Stockman and Tesar (1995), Huggett (1997), Santos and Vigo-Aguiar (1998), Boldrin *et al.* (2001), Clementi and Hopenhayn (2006), Cooper and Haltiwanger (2006) and Arellano (2008).

1.2. **Outline.** Section 2 of the paper provides background concepts and notation. Section 3 defines the model, and section 4 introduces the algorithm. Section 5 provides convergence results, section 6 discusses rates of convergence, and section 7 concludes. Remaining proofs can be found in section 8.

## 2. PRELIMINARIES

We begin by introducing notation. For topological space $\mathbb{T}$, the symbol $\mathscr{C}(\mathbb{T})$ denotes the collection of continuous, bounded, real-valued functions on $\mathbb{T}$, while $\|\cdot\|$ is the supremum norm on $\mathscr{C}(\mathbb{T})$. Operator $S\colon \mathscr{C}(\mathbb{T}) \to \mathscr{C}(\mathbb{T})$ is called a contraction of modulus $\rho$ if $0 \leq \rho < 1$ and

$$\|Sv - Sw\| \leq \rho\|v - w\| \text{ for all pairs } v, w \in \mathscr{C}(\mathbb{T}) \tag{1}$$

$S$ is called nonexpansive if (1) holds with $\rho = 1$. By Banach's contraction mapping theorem, every contraction $S$ of modulus $\rho$ on $\mathscr{C}(\mathbb{T})$ has a unique fixed point $W \in \mathscr{C}(\mathbb{T})$, and, moreover, $\|S^n w - W\| = O(\rho^n)$ for each $w \in \mathscr{C}(\mathbb{T})$.

**Lemma 2.1.** *Let $S$ and $S'$ be operators from $\mathscr{C}(\mathbb{T})$ to itself.*

1. *If $S$ is nonexpansive and $S'$ is a contraction of modulus $\rho$, then the composition $S \circ S'$ is a contraction of modulus $\rho$.*
2. *If $S$ and $S'$ are both contractions of modulus $\rho$ with fixed points $W$ and $W'$ respectively, then $\|W - W'\| \leq (1-\rho)^{-1}\|SW' - W'\|$.*

Part 1 is trivial. For a proof of part 2, see, for example, Rust (1997, lemma 2.1).

In what follows, all random variables are defined on a common probability space $(\Omega, \mathscr{F}, \mathbf{P})$, and $\mathbf{E}$ is the expectation with respect to $\mathbf{P}$. If $X$ is a map from $\Omega$ into $\mathbb{R}$ that is not necessarily measurable, then the outer expectation of $X$ is $\mathbf{E}^* X := \inf_Y \mathbf{E}Y$, where the infimum is over all real random variables $Y$ such that $X \leq Y$ and $\mathbf{E}Y$ exists. For a sequence of possibly nonmeasurable maps $(U_n)$ from $\Omega$ into a metric space $(\mathbb{T}, d)$ and a $\mathbb{T}$-valued random variable $U$, we say that $U_n \to U$ holds $\mathbf{P}^*$-almost surely if there exists a measurable real-valued sequence $\Delta_n$ with $d(U_n, U) \leq \Delta_n$ and $\mathbf{P}\{\Delta_n \to 0\} = 1$. We say that $U_n$ converges in distribution to $U$ if $\mathbf{E}^* g(U_n) \to \mathbf{E}g(U)$ for every $g \in \mathscr{C}(\mathbb{T})$. For the former convergence we write $U_n \overset{a.s.^*}{\to} U$, while for the latter we write $U_n \overset{d^*}{\to} U$.

The continuous mapping theorem continues to hold in this setting:

**Lemma 2.2.** *If $\mathbb{T}'$ is another metric space and $g\colon \mathbb{T} \to \mathbb{T}'$ is continuous, then*

$$U_n \xrightarrow{d^*} U \implies g(U_n) \xrightarrow{d^*} g(U)$$

Let $(X_n)_{n\geq 1}$ be a sequence of (not necessarily measurable) maps from $\Omega$ into $\mathbb{R}$. We write $X_n = O_{P^*}(n^{-1/2})$ if there exists a sequence of real-valued random variables $(\Delta_n)_{n\geq 1}$ such that $|X_n| \leq \Delta_n$ for all $n$ and $\Delta_n = O_P(n^{-1/2})$.

## 3. Set Up

In this section we introduce a general stochastic dynamic programming problem and describe the value function iteration algorithm.

3.1. **The Model.** Consider an SDP of the following form. A controller observes the state $x \in \mathbb{X}$ of a given system, and responds with an action $a$ from a feasible set $\Gamma(x) \subset \mathbb{A}$. Given this state-action pair $(x, a)$, the controller receives current reward $r(x, a)$, and the new state is determined as $x' = F(x, a, U)$, where $U$ is a draw from a fixed distribution $\phi$. The process now repeats. The controller's objective is to maximize the sum of expected discounted rewards given discount factor $\rho$.

The sets $\mathbb{X}$ and $\mathbb{A}$ are referred to as the state and action spaces respectively, and $\Gamma$ is called the feasible correspondence. We let

$$\mathbb{G} := \operatorname{graph}\Gamma := \{(x, a) \in \mathbb{X} \times \mathbb{A} : a \in \Gamma(x)\}$$

The set $\mathbb{G}$ is called the set of feasible state-action pairs.

A feasible policy is a Borel measurable map $\sigma\colon \mathbb{X} \to \mathbb{A}$ such that $\sigma(x) \in \Gamma(x)$ for all $x \in \mathbb{X}$. Let $\Sigma$ be the set of all such policies. The controller's problem is

$$\max_{\sigma\in\Sigma} \left\{ \mathbf{E} \sum_{t=0}^{\infty} \rho^t r(X_t, \sigma(X_t)) \right\}$$

subject to

$$X_{t+1} = F(X_t, \sigma(X_t), U_{t+1}) \quad \text{with } x_0 \text{ given} \tag{2}$$

Almost any stationary infinite horizon dynamic program with additively separable preferences can be formulated in this way.

We assume throughout the paper that

(1) $\mathbb{X}$ and $\mathbb{A}$ are compact metric spaces.
(2) $\Gamma$ is continuous and compact-valued.
(3) The shocks $(U_t)_{t\geq 1}$ are IID with common distribution $\phi$.[4]
(4) $\phi$ is a Borel probability measure over metric space $\mathbb{U}$.
(5) The reward function $r\colon \mathbb{G} \to \mathbb{R}$ is continuous.
(6) The function $\mathbb{G} \ni (x,a) \mapsto F(x,a,u) \in \mathbb{X}$ is continuous for all $u \in \mathbb{U}$.

For $(X_t)$ as given by (2), let $V_\sigma(x_0) = \mathbf{E}\sum_{t=0}^{\infty} \rho^t r(X_t, \sigma(X_t))$. Let $T\colon \mathscr{C}(\mathbb{X}) \to \mathscr{C}(\mathbb{X})$ be the Bellman operator, defined at $v \in \mathscr{C}(\mathbb{X})$ by

$$Tv(x) := \max_{a\in\Gamma(x)} \left\{ r(x,a) + \rho \int v[F(x,a,u)]\phi(du) \right\} \qquad (x \in \mathbb{X}) \tag{3}$$

For $v \in \mathscr{C}(\mathbb{X})$, a policy $\sigma \in \Sigma$ is called $v$-greedy if $\sigma(x)$ is a maximizer of the right-hand size of (3) for all $x \in \mathbb{X}$. The value function $V_T$ is defined pointwise on $\mathbb{X}$ by $V_T(x) = \sup_{\sigma\in\Sigma} V_\sigma(x)$. A policy $\sigma \in \Sigma$ is called optimal if $V_\sigma = V_T$.

3.2. **Value Function Iteration.** The following result is standard:

**Theorem 3.1.** *Under assumptions 1–6 above,*

(1) *$T$ is a contraction of modulus $\rho$ on $\mathscr{C}(\mathbb{X})$, and $V_T$ is the unique fixed point;*
(2) *a policy $\sigma \in \Sigma$ is optimal if and only if it is $V_T$-greedy; and*
(3) *at least one such policy exists.*

In principle, $V_T$ can be calculated by value function iteration (VFI), which involves fixing an initial $v \in \mathscr{C}(\mathbb{X})$ and iterating with $T$. From theorem 3.1

[4]The assumption of IID shocks is not restrictive. For example, consider the following macroeconomic model with exogenous Markov shock sequence: The state space is a product space $K \times Z \subset \mathbb{R}^m \times \mathbb{R}^n$, where $k \in K$ is a vector of endogenous variables and $z \in Z$ is a vector of exogenous variables. Technology is summarized by a feasible set $\Theta \subset K \times Z \times K$. The exogenous process $(z_t)_{t\geq 0}$ evolves according to $z_{t+1} = g(z_t, \epsilon_{t+1})$, where $(\epsilon_t)_{t\geq 1}$ is IID. Instantaneous rewards are given by $v(k,z,k')$. This formulation is a special case of our SDP. To see this, for the state take $x := (k,z) \in K \times Z$, and for the action take $a := k' \in K$. The feasible correspondence is $\Gamma(x) := \Gamma(k,z) := \{k' \in K : (k,z,k') \in \Theta\}$. The shock is $u := \epsilon$, and the transition function is $F(x,a,u) := F(k,z,k',\epsilon) := (k', g(z,\epsilon)) \in K \times Z$. The reward function is $r(x,a) := r(k,z,k') := v(k,z,k')$.

we have $\|T^k v - V_T\| = O(\rho^k)$. Using this fact and optimality of $V_T$-greedy policies, one can show that a $T^k v$-greedy policy is approximately optimal when $k$ is sufficiently large.[5]

## 4. RANDOM FITTED VFI

Evaluation of the expression $r(x,a) + \rho \int v[F(x,a,u)]\phi(du)$ on the right-hand side of (3) requires approximation of the integral. To compute it, one can use deterministic methods or Monte Carlo. Monte Carlo tends to perform better in multiple dimensions, and preserves the contractiveness of the Bellman operator. In addition, we can evaluate every integral by drawing a *single* sample

$$U_1, \ldots, U_n \overset{\text{IID}}{\sim} \phi \tag{4}$$

once off, and then iterating with the random Bellman operator $R_n$ defined by

$$R_n v(x) := \max_{a \in \Gamma(x)} \left\{ r(x,a) + \rho \frac{1}{n} \sum_{i=1}^{n} v[F(x,a,U_i)] \right\} \qquad (x \in \mathbb{X}) \tag{5}$$

A realization of $\omega \in \Omega$ determines a particular realization $(U_i(\omega))_{i=1}^n$ of the sample (4), which in turn defines a realization $R_n(\omega)$ of $R_n$. Each realization $R_n(\omega)$ is an operator from $\mathscr{C}(\mathbb{X})$ to itself.

A second numerical issue is as follows: If $\mathbb{X}$ is infinite, then, for arbitrary given $w \in \mathscr{C}(\mathbb{X})$, one cannot evaluate either $Tw(x)$ or $R_n w(x)$ at each $x \in \mathbb{X}$ in finite time (or store the functions in a look-up table). Hence, we approximate $R_n w$ using a finite parametric representation. To do so, we introduce an approximation operator $A \colon \mathscr{C}(\mathbb{X}) \to \mathscr{A}(\mathbb{X}) \subset \mathscr{C}(\mathbb{X})$, where, given function $w \in \mathscr{C}(\mathbb{X})$, $Aw$ is an approximation of $w$, and $\mathscr{A}(\mathbb{X})$ is a class of functions such that each element can be represented by a finite number of parameters. In addition, we assume that $Aw$ can be computed on the basis of a finite number of observations (i.e., by observing the value of $w(x)$ at a finite number of $x \in \mathbb{X}$). For example, the mapping $w \mapsto Aw$ might proceed by evaluating $w$ on a fixed and finite grid of points $\{x_j\}_{j=1}^J$, and then

[5] See, e.g., (Puterman, 1994, theorem 6.3.1). An appropriate $k$ is usually chosen according to some stopping criterion that depends on the deviation between successive iterates of $T$.

constructing $Aw$ based on these "interpolation" points. Finally, we assume throughout that $A$ is nonexpansive.

**Example 4.1.** Continuous piecewise linear interpolation in $\mathbb{R}^d$ is a nonexpansive approximation scheme.[6]

**Example 4.2.** Another nonexpansive approximation scheme is provided by kernel smoothers. Let $d$ be a metric on $\mathbb{X}$, and let $\{x_i\}_{i=1}^m$ be a finite subset of $\mathbb{X}$. For given $v \in \mathscr{C}(\mathbb{X})$, let

$$Av(x) = \sum_{i=1}^{m} p_i(x) \psi \left\{ \frac{d(x, x_i)}{h} \right\} \tag{6}$$

where $h$ is a parameter, and $\psi\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous monotone decreasing function. The values $p_i(x)$ are called weights. In most cases, the weights are set to $p_i(x) = v(x_i) / \{\sum_{i=1}^m \psi[d(x, x_i)/h]\}$. For this choice of weights, the operator $A$ is nonexpansive (see, e.g., Stachurski, 2008).

Other nonexpansive schemes include $k$-nearest neighbors, shape-preserving Schumaker splines, and the variation-diminishing splines of Schoenberg.[7]

The complete procedure for random fitted value function iteration is given in the following algorithm:

**Algorithm 1**: Random Fitted VFI

1 generate the sample $(U_1, \ldots, U_n) \overset{\text{IID}}{\sim} \phi$ in (4) ;
2 fix $v \in \mathscr{C}(\mathbb{X})$ ;
3 compute $v_k := (AR_n)^k v$ iteratively ;
4 compute a $v_k$-greedy policy $\sigma$ ;

[6] To describe it, let $\mathbb{X}$ be a convex subset of $\mathbb{R}^d$, let $\mathbb{V}$ be a finite subset of $\mathbb{X}$ such that the convex hull of $\mathbb{V}$ is $\mathbb{X}$, and let $T$ be a $\mathbb{V}$-triangularization of $\mathbb{X}$. (That is, $T$ is a finite collection of non-degenerate simplexes such that the vertices of each simplex lie in $\mathbb{V}$ and any two simplexes intersect on a common face or not at all.) Given a simplex $\Delta \in T$ with vertices $\zeta_1, \ldots, \zeta_{d+1}$, each $x \in \Delta$ can be represented uniquely as $\sum_{i=1}^{d+1} \lambda(x, i)\zeta_i$, where $\lambda(x, i)$ is its $i$-th barycentric coordinate relative to $\Delta$. (By definition, $\lambda(x, i) \geq 0$ and $\sum_{i=1}^{d+1} \lambda(x, i) = 1$.) For $v \in \mathscr{C}(\mathbb{X})$, we define $A$ by $Av(x) = \sum_{i=1}^{d+1} \lambda(x, i) v(\zeta_i)$. The operator $A$ is nonexpansive (see, e.g., Stachurski, 2008).

[7] See Stachurski (2008) for further discussion and references.

In step 3, $(AR_n)^k$ is the $k$-th iterate of $AR_n$. In practice, when applying the operator $AR_n$ to a given function $w$, first $R_n w$ is evaluated on a finite grid of points $\{x_j\}_{j=1}^J$ by solving the maximization problem in (5) $J$ times. $A$ is then applied to produce the fitted function $AR_n w$.

## 5. Analysis

We begin our analysis with the following lemma:

**Lemma 5.1.** *The operator $AT$ is a contraction on $\mathscr{C}(\mathbb{X})$ of modulus $\rho$. The operator $AR_n(\omega)$ is also a contraction on $\mathscr{C}(\mathbb{X})$ of modulus $\rho$ for all $n \in \mathbb{N}$ and all $\omega \in \Omega$.*

As a consequence, there exists

(1) a unique fixed point $V_{AT} \in \mathscr{C}(\mathbb{X})$ of $AT$
(2) a unique fixed point $V_{AR_n(\omega)} \in \mathscr{C}(\mathbb{X})$ of $AR_n(\omega)$ for each $\omega \in \Omega$

The operator $AT$ is the fitted Bellman operator where function approximation is included, but the integral is computed exactly. Its fixed point $V_{AT}$ is deterministic. On the other hand, $V_{AR_n}$ is random. In what follows, we refer to $V_{AR_n}$ as a random function, although $\omega \mapsto V_{AR_n(\omega)}$ may not be Borel measurable as a mapping from $\Omega$ to $\mathscr{C}(\mathbb{X})$.

Our primary goal is to study the convergence of $V_{AR_n}$ to the value function $V_T$.[8] By the triangle inequality, the error can be decomposed as

$$\|V_T - V_{AR_n}\| \leq \|V_T - V_{AT}\| + \|V_{AT} - V_{AR_n}\| \qquad \forall\, n \in \mathbb{N} \tag{7}$$

Let us consider the two terms on the right-hand side of (7). The first term is the function approximation error, caused by replacing $T$ with $AT$. The second is the integral approximation error, caused by replacing $AT$ with $AR_n$. The following two sections consider these two errors in turn.

[8]The relative optimality of the $(AR_n)^k v$-greedy policy $\sigma$ computed by algorithm 1 depends on the deviation between $(AR_n)^k v$ and $V_T$. Using the triangle inequality, we can bound the latter by $\|(AR_n)^k v - V_{AR_n}\| + \|V_{AR_n} - V_T\|$. By lemma 5.1, the first term is $O(\rho^k)$ in $k$. Convergence of $V_{AR_n}$ to $V_T$ is less clear, and hence we focus on this term.

5.1. **Function Approximation Error.** Analysis of the function approximation error $\|V_T - V_{AT}\|$ is relatively straightforward. The details will depend on the particular function approximation scheme used in a given implementation, but sufficiently "fine" approximations will make the error arbitrarily small. To give an example, consider the kernel smoother $A$ in (6). The next result is proved in the appendix.

**Lemma 5.2.** *For any $\epsilon > 0$, there exists a choice of $\{x_i\}_{i=1}^m$, $\psi$ and $h$ such that the corresponding operator $A$ in* (6) *satisfies $\|V_T - V_{AT}\| < \epsilon$.*

5.2. **Integral Approximation Error.** Next, we turn our attention to the integral approximation error, which is the second term on the right-hand side of (7). Our first major result for the paper shows probability one convergence without any additional assumptions.[9]

**Theorem 5.1.** $\|V_{AT} - V_{AR_n}\| \overset{a.s.^*}{\to} 0$ *as* $n \to \infty$.

**Remark 5.1.** Theorem 5.1 provides no indication of the rate of convergence. This problem is treated in section 6.

*Proof of theorem 5.1.* By lemma 2.1 and the nonexpansiveness of $A$, we have

$$\begin{aligned} \|V_{AR_n} - V_{AT}\| &\leq \frac{1}{1-\rho}\|AR_nV_{AT} - V_{AT}\| \\ &= \frac{1}{1-\rho}\|AR_nV_{AT} - ATV_{AT}\| \\ &\leq \frac{1}{1-\rho}\|R_nV_{AT} - TV_{AT}\| \end{aligned}$$

Hence, to prove theorem 5.1, it is sufficient to prove that $\|R_nV_{AT} - TV_{AT}\|$ converges to zero with probability one. To bound this term, we make use of the following standard inequality: If $g, g' \in \mathscr{C}(\mathbb{Y})$ for compact set $\mathbb{Y}$, then

$$|\max g - \max g'| \leq \max|g - g'| =: \|g - g'\| \tag{8}$$

[9]Since Borel measurability of $\omega \mapsto V_{AR_n(\omega)}$ is problematic, the theorem uses the concept of $\mathbf{P}^*$-almost sure convergence.

Using (8), we obtain

$$|R_n V_{AT}(x) - TV_{AT}(x)| \leq \rho \max_{a \in \Gamma(x)} \left| \frac{1}{n} \sum_{i=1}^{n} V_{AT}[F(x,a,U_i)] - \int V_{AT}[F(x,a,u)]\phi(du) \right|$$

where $x \in \mathbb{X}$ is arbitrary. Taking the supremum over all $x \in \mathbb{X}$, we now have

$$\|R_n V_{AT} - TV_{AT}\| \leq \rho \max_{(x,a) \in \mathbb{G}} \left| \frac{1}{n} \sum_{i=1}^{n} V_{AT}[F(x,a,U_i)] - \int V_{AT}[F(x,a,u)]\phi(du) \right| \tag{9}$$

Let $y = (x,a)$ denote a typical element of $\mathbb{G}$, and let

$$h_y(u) :=: h_{(x,a)}(u) := V_{AT}[F(x,a,u)] :=: V_{AT}[F(y,u)] \tag{10}$$

Also, for $h \colon \mathbb{U} \to \mathbb{R}$, let $\phi_n(h) := \frac{1}{n}\sum_{i=1}^{n} h(U_i)$ and $\phi(h) := \int h d\phi$. Using this notation, (9) becomes

$$\|R_n V_{AT} - TV_{AT}\| \leq \rho \max_{y \in \mathbb{G}} |\phi_n(h_y) - \phi(h_y)| \tag{11}$$

A class $\mathscr{H}$ of bounded measurable functions mapping $\mathbb{U}$ into $\mathbb{R}$ is called $\phi$-Glivenko-Cantelli if $\sup_{h \in \mathscr{H}} |\phi_n(h) - \phi(h)| \to 0$ $\mathbf{P}^*$-almost surely as $n \to \infty$. A sufficient condition for this property[10] is that $\mathscr{H}$ consists of functions $h_\alpha \colon \mathbb{U} \to \mathbb{R}$ with index $\alpha$ in metric space $\Lambda$, and, moreover:

(1) $\Lambda$ is compact;
(2) $\Lambda \ni \alpha \mapsto h_\alpha(u) \in \mathbb{R}$ is continuous for every $u \in \mathbb{U}$; and
(3) there exists a measurable function $H \colon \mathbb{U} \to \mathbb{R}$ such that $\int H d\phi < \infty$ and $|h_\alpha| \leq H$ for every $\alpha \in \Lambda$.

In our case, the relevant class of functions is $\{h_y\}_{y \in \mathbb{G}}$, where $h_y$ is defined in (10). This family of functions satisfies the sufficient conditions in 1–3 above. First, $\mathbb{G}$ is a compact metric space, due to our assumptions on $\mathbb{X}$, $\mathbb{A}$ and $\Gamma$. Second, $\mathbb{G} \ni y \mapsto h_y(u) := V_{AT}[F(y,u)] \in \mathbb{R}$ is continuous for every $u \in \mathbb{U}$, due to continuity of $V_{AT}$ and $F$. Third, $|h_y(u)|$ is bounded above by the finite constant $\|V_{AT}\|$ for all $y \in \mathbb{G}$ and $u \in \mathbb{U}$. Hence, $\{h_y\}_{y \in \mathbb{G}}$ is $\phi$-Glivenko-Cantelli. This concludes the proof. □

[10] See, for example, van der Vaart, 1998, p. 272.

## 6. Rates of Convergence

The result in theorem 5.1 gives no indication as to the rate of convergence. To obtain a rate, we need to give a rate for the right-hand side of (11). The $\phi$-Glivenko-Cantelli property used in the proof of theorem 5.1 is not sufficient for rates, so further restrictions on $\{h_y\}_{y\in\mathbb{G}}$ are required.

6.1. **Donsker Classes.** Let $\mathscr{H}$ be a class of uniformly bounded, measurable functions from $\mathbb{U}$ into $\mathbb{R}$, and let $(b\mathscr{H}, \|\cdot\|)$ be the Banach space of bounded, real valued functions on $\mathscr{H}$ with the supremum norm. The class $\mathscr{H}$ is called $\phi$-Donsker if

$$\nu_n(h) := \sqrt{n}(\phi_n(h) - \phi(h)) \qquad (n \in \mathbb{N},\, h \in \mathscr{H})$$

converges in distribution to a tight Gaussian process $\nu$ in the space $b\mathscr{H}$. Here $\omega \mapsto \nu_n(\cdot)(\omega)$ and $\omega \mapsto \nu(\cdot)(\omega)$ are maps from $\Omega$ into $b\mathscr{H}$. The maps $\omega \mapsto \nu_n(\cdot)(\omega)$ are not necessarily measurable, and convergence in distribution is to be understood in the sense of $\nu_n \xrightarrow{d^*} \nu$.

**Proposition 6.1.** *If $\{h_y\}_{y\in\mathbb{G}}$ is $\phi$-Donsker, then $\|V_{AR_n} - V_{AT}\| = O_{P^*}(n^{-1/2})$.*

*Proof.* We need some preliminary results and additional notation. Let

$$G_n(y) := \nu_n(h_y) := \sqrt{n}(\phi_n(h_y) - \phi(h_y)) \qquad (n \in \mathbb{N},\, y \in \mathbb{G})$$

$G_n$ can be understood as a real-valued stochastic process indexed by $y \in \mathbb{G}$:

$$G_n(y)(\omega) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} h_y(U_i(\omega)) - \int h_y(u)\phi(du)\right) \in \mathbb{R}$$

Regarding measurability, we have the following result, proved in the appendix:

**Lemma 6.1.** *For each $n \in \mathbb{N}$, the following measurability results hold:*

1. *$\omega \mapsto G_n(\cdot)(\omega)$ is a $\mathscr{C}(\mathbb{G})$-valued random variable, and*
2. *$\omega \mapsto \|G_n(\cdot)(\omega)\| = \sup_{y\in\mathbb{G}} |G_n(y)(\omega)|$ is a real-valued random variable.*

In view of (11), we have

$$\|V_{AR_n} - V_{AT}\| \leq \frac{\rho}{1-\rho} n^{-1/2} \sup_{y\in\mathbb{G}} |G_n(y)|$$

Since $\mathscr{H} := \{h_y\}_{y \in \mathbb{G}}$ is $\phi$-Donsker, we have $\nu_n \xrightarrow{d^*} \nu$, where $\nu$ is a Gaussian process on $\mathscr{H}$. By lemma 2.2 and continuity of the norm $\|\cdot\|$ on $b\mathscr{H}$, we then have $\|\nu_n\| \xrightarrow{d^*} \|\nu\|$ in $\mathbb{R}$. Observe that

$$\|\nu_n\| = \sup_{h \in \mathscr{H}} |\nu_n(h)| = \sup_{y \in \mathbb{G}} |\nu_n(h_y)| = \sup_{y \in \mathbb{G}} |G_n(y)|$$

and hence $\sup_{y \in \mathbb{G}} |G_n(y)| \xrightarrow{d^*} \|\nu\|$. By part 2 of lemma 6.1, this is convergence in distribution in the regular sense, and, as a consequence, we have $\sup_{y \in \mathbb{G}} |G_n(y)| = O_P(1)$. Therefore

$$\|V_{AR_n} - V_{AT}\| \leq \frac{\rho}{1-\rho} n^{-1/2} O_P(1) = O_P(n^{-1/2})$$

This concludes the proof of proposition 6.1. □

6.2. **The Lipschitz Case.** In this section and the next, we use proposition 6.1 to obtain sufficient conditions for rates of convergence in different (somewhat specialized) settings. Our first result is based on a Lipschitz condition. To apply the method, we add the following assumptions:

(i) $\mathbb{G} \subset \mathbb{R}^d$.
(ii) $Aw$ is Lipschitz continuous for every $w \in \mathscr{C}(\mathbb{X})$.[11]
(iii) There exists a measurable function $m_0 \colon \mathbb{U} \to \mathbb{R}$ with $\int m_0^2 \, d\phi < \infty$ and[12]

$$\|F(y,u) - F(y',u)\|_2 \leq m_0(u) \|y - y'\|_2 \qquad \forall \, y, y' \in \mathbb{G}, \; u \in \mathbb{U} \tag{12}$$

The next result is proved in the appendix.

**Proposition 6.2.** *If* (i)–(ii) *hold, then* $\|V_{AR_n} - V_{AT}\| = O_{P^*}(n^{-1/2})$.

Notice that the assumptions concern only the transition function, not the reward function. Many dynamic macroeconomic models have Lipschitz transition rules. The consumer's problem in the incomplete markets models of Aiyagari (1994) and Huggett (1997) are obvious examples, and many recent

[11] This condition depends on the approximation architecture used in the fitted VFI routine, and is satisfied by, for example, the piecewise linear interpolation operator in Example 4.1.

[12] Here $\|\cdot\|_2$ represents the euclidean norm on $\mathbb{R}^d$.

variations have a similar structure (see, e.g., Pijoan-Mas, 2006, or Ábrahám and Cárceles-Poveda, 2010).

An important special case of our Lipschitz assumption is models with linear transition rules. The next lemma provides details.

**Lemma 6.2.** *If $\mathbb{U} \subset \mathbb{R}^k$, and $F$ is linear, in the sense that*

$$F(x, a, u) = Ax + Ba + Cu \qquad (x \in \mathbb{X},\ a \in \Gamma(x),\ u \in \mathbb{U}) \tag{13}$$

*for matrices $A$, $B$ and $C$, then assumption (ii) is satisfied.*

6.3. **The Monotone Case.** Another way to establish the $\phi$-Donsker property is via monotonicity. To this end, we drop the Lipschitz assumptions of section 6.2 and replace them with the following:

(i) $\mathbb{X} \subset \mathbb{R}^d$ and $\mathbb{U} \subset \mathbb{R}$.
(ii) $A$ maps $i\mathscr{C}(\mathbb{X})$ to itself, where $i\mathscr{C}(\mathbb{X})$ is the increasing functions in $\mathscr{C}(\mathbb{X})$.
(iii) For all $x, x' \in \mathbb{X}$ with $x \leq x'$, $\Gamma(x) \subset \Gamma(x')$; $r(x, a) \leq r(x', a)$ for all $a \in \Gamma(x)$; and $F(x, a, u) \leq F(x', a, u)$ for all $a \in \Gamma(x)$ and $u \in \mathbb{U}$.
(iv) For all $y \in \mathbb{G}$, $F(y, u) \leq F(y, u')$ whenever $u \leq u'$.

**Proposition 6.3.** *If* (i)–(iv) *hold, then* $\|V_{AR_n} - V_{AT}\| = O_{P^*}(n^{-1/2})$.

The proof is given in the appendix.

Assumption (ii) depends on the approximation architecture, and is satisfied by, for example, the linear interpolation operator in Example 4.1. The other assumptions are satisfied by a number of standard models.

**Example 6.1.** Consider the growth model

$$\max \mathbb{E} \sum_{t=0}^{\infty} \rho^t u(c_t) \quad \text{subject to}$$
$$c_t \geq 0, \quad k_{t+1} \geq 0, \quad c_t + k_{t+1} \leq f(k_t, z_t)$$

Suppose that $z_t$ is Markov, following transition rule $z_{t+1} = g(z_t, U_{t+1})$, where $(U_t)_{t \geq 1}$ is IID. The state is $(k, z) \in \mathbb{R}^2_+$. To write the model in our framework, we take $F(k, z, k', u) = (k', g(z, u))$, $r(k, z, k') = u(f(k, z) - k')$ and $\Gamma(k, z) = [0, f(k, z)]$. If $f$ and $g$ are both increasing, then assumptions (iii) and (iv) above are satisfied.

## 7. CONCLUSION

We studied a Monte Carlo VFI algorithm with function approximation. We proved that the algorithm is consistent for a wide variety of models. This guaranteed convergence stands in contrast to many other numerical techniques proposed in the literature. Under additional restrictions, we established a parametric rate of convergence, independent of the dimension of the state, action and shock spaces.

Many avenues for future research exist. First, we identified only two cases where the $\phi$-Donsker property is satisfied (the Lipschitz and monotonicity conditions of sections 6.2 and 6.3). Additional research should illuminate other cases. In addition, we treated only stationary, additively separable, infinite horizon SDPs, leaving open the cases of nonstationary models, optimal stopping, and general recursive utility. All of these issues are left for future study.

## 8. REMAINING PROOFS

*Proof of lemma 5.1.* The contractiveness of $AT$ follows from lemma 2.1. Next we consider contractiveness of $R_n$. Fix $n \in \mathbb{N}$ and $\omega \in \Omega$. Let $R := R_n(\omega)$. Fix $w, w' \in \mathscr{C}(\mathbb{X})$ and $x \in \mathbb{X}$. In view of (8), we have

$$|Rw(x) - Rw'(x)| \leq \rho \max_{a \in \Gamma(x)} \left| \frac{1}{n} \sum_{i=1}^{n} w[F(x, a, U_i(\omega))] - \frac{1}{n} \sum_{i=1}^{n} w'[F(x, a, U_i(\omega))] \right|$$

Using the triangle inequality and the definition of $\| \cdot \|$, we obtain

$$|Rw(x) - Rw'(x)| \leq \rho \|w - w'\|$$

Taking the supremum over $x \in \mathbb{X}$ yields the desired result.

Finally, contractiveness of $AR_n$ now follows from lemma 2.1. □

*Proof of lemma 5.2.* Fix $\epsilon > 0$. By lemma 2.1, we have

$$\|V_T - V_{AT}\| \leq (1 - \rho)^{-1} \|AV_T - V_T\| \tag{14}$$

Since $\mathbb{X}$ is compact, $V_T$ is uniformly continuous, and we select $\delta > 0$ with $|V_T(x) - V_T(y)| < (1 - \rho)\epsilon$ whenever $d(x, y) < \delta$. Using compactness again, we choose $\{x_i\}_{i=1}^m$ such that, given any $x \in \mathbb{X}$, there exists at least

one $x_i$ with $d(x, x_i) < \delta$. Finally, we choose $\psi$ such that $\psi(u) = 0$ whenever $u$ is greater than some constant $M$,[13] and $h$ such that $Mh < \delta$.

Now fix any $x \in \mathbb{X}$. Letting $\lambda(x, i) := \psi[d(x, x_i)/h] / \sum_j \psi[d(x, x_j)/h]$, we can write

$$|AV_T(x) - V_T(x)| = \left| \sum_i \lambda(x, i) V_T(x_i) - V_T(x) \right| \leq \sum_i \lambda(x, i) |V_T(x_i) - V_T(x)|$$

If $d(x, x_i) \geq \delta$, then $d(x, x_i)/h \geq M$, and hence $\psi[d(x, x_i)/h] = \lambda(x, i) = 0$. For the remaining terms in the sum we have $d(x, x_i) < \delta$, and hence $|V_T(x_i) - V_T(x)| < (1 - \rho)\epsilon$. Since $x$ is arbitrary, we have $\|AV_T - V_T\| < (1 - \rho)\epsilon$. Combining this bound with (14) completes the proof of the lemma. □

*Proof of lemma 6.1.* We begin by proving measurability of $\omega \mapsto H(\cdot)(\omega)$, where

$$H(y)(\omega) = h_y(U(\omega)) = V_{AT}[F(y, U(\omega))]$$

Since $\mathbb{G}$ is compact in the product topology, the Stone–Weierstrass theorem implies that $\mathscr{C}(\mathbb{G})$ is separable. Hence, by the Pettis measurability theorem, we need only show that $\omega \mapsto \ell(H(\cdot)(\omega))$ is measurable for each $\ell$ in the dual space $\mathscr{C}(\mathbb{G})^*$ of $\mathscr{C}(\mathbb{G})$. By the Riesz representation theorem, $\mathscr{C}(\mathbb{G})^*$ can be identified with $\mathscr{M}(\mathbb{G})$, the space of finite signed Borel measures on $\mathbb{G}$. Thus, it remains to show that

$$\Omega \ni \omega \mapsto \int H(y)(\omega)\, \gamma(dy) \in \mathbb{R} \text{ is measurable} \quad \forall\, \gamma \in \mathscr{M}(\mathbb{G})$$

To this end it is sufficient to show that $H(y)(\omega) = V_{AT}[F(y, U(\omega)]$ is measurable with respect to the product $\sigma$-algebra $\mathscr{B}_{\mathbb{G}} \otimes \mathscr{F}$, where $\mathscr{B}_{\mathbb{G}}$ is the Borel $\sigma$-algebra on $\mathbb{G}$. Since $H$ is continuous with respect to $y$ and measurable with respect to $\omega$, $H$ is a Carathéodory function (Aliprantis and Border, 2006, definition 4.50). As $\mathbb{G}$ is separable, measurability with respect to $\mathscr{B}_{\mathbb{G}} \otimes \mathscr{F}$ is established (Aliprantis and Border, 2006, lemma 4.51).

Given measurability of $\omega \mapsto H(\cdot)(\omega)$, measurability of $\omega \mapsto G_n(\cdot)(\omega)$ follows from the fact that linear combinations of measurable random elements of a separable Banach space are themselves measurable.

[13] A typical example is the Epanechnikov kernel.

Regarding the second claim in the lemma, measurability of $\omega \mapsto \|G_n(\cdot)(\omega)\|$ follows from measurability of $\omega \mapsto G_n(\cdot)(\omega)$, continuity of the norm as a map from $\mathscr{C}(\mathbb{G})$ to $\mathbb{R}$, and the fact that continuous transformations of measurable mappings are measurable. □

*Proof of proposition 6.2.* By proposition 6.1, it suffices to show that the class $\{h_y\}_{y\in\mathbb{G}}$ is $\phi$-Donsker when (i)–(iii) hold. A sufficient condition for $\{h_y\}_{y\in\mathbb{G}}$ to be $\phi$-Donsker is the existence of a measurable function $m\colon \mathbb{U} \to \mathbb{R}$ such that $\int m^2 d\phi < \infty$ and

$$|h_y(u) - h_{y'}(u)| \leq m(u)\|y - y'\|_2 \qquad \forall\, y, y' \in \mathbb{G},\ u \in \mathbb{U} \tag{15}$$

(see, e.g., van der Vaart, 1998, p. 271). To find such an $m$, observe that $V_{AT}$ is Lipschitz, as follows from (ii) and the relation $V_{AT} = ATV_{AT}$. As a consequence, there exists a $K < \infty$ such that, for any $y, y' \in \mathbb{G}$ and $u \in \mathbb{U}$, we have

$$\begin{aligned} |h_y(u) - h_{y'}(u)| &:= |V_{AT}[F(y,u)] - V_{AT}[F(y',u)]| \\ &\leq K\|F(y,u) - F(y',u)\|_2 \leq Km_0(u)\|y - y'\|_2 \end{aligned}$$

where $m_0$ is the function in (iii). Letting $m := Km_0$, we see that $\int m^2 d\phi = K^2 \int m_0^2 d\phi < \infty$. All the conditions are now verified, and hence $\{h_y\}_{y\in\mathbb{G}}$ is $\phi$-Donsker. □

*Proof of lemma 6.2.* To see this, observe that for any $y = (x, a) \in \mathbb{G}$, $y' = (x', a') \in \mathbb{G}$, and $u \in \mathbb{U}$,

$$\begin{aligned} &\|Ax + Ba + Cu - Ax' - Ba' - Cu\|_2 \\ &\qquad = \|A(x - x') + B(a - a')\|_2 \leq \gamma(\|x - x'\|_2 + \|a - a'\|_2) \end{aligned}$$

where $\gamma$ is the maximum of the operator norms of $A$ and $B$. Since $y = (x, a) \mapsto \|x\|_2 + \|a\|_2 \in \mathbb{R}$ defines a norm on $\mathbb{R}^d$, and since all norms on $\mathbb{R}^d$ are equivalent, we obtain

$$\|F(y,u) - F(y',u)\|_2 \leq M\gamma\|y - y'\|_2 \qquad \forall\, y, y' \in \mathbb{G},\ u \in \mathbb{U}$$

for some $M < \infty$. This verifies (ii). □

*Proof of proposition 6.3.* From van der Vaart (1998, p. 273), it suffices to show that the class $\{h_y\}_{y\in\mathbb{G}}$ is uniformly bounded on $\mathbb{U}$, and that each element $h_y$

is monotone increasing on $\mathbb{U}$. Since $h_y(u) = V_{AT}[F(y,u)]$, uniform boundedness will hold if $V_{AT}$ is bounded on $\mathbb{X}$. That this is the case follows from the fact that $\mathbb{X}$ is compact and $V_{AT} \in \mathscr{C}(\mathbb{X})$.

Regarding monotonicity, we begin by showing that $V_{AT}$ is monotone increasing. To see that this is the case, observe that $V_{AT}$ is the fixed point of $AT$ in $\mathscr{C}(\mathbb{X})$. Since $i\mathscr{C}(\mathbb{X})$ is a closed subset of $\mathscr{C}(\mathbb{X})$, we need only show that $AT$ maps $i\mathscr{C}(\mathbb{X})$ into itself. Since $A\colon i\mathscr{C}(\mathbb{X}) \to i\mathscr{C}(\mathbb{X})$ by assumption, it remains to verify that $T$ also has this property. For a proof of this fact, see Stachurski (2009, theorem 12.1.2). As a result, $V_{AT}$ is increasing, and the claim in the proposition now follows from assumption (iv) above. □

## REFERENCES


[1] Ábrahám, A. and E. Cárceles-Poveda (2010): "Endogenous Trading Constraints with Incomplete Asset Markets," *Journal of Economic Theory,* 145, 974–1004

[2] Aiyagari, Rao S. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, Vol. 109(3), 659-84.

[3] Aliprantis, C. D. and K. C. Border (2006): *Infinite Dimensional Analysis,* 3rd Edition, Springer-Verlag, New York.

[4] Arellano, C. (2008): "Default Risk and Income Fluctuations in Emerging Economies," *The American Economic Review,* 98 (3), 690-712.

[5] Arouba, S. B., J. Fernandez-Villaverde and J.F. Rubio-Ramirez (2006): "Comparing Solution. Methods for Dynamic Equilibrium Economies," *Journal of Economic Dynamics and Control,* 30, 2477–2508.

[6] Baird, L.C. (1995): "Residual Algorithms: Reinforcement Learning with Function Approximation," *Machine Learning: Proceedings of the Twelfth International Conference,* Morgan Kaufman Publishers, San Francisco, CA, 9–12.

[7] Bellman, R. E., R. Kalaba and B. Kotkin (1963): "Polynomial Approximation–A New Computational Technique in Dynamic Programming," *Mathematics of Computation,* 17 (82), 155–161.

[8] Boldrin, M. L. J. Christiano and J. D. M. Fisher (2001): "Habit Persistence, Asset Returns, and the Business Cycle," *The American Economic Review,* 91 (1), 149–166.

[9] Boyan, J. and A. Moore (1995): "Generalization in Reinforcement Learning: Safely Approximating the Value Function," in G. Tesauro, D.S. Touretzky, and T.K. Lee (eds), *Neural Information Processing Systems 7,* MIT Press, Cambridge, MA, 369–376.

[10] Clementi, G. L. and H. Hopenhayn (2006): "A Theory of Financing Constraints and Firm Dynamics," *Quarterly Journal of Economics,* 121 (1), 229–265.

[11] Cooley, T. F. and G. D. Hansen (1989): "The Inflation Tax in a Real Business Cycle Model," *American Economic Review,* 79 (4), 733-48.

[12] Cooper, R. W. and J. C. Haltiwanger (2006): "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, 73 (3), 611-633.

[13] Fukushima, K. and Y. Waki (2011): "A Polyhedral Approximation Approach to Concave Numerical Dynamic Programming," mimeo, University of Wisconsin.

[14] Imrohoroglu, A. (1989): "Cost of Business Cycles with Indivisibilities and Liquidity Constraints," *Journal of Political Economy*, 97 (6), 1364-1383.

[15] Keane, M. P. and K. I. Wolpin (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation," *The Review of Economics and Statistics*, 76 (4), 648–672.

[16] Kydland, F. E. and E. C. Prescott (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50 (6), 1345-1370.

[17] Longstaff, F. A. and E. S. Schwartz (2001): "Valuing American Options by Simulation: A Simple Least-squares Approach," *Review of Financial Studies*, 14, 113–147.

[18] Marimon, R. and A. Andrew Scott, eds., (2001): *Computational Methods for the Study of Dynamic Economies*, Oxford University Press, USA.

[19] Martins-da-Rocha, F. V. and Y. Vailakis (2010): "Existence and Uniqueness of a Fixed Point for Local Contractions," *Econometrica*, 78 (3), 1127–1141.

[20] Munos, R and Szepesvári, C. (2008): "Finite Time Bounds for Fitted Value Iteration," *Journal of Machine Learning Research*, 1, 815–857.

[21] Pakes, A. and P. McGuire (2001): "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the 'Curse' of Dimensionality," *Econometrica*, 69 (5), 1261–1281.

[22] Pijoan-Mas, J. (2006): "Precautionary Savings or Working Longer Hours?" *Review of Economic Dynamics*, 9 (2), 326–352.

[23] Puterman, M. (1994): *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, New York.

[24] Rincon-Zapatero, J. P., and C. Rodriguez-Palmero (2003): "Existence and Uniqueness of Solutions to the Bellman Equation in the Unbounded Case," *Econometrica* 71, 1519–55.

[25] Rust, J. (1997): "Using Randomization to Break the Curse of Dimensionality," *Econometrica*, 65 (3), 487–516.

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

[27] Stachurski, J. (2008): "Continuous State Dynamic Programming via Nonexpansive Approximation," *Computational Economics*, 31 (2), 141–160.

[28] Stachurski, J. (2009): *Economic Dynamics: Theory and Computation*, The MIT Press, Cambridge MA.

[29] Stockman, A. C. and L. L. Tesar (1995): "Tastes and Technology in a Two-Country Model of the Business Cycle: Explaining International Comovements," *American Economic Review*, 85 (1), 168-85.

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

[31] Tsitsiklis, J. N. and B. Van Roy (1996): "Feature-Based Methods for Large Scale Dynamic Programming," *Machine Learning*, 22, 59–94.

[32] van der Vaart, A. W. (1998): *Asymptotic Statistics*, Cambridge series in statistical and probabilistic mathematics, Cambridge University Press, UK.

DEPARTMENT OF ECONOMICS, CENTRAL EUROPEAN UNIVERSITY, 1051 BUDAPEST, NÁDOR U. 9., HUNGARY

*E-mail address*: pal_jeno@ceu-budapest.edu

RESEARCH SCHOOL OF ECONOMICS, AUSTRALIAN NATIONAL UNIVERSITY, CANBERRA 0200, AUSTRALIA

*E-mail address*: john.stachurski@anu.edu.au