working paper

12 10R

Approximating High-Dimensional Dynamic Models: Sieve Value Function Iteration

Peter Arcidiacono, Patrick Bayer,
Federico A. Bugni, and Jonathan James

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 12-10R November 2012*

Approximating High-Dimensional Dynamic Models: Sieve Value Function Iteration

Peter Arcidiacono, Patrick Bayer, Federico A. Bugni, and Jonathan James

Many dynamic problems in economics are characterized by large state spaces which make both computing and estimating the model infeasible. We introduce a method for approximating the value function of high-dimensional dynamic models based on sieves and establish results for the: (a) consistency, (b) rates of convergence, and (c) bounds on the error of approximation. We embed this method for approximating the solution to the dynamic problem within an estimation routine and prove that it provides consistent estimates of the model's parameters. We provide Monte Carlo evidence that our method can successfully be used to approximate models that would otherwise be infeasible to compute, suggesting that these techniques may substantially broaden the class of models that can be solved and estimated.

Keywords: Large state space, Dynamic decision problem, Sieve approximation, Value function, Value function iteration.

JEL Codes: C02, C44, C60, C63

*First version April 2012.

Peter Arcidiacono is at Duke University (psarcidi@econ.duke.edu), Patrick Bayer is at Duke University (patrick.bayer@duke.edu), Federico A. Bugni is at Duke University (federico.bugni@duke.edu), and Jonathan James is at the Federal Reserve Bank of Cleveland (jwj8@duke.edu). The authors thank Pedro Mira, R. Vijay Krishna, and Peng Sun for useful comments and suggestions. Arcidiacono, Bayer, and Bugni thank the National Science Foundation for research support via grant SES-1124193.

1 Introduction

Dynamic problems in economics (and many other disciplines) are characterized by large state spaces that routinely make the model difficult or impossible to compute. Faced with these challenges, researchers proceed in a number of different ways including analyzing a static version of the problem, simplifying the dynamic problem to a setting where the state space is manageable, or employing a number of techniques for approximating the value and policy functions that characterize the solution to the problem. The problems analyzed in many broad literatures in economics ranging from models of market equilibrium (which are almost always treated in a static framework) to strategic or social interactions in networks (static or myopic dynamics) to dynamic games (typically characterized by a small number of players and states) to matching problems (almost always analyzed as static problems) remain limited by this large state space problem.

In recent years, techniques for approximating value functions in large state space problems have been developed using simulation and interpolation (Keane and Wolpin (1994)) or parametric policy iteration (PPI) (Rust (2000) and Bénitez-Silva, Hall, Hitsch, Pauletto, and Rust (2000)). The main idea behind these methods is to approximate the value function using flexible functions of the relevant state variables. Simulation and interpolation has been applied by Keane and Wolpin (1997) and Crawford and Shum (2005) among many others. PPI has been applied to a single agent setting in Hendel and Nevo (2006) and to a dynamic game in Sweeting (2011). While the potential application of these methods is great, the literature contains very few formal results regarding the quality of the approximation and, therefore, provides little formal or practical foundation for researchers to use as they apply these methods.

In this paper, we consider an alternative, but related, method for approximating value functions that we refer to as sieve value function iteration (SVFI).[1] We develop the method in the context of a generic single agent dynamic programming problem that can have either a finite or infinite horizon. The SVFI method involves approximating the integrated value function with a nonparametric sieve function. For any sieve function (i.e., for a particular choice of parameters), one can evaluate the Bellman equation and compute a notion of distance between the approximation and its contraction, and thus characterize how close the Bellman equation comes to holding. We approximate the value function by choosing the parameters of the sieve function that come as close as possible to making the Bellman equation hold exactly. Since the sieve space is a simple space, this minimization problem is relatively easy to solve. Moreover, as the sequence of sieve spaces becomes a better and better approximation of the original functional space $\mathcal{F}$, our approximation converges to the true value function.

In order to analyze the formal properties of the SVFI method, we assume that the complexity of the sieve space, n, increases to infinity. In this sense our SVFI approximation technique becomes non-parametric and we establish a number of results:

1. *Consistency:* we show that the sieve approximation converges to the true value function as

[1] For an excellent review of the method of sieves see Chen (2007).

the richness of the sieve space increases.

2. *Rates of convergence:* we provide the first results in the literature for the rate at which the approximating function converges to the true value function.

3. *Iteration of Bellman operator:* we characterize how rates of convergence are affected by iterating the Bellman operator, pointing out for the first time that this has the potential to improve performance in certain applications.

4. *Bound on the error of approximation:* following arguments in Rust (2000), we provide an upper bound to the error of approximation to the unknown true value function that is feasible to compute.

While the consistency of the SVFI might be an expected result, establishing the rate of convergence of the SVFI approximation and understanding the effects of increasing the number of iterations on the quality of the approximation are two useful formal contributions to the literature. The fourth result can be related to results available in Rust (2000). The availability of bounds on the error of approximation is an incredibly valuable feature of SVFI because it ensures that one can bound the extent of approximation error relative to the true value function even when the true value function cannot be computed, i.e., the case of relevance in very large state space problems. This also serves as a natural guide to researchers as they implement the SVFI, providing a clear sense of whether the approximation is reasonable for any given specification of the sieve function. Taken together, these results provide a formal foundation for the use of the SVFI for approximating large state space problems.

The SVFI approach is quite straightforward to understand and implement and, thus, has the potential to be widely applied in economics and other disciplines. The framework can also be flexibly implemented. It is possible, for example, to estimate the parameters of the sieve approximation function by minimizing the distance in the Bellman operator for only a large subset of the states in the state space. This is attractive for problems with large finite or infinite state spaces.

The method can be applied equally well to infinite and finite horizon problems. For finite horizon problems, we develop two strategies for approximating value function. First, we show how SVFI can be used to approximate the value functions at each time period using a traditional backwards recursion solution method. More interestingly, by including the time to the horizon as another state variable in the sieve function, we show how it is possible to approximate a single sieve function that provides an approximation of the value function at each point in time without solving the problem backwards. The particular features of the application in question will generally determine which of these approaches is computationally lighter or easier to implement.

Having developed this general method for approximating value functions, we then formally show how the SVFI can be used to estimate the structural parameters of large state space dynamic problems in empirical applications. Conceptually it is easy to see that for any particular guess of the model's structural parameters, the SVFI can be used to approximate the solution to the dynamic

problem and thus compute the approximate value and policy functions as well as conditional choice probabilities. By comparing these objects to their empirical counterparts in the data it is possible to compute a wide variety of objective functions that would be appropriate for estimating the structural parameters. We show that it is possible to consistently estimate the model's structural parameters by embedding the SVFI within the estimation algorithm.

We close the paper by demonstrating the performance of SVFI in a particular Monte Carlo application, inspired by the famous bus-engine replacement problem of Rust (1987), in which a firm must dynamically manage its entire fleet of buses. We begin by analyzing an infinite horizon problem with a state space that is finite and relatively large but manageable, i.e., for which it is still possible to compute the value function exactly at each state. We demonstrate that SVFI approximates value functions for this problem very closely in a tiny fraction of the time that it takes to compute the exact solution to the problem. We show that an accurate approximation can be obtained even when the Bellman operator is evaluated at a small, randomly, drawn subset of the full set of states and compare the speed and accuracy of alternative methods for minimizing this distance, including non-linear least squares, an iterative least squares method, and methods that iterate the Bellman operator. We extend this problem to an infinite state space by adding a continuous state variable, again demonstrating that it is possible to approximate the solution exceptionally well in a reasonably short amount of computation time. We then analyze an analogous finite horizon problem, comparing the speed and accuracy of SVFI approximations using (a) the traditional backwards recursion solution method in which we compute a separate sieve approximation function at each point in time and (b) an approach that treats time to the horizon as a state variable, yielding a single time-interacted sieve approximation function.

We complete the Monte Carlo section of the paper, by evaluating the performance of SVFI approximation within the context of the estimation of the structural parameters of the model. Returning to the infinite horizon problem, we demonstrate that SVFI accurately estimates with only a minimal impact on the effective standard errors while being much faster to compute. In particular, we propose estimating the sieve parameters outside of the estimation of the structural parameters in a manner similar to Aguirregabiria and Mira (2002) swapping of the nested fixed point algorithm. Further, it is straightforward to examine how the estimated structural parameters are affected by the dimension of the sieve and, therefore, to ensure that the sieve is sufficiently rich such that its impact on the estimated structural parameters is minimal.

At this point, it is important to acknowledge that there is an existent operations research literature that develops approximation methods to dynamic programming problems like the ones addressed in this paper. Extensive surveys of major developments can be found in Bertsekas and Tsitsiklis (1996), Bertsekas (2001), Powell (2008), and, more recently, Bertsekas (2012). These references (and references therein) describe a variety of *approximation methods* or *architectures* based on polynomial functions. For example, Bertsekas and Tsitsiklis (1996, Section 6.10) consider parametric approximation based on "Bellman Error methods" which is closely related to our SVFI approximation. Also, according to Bertsekas (2012, chapter 6), our SVFI can be characterized as

an "indirect approximation method" of the value function based on a "parametric architecture".

Just like in this paper, the operations research literature derives bounds on the error of approximation and, with these, provide consistency type results. See, e.g., Bertsekas and Tsitsiklis (1996, page 332), Bertsekas (2001, page 34), or Bertsekas (2012, page 500). The main difference between our methodological contribution and this literature lies on the nature of the asymptotic analysis. On the one hand, the operations research literature considers the behavior of the approximation using an iterative procedure *keeping the approximating parametric space constant*. On the other hand, our analysis does not involve iterative procedures[2] but it does require the sieve space to increase in complexity in order to derive formal results. Given that our interest is in deriving formal (asymptotic) results, the difference in the nature of the asymptotics implies that these two approaches are clearly distinct contributions. In addition to being theoretically interesting in its own right, the type of asymptotic analysis considered in this paper becomes essential in order to establish the asymptotic properties of the estimation procedure that results from embedding the SVFI within the estimation algorithm.

The rest of the paper proceeds as follows. In section 2 we describe the optimization problem. Sections 3 and 4 develop formal results for infinite horizon problems and finite horizon problems, respectively. In particular, we establish properties of the approximation including consistency and rates of convergence. In section 5, we show how to use SVFI in estimation as well as properties of the estimator. Section 6 investigates the small sample properties of our approximations, establishing both the speed and reliability of our methods. Section 7 concludes.

2 The dynamic decision problem

In this paper, we are interested in approximating the value function of a single agent solving a dynamic decision problem. We begin by introducing the general problem. In every period $t \in \mathbb{N}$, the agent observes the current value of a state (x_t, ε_t) and chooses an action a_t in a finite choice set $A(x_t)$. The first component x_t is observed by the researcher whereas the second component $\varepsilon_t \equiv \{\varepsilon_t(a) : a \in A(x_t)\}$ is not. We assume that $x_t \in X$ and $\varepsilon_t \in E$. Conditional on this information, the agent solves the following optimization problem:

$$V_t(x_t, \varepsilon_t) = \sup_{\Pi_t} E\left\{ \sum_{j=t}^{T} \beta^{j-t}[u(x_j, a_j) + \varepsilon_j(a_j)] \middle| x_t, \varepsilon_t \right\}, \tag{2.1}$$

where $\Pi_t = \{\{a_j\}_{j=t}^{\infty} : a_t = A(x_t)\}$ and $u(x_t, a_t) + \varepsilon_t(a_t)$ denotes the period utility of making decision a_t in state (x_t, ε_t), $u(x_t, a_t)$ representing the structural part (possibly known up to a finite dimensional parameter) and $\varepsilon_t(a_t)$ representing the residual part, unobserved by the researcher.

[2] As we explain later the minimization problem we propose can be computationally complex to implement. For this reason, we provide an iterative algorithm with the sole purpose of approximating a solution to this minimization problem. Nevertheless, if it were possible to solve the minimization problem directly (i.e., without using the iterative algorithm), then our methodology does not really require the use of iterations.

The objective of this paper is to provide a computationally feasible approximation to the value functions (i.e., $\{V_t(\cdot)\}_{t=1}^T$) and study its properties. In general, we might be interested in approximating these function because we want to do welfare analysis (i.e., which can be conducted directly using these functions) or because we are interested in any other feature of the problem that can be computed from these functions (i.e., optimal decision rules, conditional choice probabilities (CCPs), etc.).

The formulation in Eq. (2.1) encompasses both finite horizon problems (i.e., $T = \infty$) and infinite horizon problems (i.e., $T < \infty$). Following the dynamic programming literature, our strategy to approximate the value function in Eq. (2.1) is to show that it is the unique solution to a (functional) Bellman equation, and construct an approximation based on this representation. Given that the Bellman equation representation of the finite and infinite horizon problems are fundamentally different, it is convenient to divide the rest of the discussion into these two cases.

3 Approximation in infinite horizon problems

This section describes the application of non-parametric sieve methods to approximate the value functions in Eq. (2.1) when the dynamic problem has an infinite horizon (i.e., $T = \infty$). The distinctive feature of the infinite horizon problem is that the value function of the problem is stationary (i.e., $V_t(\cdot) = V(\cdot)$, $\forall t \in \mathbb{N}$), provided that we impose mild additional assumptions to the problem. In particular, we follow Rust (1987, 1988) and assume that the joint stochastic process $\{x_t, \varepsilon_t, a_t\}$ is a controlled Markov process that satisfies the following conditional independence (CI) assumption:

$$\begin{aligned} dP(x_{t+1}, \varepsilon_{t+1}|x_t, \varepsilon_t, a_t, x_{t-1}, \varepsilon_{t-1}, a_{t-1}, \dots) &= dP(x_{t+1}, \varepsilon_{t+1}|x_t, \varepsilon_t, a_t), \text{ (Markov)} \\ &= dP(\varepsilon_{t+1}|x_{t+1}, x_t, \varepsilon_t, a_t) dP(x_{t+1}|x_t, \varepsilon_t, a_t), \\ &= dP(\varepsilon_{t+1}|x_{t+1}) dP(x_{t+1}|x_t, a_t). \text{ (CI)} \end{aligned}$$

Under these assumptions, the literature provides several ways of formulating the value function in Eq. (2.1) as the recursive solution of a Bellman equation. In particular, these include: (a) the *social surplus function* formulation discussed in Rust (1987, 1988), (b) the *conditional value function* formulation of Rust (1987), and (c) the *choice-specific value function* formulation of Rust (1988). Rather than describing each of these formulations in the main text, we now introduce a single *unified formulation* that encompasses all of these formulations.[3]

According to our unified formulation, the value functions in Eq. (2.1) are stationary and solve the following Bellman equation:

$$V(s) = \max_{a \in A(s)} \{u(s,a) + \beta E(F(V, s')|s, a)\}, \forall s \in S, \tag{3.1}$$

[3]Appendix B describes each of these formulations and shows that all of them are special cases of our unified formulation.

where s is the current value of the state, s' is the future value of the state, S is the state space, a represents the action chosen by the agent, $A(s)$ is the set of actions available to the agent when the state is s, and F is a known functional of the value function $V(\cdot)$ and the future state s' that satisfies certain known properties.[4]

3.1 Approximating the value function

In order to discuss the approximation of V, we must first define the space to which this function belongs. In particular, assume that V belongs to a functional space, denoted $\mathcal{F}$. For example, we can take $\mathcal{F}$ to be the space of measurable, bounded, real-valued functions from S to $\mathbb{R}$. We define a metric d in this space, making $(\mathcal{F}, d)$ a normed vector space. If we do not indicate otherwise, the metric is the sup-norm metric, i.e., for any $f_1, f_2 \in \mathcal{F}$:

$$d(f_1, f_2) = \sup_{s \in S} |f_1(s) - f_2(s)|. \tag{3.2}$$

Furthermore, we assume that this metric space is complete, i.e., it is a Banach space. Consider the following (functional) operator $\Gamma : \mathcal{F} \to \mathcal{F}$:

$$[\Gamma\theta](s) \equiv \max_{a \in A(s)} \{u(s,a) + \beta E(F(\theta, s')|s, a)\}, \forall s \in S. \tag{3.3}$$

According to the definition in Eq. (3.1), the value function is the fixed point of this operator. Furthermore, under Blackwell (1965) sufficient conditions (also see, e.g., Stokey and Lucas (1989, Theorem 3.3)), the operator Γ can be shown to be a contraction mapping. Thus, as a consequence of the contraction mapping theorem (see, e.g., Stokey and Lucas (1989, Theorem 3.2)), it follows that the value function is the *unique* fixed point of the contraction mapping, i.e., the *unique* solution to:

$$\min_{\theta \in \mathcal{F}} \; d(\theta, \Gamma\theta) \tag{3.4}$$

As a consequence, if it is possible to solve the minimization problem in Eq. (3.4), then the solution has to be unique and equal to the value function. Unfortunately, there are several situations of practical relevance in which this minimization problem is computationally infeasible. In this paper we focus on the difficulties that arise when the state space is too large, i.e., the cardinality of the set S is either infinity or finite but too large to permit the use of traditional methods.

The approximation method we propose is inspired by the sieves non-parametric estimation method. Instead of solving the original minimization problem (i.e., Eq. (3.4)), we replace the original (possibly infinite dimensional) parameter space $\mathcal{F}$ with a sequence of simpler (often finite dimensional) parameter spaces, called sieves. Throughout this paper, the sequence of sieve spaces will be denoted by $\{\Theta_n\}_{n \geq 1}$, where $n \in \mathbb{N}$ is an index that represents the complexity of the sieve

[4] Among other properties, $F : \mathcal{F} \times S$ satisfies: (a) monotonicity: for functions f_1, f_2 with $f_1(s) \leq f_2(s)$ for all $s \in S$, then $F(f_1, s') \leq F(f_2, s')$ and (b) discounting: for any function f and any $\alpha \in \mathbb{R}$, $\beta F(f + \alpha, s') = \beta F(f, s') + \beta\alpha$.

space. In order for this replacement to produce an accurate approximation of the unknown value function, it will be required that the sieve space sequence $\{\Theta_n\}_{n\geq 1}$ to become increasingly more complex (i.e., for any $n \in \mathbb{N}$, $\Theta_n \subset \Theta_{n+1} \subseteq \mathcal{F}$) and dense in $\mathcal{F}$ as $n \to \infty$.

For a given sieve space Θ_n, our method produces an approximation, denoted $\hat{\theta}_n$. In essence, we replace the original parameter space $\mathcal{F}$ by the sieve parameter space Θ_n and, loosely speaking, our approximation will be given as:

$$\hat{\theta}_n \approx \arg\min_{\theta \in \Theta_n} d(\theta, \Gamma\theta).$$

That is, we will seek to choose the parameters of the sieve to get as close to a fixed point of the Bellman operator as possible. Naturally, the quality of the approximation will be determined by the sieve space Θ_n used to approximate the original parameter space $\mathcal{F}$.

We introduce a definition of consistency of the approximation and rate of convergence of the approximation.

Definition 3.1 (Consistency)**.** *$\hat{\theta}_n$ is a consistent approximation to V if and only if:*

$$d(\hat{\theta}_n, V) = o_p(1), \text{ as } n \to \infty.$$

To be precise, this is an approximation problem and not an estimation problem. In other words, there are no data and, hence, no random sampling error[5].

Definition 3.2 (Rate of Convergence)**.** *$\hat{\theta}_n$ converges to V at a rate of γ_n^{-1} if and only if:*

$$d(\hat{\theta}_n, V) \leq O_p(\gamma_n),$$

where $\gamma_n = o(1)$ as $n \to \infty$.

3.2 Assumptions

We now provide a list of the assumptions used in this section. As we show in Example 3.1, all of these assumptions are satisfied in dynamic decision problems that have a very large but finite state space.

Assumption A.1. $(\mathcal{F}, d)$ is a complete metric space of functions that map S into $\mathbb{R}$ and Γ defined in Eq. (3.3) is a contraction mapping with modulus $\beta \in (0, 1)$.

Assumption A.2. For any $n \in \mathbb{N}$, d_n is a pseudo-metric in $(\mathcal{F}, d)$ such that $\exists K_1, K_2 > 0$,

$$K_1 \; d_n(f_1, f_2) - \eta_{1,n} \leq d(f_1, f_2) \leq K_2 \; d_n(f_1, f_2) + \eta_{1,n},$$

where $\eta_{1,n} = O_p(\gamma_{1,n})$ and $\gamma_{1,n} = o(1)$, uniformly in $f_1, f_2 \in \mathcal{F}$.

[5]The reason to use $o_p(1)$ instead of $o(1)$ is that, in general, we allow for randomness in the approximation. The randomness can occur, e.g., in the choice of the sieve space or the solution to the approximation problem.

Assumption A.3. For some $k \in \mathbb{N}$, we can find $\theta_n \in \Theta_n$ that satisfies:

$$d_n(\theta_n, \Gamma^k \theta_n) \leq \inf_{\theta \in \Theta_n} d_n(\theta, \Gamma^k \theta) + \eta_{2,n}. \tag{3.5}$$

where Γ^k is the k^{th} iteration of Γ, $\eta_{2,n} = O_p(\gamma_{2,n})$, and $\gamma_{2,n} = o(1)$.

Assumption A.4. For any $f \in \mathcal{F}$:

$$\inf_{\theta \in \Theta_n} d(\theta, f) = \eta_{3,n}(f),$$

where $\eta_{3,n}(f) = O_p(\gamma_{3,n}(f))$ and $\gamma_{3,n}(f) = o(1)$.

We now briefly comment on each of the assumptions. Even though Assumption A.1 might look innocuous, it is not. The definition of a contraction mapping is associated to a metric space $(\mathcal{F}, d)$ and, in particular, to the metric d. In other words, if $(\mathcal{F}, d)$ and $(\mathcal{F}, \check{d})$ are two metric spaces, it is possible that a mapping Γ is a contraction mapping with respect to d but not with respect to $\check{d}$. This is exactly what happens in the context of single agent dynamic decision problems. According to Blackwell (1965) sufficient conditions, Γ in Eq. (3.3) is a contraction mapping with respect to the sup-norm metric but it may not not be a contraction mapping with respect to more convenient metrics, such as the l_2-metric. See, e.g., the discussion in Bertsekas and Tsitsiklis (1996, page 369).

In cases in which the state space S is too large, it might not be computationally possible to work with the metric d, but we find convenient to replace this metric with an associated pseudo-metric d_n. In order for this replacement to produce interesting results, we need to assume that the difference between the metric d and the pseudo-metric d_n vanishes at a certain rate. This is what Assumption A.2 accomplishes.

If we had the computational capability to solve the minimization problem in Eq. (3.4), then the solution to this problem would be exactly object of interest. The motivation for this paper, however, is to consider cases in which the size of the state space S makes solving the minimization problem in Eq. (3.4) impossible. Assumption A.3 describes the operation that is within our computational possibilities.

Instead of solving the problem in the original space $\mathcal{F}$, we replace the space with a simpler sieve space Θ_n that approximates $\mathcal{F}$. Assumption A.3 implies that, by the virtue of this simplification, we are now able to solve certain computational problems. In particular, we assume that we can: (i) compute the k^{th} iteration of Γ for any function in the sieve space Θ_n and (ii) minimize an objective function within the sieve space Θ_n, possibly up to a small error denoted by $\eta_{2,n}$. In case the objective function can be exactly minimized, then $\eta_{2,n} = 0$. Even when restricted to the simpler sieve space Θ_n, the minimization problem in Assumption A.3 can appear to be computationally challenging. For this reason, Section 3.3 describes an algorithm than can be used to implement this step.

The strategy our our approximation is to replace the original parameter space $\mathcal{F}$ with a sequence

of approximating sieve spaces $\{\Theta_n\}_{n\geq 1}$. In order to guarantee that this replacement does not affect the asymptotic properties, Assumption A.4 requires that any function $f \in \mathcal{F}$ can be approximated by an element in Θ_n, up to a small error denoted by $\eta_{3,n}(f)$. In order to guarantee a small error of approximation, it is convenient to choose the sieve space Θ_n that mimics the properties of the original parameter space $\mathcal{F}$. In turn, in order to derive these properties, Stokey and Lucas (1989, Corollary 1, page 52) can prove to be a very valuable tool.

In order to illustrate these assumptions, we consider the following example.

Example 3.1 (Large but finite state space)**.** *Suppose that the agent solves the value function in Eq. (3.1) where the state-space S is finite but large, i.e., $\#S = N < \infty$. By Stokey and Lucas (1989, Corollary 1, page 52) we can show that the value function belongs to $B(S)$, i.e., the space of bounded functions that map S onto $[-B, B] \in \mathbb{R}$, for some $B < \infty$. This implies that we can take the relevant metric space to be $(\mathcal{F}, d)$, with $\mathcal{F} = B(S)$ and d equal to the sup-norm metric, i.e.,*

$$d(f_1, f_2) = \sup_{s\in S} |f_1(s) - f_2(s)| = \max_{i\leq N} |f_1(s_i) - f_2(s_i)|, \ \forall f_1, f_2 \in \mathcal{F}.$$

We now verify all of the assumptions. We begin with Assumption A.1. By arguments in Stokey and Lucas (1989, page 47), $(\mathcal{F}, d)$ is a complete metric space of functions that map S onto $\mathbb{R}$. By the Blackwell (1965) sufficient conditions, it is easy to see that $\Gamma : \mathcal{F} \to \mathcal{F}$ is a contraction mapping with modulus $\beta \in (0, 1)$.

Given that $\#S = N$ is a large number, we might not be able to compute d exactly. Instead, for any $n \in \mathbb{N}$ with $n < N$, we might be able to compute:

$$d_n(f_1, f_2) = \max_{i\leq n} |f_1(s_i) - f_2(s_i)|, \ \forall f_1, f_2 \in \mathcal{F}.$$

This is a pseudo-norm in $(\mathcal{F}, d)$, which we refer as the sup-norm pseudo-metric (this pseudo-norm becomes the sup-norm metric when $n = N$). Notice that:

$$|d_n(f_1, f_2) - d(f_1, f_2)| = \eta_{1,n}, \ \forall f_1, f_2 \in \mathcal{F},$$

with $\eta_{1,n} = \max_{i>n} |f_1(s_i) - f_2(s_i)| \leq \max\{N - n, 0\}B$, and so $\eta_{1,n} = O_p(\gamma_{1,n})$ and $\gamma_{1,n} = o(1)$. This verifies Assumption A.2 with $K_1 = K_2 = 1$.

As we have already pointed out, the sup-norm (pseudo-)metric can complicated the associated optimization problem in Assumption A.3. For this reason, it is convenient to consider alternative pseudo-metrics. For example, one possibility is to use the l_2 pseudo-metric, given by:

$$d_n(f_1, f_2) = \sqrt{\sum_{i\leq n} |f_1(s_i) - f_2(s_i)|^2}, \ \forall f_1, f_2 \in \mathcal{F},$$

In order to verify Assumption A.2 we notice that:

$$d_n(f_1, f_2) \leq \max_{i \leq n} |f_1(s_i) - f_2(s_i)| \leq \sqrt{N}\ d_n(f_1, f_2),$$

and, therefore:

$$d_n(f_1, f_2) + \eta_{1,n} \leq d(f_1, f_2) \leq N^{1/2}\ d_n(f_1, f_2) + \eta_{1,n},$$

with $\eta_{1,n} = \max_{i>n} |f_1(s_i) - f_2(s_i)| \leq \max\{N - n, 0\}B$*, and so it follows that* $\eta_{1,n} = O_p(\gamma_{1,n})$ *and* $\gamma_{1,n} = o(1)$*. This verifies Assumption A.2 with* $K_1 = 1$ *and* $K_2 = \sqrt{N}$*.*

Assumption A.3 assumes that we have computational capabilities to (approximately) minimize a certain objective function. As we have already mentioned, the minimization problem can be approached by using the algorithm described in Section 3.3.

Finally, Assumption A.4 requires that the sieve space Θ_n *can approximate the original space* $\mathcal{F}$*. The non-parametric sieve estimation literature describes many possible choices of sieve spaces that will produce this type of result. For example, for any* $n \in \mathbb{N}$*, we take* Θ_n *to be the set of polynomial functions defined in* S *of degree* n*. Then, the Weierstrass Theorem (see, e.g., Royden (1988, page 128)) implies that:*

$$\inf_{\theta \in \Theta_n} d(\theta, v) = o(1),\ \forall v \in \mathcal{F},$$

as required by Assumption A.4. Results on the rate of convergence of the $o(1)$ *term are available from Chen (2007), and references therein (e.g., Lorentz (1966) and Powell (1981, chapters 15-16)).*

3.3 An iterative procedure to implement the minimization

This paper proposes an approximation method by replacing the original parameter space $\mathcal{F}$ in Eq. (3.4) with a simpler sieve space Θ_n in Eq. (3.5). Unfortunately, the single-step minimization problem in Eq. (3.5) can be computationally challenging to approach as is.

In order to deal with these challenges, we suggest the Algorithm 3.1. Relative to the original minimization problem, this algorithm replaces the single-step minimization problem with a iterative multi-step procedure. We explain the relative advantages of this alternative procedure after introducing the algorithm.

Algorithm 3.1. *Let* $\{\varepsilon_n\}_{n \geq 1}$ *be a tolerance sequence that satisfies* $\varepsilon_n = o(1)$*. For an arbitrary initial function* $f \in \Theta_n$*, consider the following iterative procedure:*

1. *Given* f*, choose a function* $\theta_m \in \Theta_n$ *such that:*

$$\theta_m = \arg\min_{\theta \in \Theta_n} d_n(\theta, \Gamma^k f). \tag{3.6}$$

2. *If* $\max\{d_n(\theta_m, \Gamma^k f), d_n(\theta_m, f)\} \leq \varepsilon_n$*, then stop the algorithm and define* $\theta_n \equiv \theta_m$*. Otherwise, set* $f = \theta_m$ *and return to step 1.*

If the algorithm converges, then it can be shown that the resulting $\theta_n \in \Theta_n$ satisfies Eq. (3.5) with $\eta_{2,n} = O(\max\{\varepsilon_n, \eta_{1,n}\})$.[6]

The pseudo-norm d_n should be chosen so that it makes the minimization problem in Eq. (3.6) easy to solve. For example, if d_n is a weighted l_2 norm and if Θ_n is a finite dimensional linear sieve (i.e., a linear span of finitely many known basis functions[7]), then the minimization problem in Eq. (3.6) can be solved with a closed form by a least squares procedure.

Having defined the dynamic decision problem and laid out basic definitions and assumptions, we now introduce our method for approximating the value function and prove a series of results regarding the properties of this approximation. We begin by analyzing infinite horizon problems, taking up finite horizon problems in the following section of the paper.

3.4 Definition of the approximation

Under Assumption A.1, the Contraction Mapping Theorem (see, e.g., Stokey and Lucas (1989, Theorem 3.2, page 50)) indicates that the contraction mapping Γ defined by Eq. (3.3) has a unique fixed point. As we have explained, Eq. (3.1) implies this fixed point is the value function of interest, which we have denoted by $V \in \mathcal{F}$. As we have explained, this motivates us to consider the (unfeasible) optimization problem in Eq. (3.4) or, more generally, for some $k \in \mathbb{N}$, the following (equally unfeasible) optimization problem:

$$\inf_{\theta \in \mathcal{F}} d(\theta, \Gamma^k \theta).$$

Even though the objective function can be computed, the sample space of the problem under consideration makes the domain of the optimization, $\mathcal{F}$, too complex to handle. In order to circumvent this issue, we consider replacing the space $\mathcal{F}$ with a sequence of approximating spaces or sieves. According to Assumption A.4, the sequence of spaces $\{\Theta_n\}_{n \geq 1}$ becomes a good approximation of $\mathcal{F}$ as our computational possibilities, n, diverge to infinity. In particular, this motives us to consider the following alternative optimization problem:

$$\inf_{\theta \in \Theta_n} d(\theta, \Gamma^k \theta).$$

In certain situations, minimizing with respect to the metric d might not be computationally easy or even possible. In those cases, it is convenient to consider replacing the metric d with a suitable pseudo-metric d_n according to Assumption A.2. This leads us to consider the following optimization problem:

$$\inf_{\theta \in \Theta_n} d_n(\theta, \Gamma^k \theta).$$

[6] See Lemma A.1 for the formal statement and its proof.

[7] This is a widely used class of sieve spaces that includes polynomials, trigonometric functions, splines, and orthogonal wavelets. See, e.g., Chen (2007, page 5570).

In certain settings, the above minimization problem might only be feasible up to a certain residual term that will vanish as n diverges to infinity. This is precisely the case described in Assumption A.3. This progression naturally leads to the definition of our SVFI approximation.

Definition 3.3 (Sieve Value Function Approximation)**.** *Assume Assumption A.3. Then the SVFI approximation of V is any $\hat{\theta}_n \in \Theta_n$ that satisfies:*

$$d_n(\hat{\theta}_n, \Gamma^k \hat{\theta}_n) \leq \inf_{\theta \in \Theta_n} d_n(\theta, \Gamma^k \theta) + \eta_{2,n}. \tag{3.7}$$

where $\eta_{2,n} = O_p(\gamma_{2,n})$ with $\gamma_{2,n} = o(1)$.

3.5 Properties of the approximation

All of the findings of this section are corollaries of the following lemma.

Lemma 3.1. *Assume Assumptions A.1-A.4. Then, the SVFI approximation satisfies:*

$$d(\hat{\theta}_n, V) \leq \frac{(1 + K_2 K_1^{-1})\eta_{1,n} + K_2 \eta_{2,n} + K_1^{-1} K_2 (1 + \beta^k)\eta_{3,n}(V)}{1 - \beta^k}, \tag{3.8}$$

where $V \in \mathcal{F}$ is the unique fixed point of Γ in $(\mathcal{F}, d)$.

Lemma 3.1 is the key result to establish the consistency of the SVFI approximation, derive its rate of convergence, and investigate the finite sample properties of its approximation error. The following result establishes the consistency and the rate of convergence of the approximation.

Theorem 3.1. *Assume Assumptions A.1-A.4. Then, the SVFI approximation satisfies:*

$$d(\hat{\theta}_n, V) = O_p(\max\{\gamma_{1,n}, \gamma_{2,n}, \gamma_{3,n}(V)\}),$$

where $\max\{\gamma_{1,n}, \gamma_{2,n}, \gamma_{3,n}(V)\} = o(1)$ as $n \to \infty$, and $V \in \mathcal{F}$ is the unique fixed point of Γ in $(\mathcal{F}, d)$. This implies that the SVFI approximation:

1. *is consistent approximation of V, i.e., $d(\hat{\theta}_n, V) = o_p(1)$, as $n \to \infty$.*
2. *converges to V at a rate of $\min\{\gamma_{1,n}^{-1}, \gamma_{2,n}^{-1}, \gamma_{3,n}^{-1}(V)\}$.*

Theorem 3.1 implies that the rate of convergence of the approximation depends on the rate at which three errors converge to zero. These errors are:

1. The error of approximating the metric d with the approximate pseudo-metric d_n, denoted by $\eta_{1,n}$, which converges to zero at a rate of $\gamma_{1,n}^{-1}$.
2. The error of approximation when minimizing the objective function $d_n(\theta, \Gamma^k \theta)$, denoted by $\eta_{2,n}$, which converges to zero at a rate of $\gamma_{2,n}^{-1}$.

Table 1: Value of constants associated to the upper bound on the error of approximation for different number of iterations.[a]

	Number of iterations: k						
	1	2	3	4	5	6	7
$1/(1-\beta^k)$	10	5.26	3.69	2.91	2.44	2.13	1.92
$(1+\beta^k)/(1-\beta^k)$	19	9.53	6.38	4.82	3.88	3.27	2.83

[a] The discount factor β is set to 0.9.

3. The error of approximating the value function $V \in \mathcal{F}$ with an element in the sieve space Θ_n, denoted by $\eta_{3,n}(V)$, which converges to zero at a rate of $\gamma_{3,n}^{-1}(V)$.

The slowest of these three rates determines the rate of convergence of the approximation. The motivation for introducing sieve approximations was the fact that working with the original space $\mathcal{F}$ was computationally infeasible, i.e., the third source of error $\eta_{3,n}(V)$ cannot be avoided. However, it might be possible to avoid the other sources of error. In other words, it might be possible to use $d_n = d$, leading to $\gamma_{1,n} = 0$, or it might be possible to solve the minimization of $d_n(\theta, \Gamma^k\theta)$ exactly, leading to $\gamma_{2,n} = 0$. If so, then, the convergence rate of this source of error is infinity, effectively disappearing from the expression for the rate of convergence $\min\{\gamma_{1,n}^{-1}, \gamma_{2,n}^{-1}, \gamma_{3,n}^{-1}(V)\}$.

It is interesting to notice that the findings in Theorem 3.1 do not depend on the number of contraction mapping iterations k in Assumption A.3 (in particular, they hold even if $k = 1$). The choice of k affects several constants associated with the rates of convergence, which are "hidden" in the O_p notation. While these constants are not relevant for the asymptotic results, they can be very relevant for finite values of the computational power n. The right tool for this analysis is Eq. (3.8) in Lemma 3.1, which provides a concrete upper bound on the error of approximation that can be used to study the effect of changes in k that is valid for any value of n. This result reveals that the three sources of error are each associated with a constant that depend on k. In particular, the error terms $\eta_{1,n}$, $\eta_{2,n}$, and $\eta_{3,n}$ are associated to the constants $(1 + K_2K_1^{-1})/(1-\beta^k)$, $K_2/(1-\beta^k)$, and $K_1^{-1}K_2(1+\beta^k)/(1-\beta^k)$, respectively. A corollary of this is that, *ceteris paribus*, an increase in the number of value function iterations k reduces the value of the upper bound of the error of approximation. In particular, Table 1 illustrates that there are significant gains in precision from raising the value of k when the discount factor is $\beta = 0.9$. For example, changing the number of iterations from $k = 1$ to $k = 2$ reduces the (upper bound on the) error of approximation by approximately 50%.

We illustrate the tradeoffs associated with increasing the number of contraction mapping iterations k in the Monte Carlo analysis in Section 6.

4 Approximation in finite horizon dynamic problems

This section describes the application of non-parametric sieve methods to approximate the value functions in Eq. (2.1) when the dynamic problem has a finite horizon (i.e. $T < \infty$). The distinctive feature of the finite horizon problem is that the value function of the problem is no longer stationary (i.e. $V_t(\cdot)$ depends on the time index). As time passes and the terminal period T approaches, the agent's value function changes.

Using backward induction, the value function in any given period can be expressed the optimized choice between instantaneous utility and the discounter value for the immediately proceeding period. By repeating the arguments used to develop the unified formulation used in the infinite horizon case, we use s to denote the current value of the state, s' to denote the future value of the state, S to denote the state space, a to denote the action chosen by the agent, $A(s)$ to denote the space of actions available to the agent when the state is s, and F to be a known functional of $V_{t+1}(\cdot)$ and s' that satisfies certain known properties. Based on this notation, in every period $t = 1, \ldots, T$, the agent solves an optimization problem characterized by the following value function:

$$V_t(s) = \max_{a \in A(s)} \{u(s,a) + \beta E(F(V_{t+1}, s')|s,a)\}, \ \forall s \in S, \tag{4.1}$$

where the variables and functions are defined as in Eq. (3.1) and:

$$V_{T+1}(s) = 0, \ \forall s \in S. \tag{4.2}$$

Using the notation developed in Eq. 3.3, the sequence of value functions $\{V_t\}_{t=1}^T$ can be defined as follows: $V_T = \Gamma V_{T+1}$ with a zero terminal value, i.e., $V_{T+1}(s) = 0$, $\forall s \in S$.[8]

The approximation procedure developed for the infinite horizon problem must be modified to accommodate several distinct features of the finite horizon setup. First, the finite horizon problem requires an approximation for the value function at each point in time; the infinite horizon problem is stationary, i.e., the agent solves the same problem every period, and thus only requires the approximation of a single value function V. Second, the approximation procedure developed for the infinite horizon problem required the value function V to be a fixed point in a contraction mapping. Clearly, this will not be true for the non-stationary finite horizon problem. As time progresses, the last period of the game approaches and this affects the value of participating in the game. Thus the value functions of the finite horizon problem are not fixed points to any mapping, but are instead a finite sequence of functions that are sequentially related.

With enough computational power, the set of functions $\{V_t\}_{t=1}^T$ could be computed exactly using backward induction. Nevertheless, for economic models with large state spaces, the exact

[8] It should be noted that each of the objects in Eq. (4.1), i.e., the set of possible actions $A(\cdot)$, the period utility function $u(\cdot)$, the expectation operator $E(\cdot)$, and the functional $F(\cdot)$, could be allowed to be time-specific without affecting any of the theoretical results to follow. We opted to keep these elements time invariant to simplify the exposition and to relate them easily to elements in Eq. (3.1).

implementation of backward induction might be too computationally challenging or even impossible. The objective of this section is to propose a sieve value function approximation for such settings. For a given computational power (i.e. for a given sieve space Θ_n), our approximation method produces a sequence of approximations: $\{\hat{\theta}_{n,t}\}_{t=1}^{T}$ where, for all $t = 1, \ldots, T$, $\hat{\theta}_{n,t}$ is the sieve value function approximation for V_t.

We consider two approximation procedures. The first is, essentially, the sieve approximation version of a traditional backward induction. The value function is first approximated for the last period and this approximate function is used to solve for the (approximate) value function for the previous period. Continuing to work backwards yields an approximate value function for each period. Implementing this procedure requires using computational routines that are specifically tailored for the finite horizon setup.

The second procedure entails expanding the state space to include time as a state variable. To the best of our knowledge, this procedure is novel to our paper. While this procedure is less intuitive than backward induction, it has the advantage of being implemented with the same computational routines developed above for the infinite horizon problem.

4.1 Approximation using backward induction

The computation of the sequence of value functions $\{V_t\}_{t=1}^{T}$ by backward induction is well understood and requires no further discussion. This section proposes an approximation to these value functions using sieves. The approximation requires the following assumptions. By repeating previous arguments, it is easy to see that these assumptions are satisfied in dynamic decision problems that have a very large but finite state space.

Assumption B.1. $(\mathcal{F}, d)$ is a complete metric space of functions that map S onto $\mathbb{R}$, where d is the sup-norm metric, i.e.,

$$d(f_1, f_2) = \sup_{s \in S} ||f_1(s) - f_2(s)||, \ \forall f_1, f_2 \in \mathcal{F}.$$

Assumption B.2. For any $n \in \mathbb{N}$, d_n is a pseudo-metric in $(\mathcal{F}, d)$ such that $\exists K_1, K_2 > 0$,

$$K_1 \ d_n(f_1, f_2) - \lambda_{1,n} \leq d(f_1, f_2) \leq K_2 \ d_n(f_1, f_2) + \lambda_{1,n},$$

where $\lambda_{1,n} = O_p(\upsilon_{1,n})$ and $\upsilon_{1,n} = o(1)$, uniformly in $f_1, f_2 \in \mathcal{F}$.

Assumption B.3. For any $f \in \mathcal{F}$, we can find $\theta_n \in \Theta_n$ that satisfies:

$$d_n(\theta_n, f) \leq \inf_{\theta \in \Theta_n} d_n(\theta, f) + \lambda_{2,n}.$$

where $\lambda_{2,n} = O_p(\upsilon_{2,n})$ and $\upsilon_{2,n} = o(1)$, uniformly in $f \in \mathcal{F}$.

Assumption B.4. For any $f \in \mathcal{F}$:

$$\inf_{\theta \in \Theta_n} d(\theta, f) = \lambda_{3,n},$$

where $\lambda_{3,n} = O_p(\upsilon_{3,n})$ and $\upsilon_{3,n} = o(1)$, uniformly in $f \in \mathcal{F}$.

Assumption B.5. For all $f_1, f_2 \in \mathcal{F}$, $a \in A(s)$, and $s \in S$:

$$E[F(f_1, s') - F(f_2, s')|s, a] \leq d(f_1, f_2).$$

We now briefly comment on each of the assumptions with focus on the differences with Assumptions A.1-A.4. With respect to Assumption A.1, Assumption B.1 eliminates the contraction mapping requirement with the requirement that d is the sup-norm metric. As we have already explained, these two assumptions are not that different, as the mapping Γ in Eq. (3.3) can be shown to be a contraction mapping with respect to the sup-norm metric but not with respect to other metrics. Assumption B.2 is identical to Assumption A.2. Assumption B.3 is very similar to Assumption A.3. The differences between the two are the following. First, Assumption A.3 assumed that one could (approximately) minimize a specific objective function within the sieve space Θ_n, whereas Assumption B.3 assumes that one can (approximately) find the best approximation within the sieve space Θ_n for any function in $\mathcal{F}$. Second, in accordance to the first point, Assumption B.3 requires that the error of minimization to converge to zero (in probability) uniformly in $f \in \mathcal{F}$. Assumption B.4 strengthens Assumption A.4 as it requires that $\inf_{\theta \in \Theta_n} d(\theta, f)$ converges to zero (in probability) uniformly in $f \in \mathcal{F}$. In other words, instead of requiring a vanishing error of approximation of any particular function, we require that the a vanishing error of approximation for the *worst* function in the class of functions $\mathcal{F}$. For references on these stronger results see, e.g., chapter 8 in Lorentz (1966). Finally, Assumption B.5 is a mild assumption about the properties of the mapping F. In particular, Lemma A.2 verifies that this assumption holds for all possible the formulations of the problem.

The approximation considered in this section is defined as follows:

Definition 4.1 (Sieve Value Function Approximation)**.** *Assume Assumption B.3. Then the approximation of* $\{V_t\}_{t=1}^T$ *is* $\{\hat{\theta}_{n,t}\}_{t=1}^T$ *constructed in the following iterative manner. For* $z = 1, \ldots, T$*, let* $t = T + 1 - z$ *and complete the following steps:*

1. *Define* $\hat{\theta}^*_{t,n} : S \to \mathbb{R}$ *as follows:*

$$\hat{\theta}^*_{t,n}(s) \equiv [\Gamma \hat{\theta}_{t+1,n}](s) = \max_{a \in A(s)} \{u(s,a) + \beta E(F(\hat{\theta}_{t+1,n}, s')|s, a)\}, \ \forall s \in S,$$

where either: $t = T$ *and* $\hat{\theta}_{t+1,n}(s) \equiv V_{T+1}(s) = 0 \ \forall s \in S$*, or* $t < T$ *and* $\hat{\theta}_{t+1,n}$ *has been defined in a previous iteration of the algorithm.*

2. *Define* $\hat{\theta}_{t,n} : S \to \mathbb{R}$ *to be any* $\hat{\theta}_{t,n} \in \Theta_n$ *that satisfies:*

$$d_n(\hat{\theta}_{t,n}, \hat{\theta}^*_{t,n}) \leq \inf_{\theta \in \Theta_n} d_n(\theta, \hat{\theta}^*_{t,n}) + \lambda_{2,n},$$

where $\lambda_{2,n} = O_p(\upsilon_{2,n})$ *and* $\upsilon_{2,n} = o(1)$.

It is evident from the description of the procedure that this method implements the traditional backward induction procedure using sieve approximations, i.e., it performs an approximation of the value function of the terminal period and uses the approximation of the value function in a given period to conduct an approximation for the value function in the immediately preceding period. The following theorem establishes the asymptotic properties of the approximation.

Theorem 4.1. *Assume Assumptions B.1-B.5. Then, the approximation satisfies:*

$$\max_{t=1,\ldots,T} d(\hat{\theta}_{t,n}, V_t) = O_p(\max\{\upsilon_{1,n}, \upsilon_{2,n}, \upsilon_{3,n}\}),$$

where $\max\{\upsilon_{1,n}, \upsilon_{2,n}, \upsilon_{3,n}\} = o(1)$ *as* $n \to \infty$. *This implies that for all* $t = 1, \ldots, T$, *the sieve value function approximation* $\hat{\theta}_{t,n}$*:*

1. *is a consistent approximation of* V_t, *i.e.,* $d(\hat{\theta}_{t,n}, V_t) = o_p(1)$, *as* $n \to \infty$.

2. *converges to* V_t *at a rate of* $\min\{\upsilon_{1,n}^{-1}, \upsilon_{2,n}^{-1}, \upsilon_{3,n}^{-1}\}$.

As in Theorem 3.1, Theorem 4.1 indicates that the rate of convergence of the approximation depends on the rate at which three errors converge to zero. The slowest of these three rates determines the rate of convergence of the approximation.

4.2 Approximation using time as a state variable

The approximation considered in this section entails considering the time dimension as part of the state of the problem. In some sense, it may seem counterintuitive to "increase" the state space for a problem in which the size of space was already deemed to large to compute directly. However, as we demonstrate in this section, the approximation is computationally feasible and can be implemented using the exact same computational tools as in the infinite horizon case.

Consider the state space that results from the cartesian product of the (time invariant) state space S with the time dimension $\{1, \ldots, T+1\}$, i.e., $\tilde{S} = S \times \{1, \ldots, T+1\}$. Throughout this section, we superscript with the symbol $\sim$ to denote objects in the new state space that includes the time dimension. For example, the new metric space is denoted by $(\tilde{\mathcal{F}}, \tilde{d})$, where $\tilde{\mathcal{F}}$ denotes a set functions from $\tilde{S}$ onto $\mathbb{R}$ and $\tilde{d}$ denotes the corresponding norm in this space. In this enlarged state space, the sequence of value functions $\{V_t\}_{t=1}^{T+1}$ defined by Eqs. (4.1) and (4.2) can be equivalently re-written as follows:

$$V(s,t) \equiv V_t(s).$$

In the state space $\tilde{S}$ we can define an analogue to the function Γ, which we denote by $\tilde{\Gamma}$ and define as follows:

$$[\tilde{\Gamma}\theta](s,t) \equiv \sup_{a \in A(s)} \left\{ u(s,a) + \beta E(F(\theta,(s',t+1))|(s,t),a) \right\} \times 1[t < T+1], \ \forall (s,t) \in \tilde{S}. \tag{4.3}$$

In order to conduct the approximation in the state space $\tilde{S}$, we use a sequence of sieve spaces, denoted by $\{\tilde{\Theta}_n\}_{n \geq 1}$, where each $\tilde{\Theta}_n$ is a space of (simple) functions that map $\tilde{S}$ onto $\mathbb{R}$. We consider the following assumptions.

Assumption B.6. $(\tilde{\mathcal{F}}, \tilde{d})$ is a complete metric space of functions that map $\tilde{S}$ onto $\mathbb{R}$, where $\tilde{d}$ is the sup-norm metric, i.e.,

$$\tilde{d}(f_1, f_2) = \sup_{(s,t) \in \tilde{S}} ||f_1(s,t) - f_2(s,t)||, \ \forall f_1, f_2 \in \tilde{\mathcal{F}}.$$

Assumption B.7. For any $n \in \mathbb{N}$, $\tilde{d}_n$ is a pseudo-metric in $(\tilde{\mathcal{F}}, \tilde{d})$ such that $\exists K_1, K_2 > 0$,

$$K_1 \tilde{d}_n(f_1, f_2) - \lambda_{1,n} \leq \tilde{d}(f_1, f_2) \leq K_2 \tilde{d}_n(f_1, f_2) + \lambda_{1,n},$$

where $\lambda_{1,n} = O_p(\upsilon_{1,n})$ and $\upsilon_{1,n} = o(1)$, uniformly in $f_1, f_2 \in \tilde{\mathcal{F}}$.

Assumption B.8. For some $k \in \mathbb{N}$, we can find $\theta_n \in \Theta_n$ that satisfies:

$$\tilde{d}_n(\theta_n, \tilde{\Gamma}^k \theta_n) \leq \inf_{\theta \in \tilde{\Theta}_n} \tilde{d}_n(\theta, \tilde{\Gamma}^k \theta) + \lambda_{2,n},$$

where $\tilde{\Gamma}^k$ is the k^{th} iteration of $\tilde{\Gamma}$, $\lambda_{2,n} = O_p(\upsilon_{2,n})$, and $\upsilon_{2,n} = o(1)$.

Assumption B.9. For any $f \in \tilde{\mathcal{F}}$:

$$\inf_{\theta \in \tilde{\Theta}_n} \tilde{d}(\theta, f) = \lambda_{3,n}(f),$$

where $\lambda_{3,n}(f) = O_p(\upsilon_{3,n}(f))$ and $\upsilon_{3,n}(f) = o(1)$.

Assumption B.10. For all $f_1, f_2 \in \tilde{\mathcal{F}}$, $a \in A(s)$, and $(s,t) \in \tilde{S}$:

$$E[F(f_1,(s',t+1)) - F(f_2,(s',t+1))|(s,t),a] \leq d(f_1, f_2).$$

With the exception of the fact that the state space has been enriched with the time index, Assumptions B.6-B.10 are analogous to assumptions that have already been discussed in the paper. On the one hand, Assumptions B.6, B.7, and B.10 are analogous to Assumptions B.1, B.2, and

B.5 used to consider the sieve approximation to the backward induction solution in finite horizon problems. On the other hand, Assumptions B.8 and B.9 are analogous to Assumptions A.3 and A.4 used to consider the SVFI approximation in infinite horizon problems.

In the context of the SVFI approximation in infinite horizon problems, we provided an iterative algorithm to (approximately) minimizing the objective function. A similar iterative procedure can be developed in the present context.

Algorithm 4.1. *Let $\{\varepsilon_n\}_{n\geq 1}$ be a tolerance sequence that satisfies $\varepsilon_n = o(1)$. For an arbitrary initial function $f \in \tilde{\Theta}_n$, consider the following iterative procedure:*

1. *Given f, choose a function $\theta_m \in \tilde{\Theta}_n$ such that:*

$$\theta_m = \inf_{\theta \in \tilde{\Theta}_n} \tilde{d}_n(\theta, \tilde{\Gamma}^k f).$$

2. *If $\max\{\tilde{d}_n(\theta_m, \tilde{\Gamma}^k f), \tilde{d}_n(\theta_m, f)\} \leq \varepsilon_n$, then stop the algorithm and define $\theta_n \equiv \theta_m$. Otherwise, set $f = \theta_m$ and return to step 1.*

If the algorithm converges, then it can be shown that the resulting $\theta_n \in \tilde{\Theta}_n$ satisfies Eq. (3.5) with $\lambda_{2,n} = O(\max\{\varepsilon_n, \lambda_{1,n}\})$.[9]

The following result is the key to the asymptotic findings of this section.

Lemma 4.1. *Assume Assumptions B.6 and B.10. Let $V \in \tilde{\mathcal{F}}$ be the function defined by: $V(t,s) \equiv V_t(s)$ for all $S \times \{1, \ldots, T+1\}$. Then:*

1. *$\tilde{\Gamma}$ is a contraction mapping with modulus β on $(\tilde{\mathcal{F}}, \tilde{d})$.*

2. *V is the unique fixed point of the contraction mapping $\tilde{\Gamma}$, i.e., $\tilde{\Gamma} V = V$.*

In the context of infinite horizon problems, Assumption A.1 indicated that the value function was the unique fixed point of a certain contraction mapping. This result was the key to proposing the SVFI approximation method. Lemma 4.1 indicates that an analogous result holds for in the context of finite horizon problem. In fact, if we combine this result with the remaining assumptions, the current setup satisfies all of the conditions required for the SVFI approximation. As a consequence, an analogous approximation to the the SVFI approximation will have the same asymptotic properties, i.e, consistency and rates of convergence. This analogous approximation is defined next.

Definition 4.2 (Sieve Value Function Approximation)**.** *Assume Assumption B.8. Then the sieve approximation of $\{V_t\}_{t=1}^T$ is $\{\hat{\theta}_{t,n}\}_{t=1}^T$ where, for every $(s,t) \in S \times \{1, \ldots, T\}$, $\hat{\theta}_{t,n}(s) \equiv \hat{\theta}_n(s,t)$ and $\hat{\theta}_n \in \tilde{\Theta}_n$ is any function that satisfies:*

$$\tilde{d}_n(\hat{\theta}_n, \tilde{\Gamma}^k \hat{\theta}_n) \leq \inf_{\theta \in \tilde{\Theta}_n} \tilde{d}_n(\theta, \tilde{\Gamma}^k \theta) + \lambda_{2,n}.$$

[9] This can be shown using a very similar argument to the one used for Algorithm 3.1 in Lemma A.1. The argument requires that $\tilde{\Gamma}$ is a contraction mapping with modulus β on $(\tilde{\mathcal{F}}, \tilde{d})$, which shown in Lemma 4.1.

where $\lambda_{2,n} = O_p(\upsilon_{2,n})$ *and* $\upsilon_{2,n} = o(1)$.

Based on the previous discussion, the following result is a simple corollary of Theorem 3.1 and Lemma 4.1.

Theorem 4.2. *Assume Assumptions B.6-B.10. Then, the function* $\hat{\theta}_n \in \tilde{\Theta}_n$ *in Definition 4.2 satisfies:*

$$\tilde{d}(\hat{\theta}_n, V) = O_p(\max\{\upsilon_{1,n}, \upsilon_{2,n}, \upsilon_{3,n}(V)\}),$$

where $\max\{\upsilon_{1,n}, \upsilon_{2,n}, \upsilon_{3,n}(V)\} = o(1)$ *as* $n \to \infty$ *and* $V \in \tilde{\mathcal{F}}$ *is the function defined by:* $V(t,s) \equiv V_t(s)$ *for all* $S \times \{1, \ldots, T+1\}$. *This implies that for all* $t = 1, \ldots, T$, *the sieve approximation* $\hat{\theta}_{t,n}$*:*

1. *is a consistent approximation of* V_t, *i.e.,* $\sup_{s \in S} |\hat{\theta}_{t,n}(s) - V_t(s)| = o_p(1)$, *as* $n \to \infty$.
2. *converges to* V_t *at a rate of* $\min\{\upsilon_{1,n}^{-1}, \upsilon_{2,n}^{-1}, \upsilon_{3,n}^{-1}(V)\}$.

As in Theorems 3.1 and 4.1, Theorem 4.2 indicates that the rate of convergence of the approximation depends on the rate at which three errors converge to zero. Once again, the slowest of these three rates determines the rate of convergence of the approximation.

5 Estimation

The results to this point in the paper characterize the (approximate) computation of the value function (and associated object of interests) for a known vector parameters π that characterize the agent's dynamic decision problem. In this section, we now consider the problem of estimating π in a parameter space Π when the researcher has data on dynamic decisions and, again, the state space is too large to permit the direct computation of V for a given value of π.[10] As before, the associated value function V incorporates all the information that is relevant to the decision problem and depends on the parameter π, i.e., $V(\cdot|\pi)$. In this setup, the approximation problem of previous sections entails the approximation of the value function V for a particular value of the parameter $\pi \in \Pi$. Let π^* denote the true parameter value, i.e., $V(\cdot) \equiv V(\cdot|\pi^*)$.

For concreteness, consider an agent solving the value function in Eq. (3.1) with:

$$AS \equiv \{(a,s) : a \in A(s) \text{ and } \forall s \in S\},$$

instantaneous utility function given by:

$$u(s,a) = u(s,a|\pi_1), \ \forall (a,s) \in AS,$$

[10] Throughout this section, we pretend that the dynamic problem we refer to is the infinite horizon single agent decision problem. Nevertheless, by making slight notational changes, the results of this section can also be applied to a finite horizon problem.

transition probabilities given by:

$$dP(\varepsilon', x'|x, a) = dP(\varepsilon', x'|x, a, \pi_2), \ \forall (a, s) \in AS,$$

and a discount factor $\beta = \pi_3$. Set $\pi \equiv (\pi_1, \pi_2, \pi_3) \in \Pi$. For each parameter value $\pi \in \Pi$, the corresponding value function is denoted $V(\cdot|\pi)$.

The SVFI approximation procedure described in Definition 3.3 provides a method for approximating the value function for any given set of primitives (i.e., choice set, utility function, discount factor, and transition probabilities). In this setting, these primitives are functions of the unknown parameter value π. To be consistent with this interpretation, we denote: $\theta_n \equiv \theta_n(\cdot|\pi)$.

If the value function (or some function derived from it) were observed for an arbitrary set of values, then the SVFI approximation procedure could be used to approximate it. Furthermore, the consistency result in Theorem 3.1 suggests that the parameter π could be estimated as follows:

$$\hat{\pi}_n = \arg\min_{\pi \in \Pi} Q(\theta_n(\cdot|\pi)), \tag{5.1}$$

where $Q(\theta_n(\cdot|\pi))$ is an appropriately chosen function that measures the distance between $\theta_n(\cdot|\pi)$ and the value function V. For instance, Q could be the following function:

$$Q(\theta_n(\cdot|\pi)) = \int (V(s) - \theta_n(s|\pi))^2 \ d\mu(s),$$

where μ is any arbitrary positive measure over S. In practice, however, the estimator in Eq. (5.1) is unfeasible because the value function is not observed by the researcher and, consequently, Q is unknown.

In practice, the value function, or some feature derived from it (e.g., CCPs), can be estimated from the data. With some abuse of notation, let V_I denote the estimated value function using a sample of size I (observations are indexed $i = 1, \ldots, I$) and let Q_I denote the function that measures the distance between $\theta_n(\cdot|\pi)$ and V_I. Our previous discussion suggests that the parameter π could be estimated according to the following definition.

Definition 5.1 (Estimator based on the Sieve Value Function Approximation)**.** *Let $Q_I : \Pi \to \mathbb{R}_+$ be the function of the data that measures the distance between V_I and $\theta_n(\cdot|\pi)$ for any $\pi \in \Pi$ and let $n = n(I)$. The estimator of the true parameter value π^*, denoted $\hat{\pi}_I$, satisfies:*

$$Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) \leq \inf_{\pi \in \Pi} Q_I(\hat{\theta}_n(\cdot|\pi)) + o_p(1), \ \text{as } I \to \infty. \tag{5.2}$$

In order to clarify the structure of the problem, we consider an illustrative example.

Example 5.1. *In this case, we estimate the CCPs from a sample of I observed choices. In partic-*

ular, if the set S is finite, then the following estimator:

$$\hat{P}_I = \{\hat{P}_I(a|s),\ \forall (a,s) \in AS\},$$

where:

$$\hat{P}_I(a|s) = \sum_{i=1}^{I} 1[a_i = a, s_i = s] / \sum_{i=1}^{I} 1[s_i = s],$$

is a $\sqrt{I}$-consistent estimator of the CCPs.

By definition, the CCPs can be derived from the value function. Let $J(\cdot|\pi) : \mathcal{F} \to [0,1]^{\#AS}$ be the mapping between the value function and the CCPs, i.e.,

$$J(V|\pi) = \{P(a|s),\ \forall (a,s) \in AS\},$$

where the conditioning on π indicates that the mapping itself could depend on the parameter value.[11] *This discussion suggests the estimation of π with $\hat{\pi}_I$ as in Eq. (5.2), where the measure μ is set to be the empirical measure, i.e.,*

$$Q_I(\theta_n(\cdot|\pi)) = \int (\hat{P}_I - J(\theta_n(\pi)|\pi))^2 \, d\hat{P}_I, \tag{5.3}$$

$$= I^{-1} \sum_{i=1}^{I} 1[(a_i, s_i) = (a,s)](\hat{P}_I(a|s) - J_{(a,s)}(V|,\pi))^2. \tag{5.4}$$

In other words, we choose the parameter value that minimizes the integrated squared distance between the observed CCPs and the approximated CCPs, where the empirical measure is used as the measure of integration.

The objective of the rest of the section is to provide conditions under which $\hat{\pi}_I$ is consistent. To this end, we now provide a list of the assumptions that are exclusively used in this section.

Assumption C.1. $(\mathcal{F}_1, d_1)$ and $(\mathcal{F}_2, d_2)$ are metric spaces. The true value function $V = V(\cdot|\pi^*) \in \mathcal{F}_1$ and the true parameter value $\pi^* \in \Pi \subseteq \mathcal{F}_2$.

Assumption C.2. The SVFI approximation satisfies:

$$\sup_{\pi \in \Pi} \ d_1(\hat{\theta}_n(\cdot|\pi), V(\cdot|\pi)) = o_p(1), \text{ as } n \to \infty.$$

Assumption C.3. There is a function $Q(\cdot) : \mathcal{F}_1 \to \mathbb{R}$ such that:

a. $\hat{c}(n,I) = o_p(\delta(n,I))$ where $\hat{c}(n,I) \equiv \sup_{f \in \Theta_n} |Q_I(f) - Q(f)|$.

[11] The function J will depend on the specific formulation of the dynamic decision problem. See Appendix B for a description of each of the formulations and the definition of the function J in each case.

b. Q is uniformly continuous in $\mathcal{F}_1$ under d_1, i.e., for any $\delta > 0$, there exists $\varepsilon > 0$ such that $f_1, f_2 \in \mathcal{F}_1$ with $d_1(f_1, f_2) < \varepsilon$ implies $Q(f_1) - Q(f_2) \leq \delta$.

c. The function $Q(V(\cdot|\pi)) : \Pi \to \mathbb{R}$ is uniquely minimized at $\pi = \pi^*$.

Assumption C.4. For every $I \in \mathbb{N}$, $n = n(I)$ with $n(I) \to \infty$ and $\delta(n(I), I) \to 0$ as $I \to \infty$.

We now briefly comment on each of the assumptions. Assumption C.1 provides a label to relevant spaces, functions, and parameters. Assumption C.2 is a high level assumption which indicates that $\hat{\theta}_n(\cdot|\pi)$ is a consistent approximation of $V(\cdot|\pi)$, uniform over $\pi \in \Pi$. By repeating the arguments in Theorem 3.1, it is not hard to see that Assumption C.2 holds as a result of using the SVFI approximation procedure described in Definition 3.3 under Assumptions A.1-A.4, with the exception that Assumption A.4 is strengthened to hold uniformly in $f \in \mathcal{F}$ (i.e., as in Assumption B.4). Assumption C.3 is similar to assumptions used in the literature on extremum estimators (see, e.g., Chen (2007, Theorem 3.1) for conditions pertaining to sieve estimators, and Amemiya (1985, Theorem 4.1.1) or McFadden and Newey (1994, Theorem 2.1) for finite dimensional estimators). Finally, Assumption C.4 describes conditions that restrict the relationship between the data sample size and the complexity of the sieve.

Under these assumptions, it is possible to obtain consistent estimators of the parameters of large state space dynamic decision problems by embedding SVFI methods in the estimation algorithm.

Theorem 5.1. *Assume Assumptions C.1-C.4. Then, the estimator in Eq. (5.1) is consistent, i.e.,*

$$d_2(\hat{\pi}_I, \pi^*) = o_p(1), \text{ as } I \to \infty.$$

6 Monte Carlo simulations

In this section we illustrate the small sample properties of sieve value function approximation in an infinite horizon single agent setting. We first demonstrate sieve approximation of the value function when the structural parameters are known. In the second part of this section, we show how sieves can be applied in the estimation of structural models. The Monte Carlo experiments are conducted in a framework similar to the bus engine replacement problem in Rust (1987). In our context though, the problem is modified so that rather than making the engine replacement decision of a single bus, the agent now endogenizes the purchasing and replacement decision over an entire fleet of buses, with the constraint that at most one bus can be replaced in any period. Allowing more buses increases the number of state variables, allowing us to study the approximation for both small and large scale problems. We consider a single bus setting where we can compare the sieve approximation to full solution methods and then study the performance of sieves in a more complex problem (10 buses) where full solution methods are infeasible.

Let J indicate the number of buses in the agent's fleet. The endogenous state vector x_t contains the current engine mileage for each bus j, $x_t(j)$. There are J endogenous state variables. In each

period the agent has the option of reseting any (but only one) of the J engine mileages to zero. The choice set is defined as $a_t = \{0, 1, \ldots, J\}$, where $a_t = 0$ represents replacing no buses and $a_t = j$ corresponds replacing the engine in bus j. For these examples we assume that if a bus engine is not replaced the engine mileage transitions as an exponential random variable that is the same for all buses.

$$x_{t+1}(j) = \begin{cases} x_t(j) + \zeta_t(j) & \text{, if } a_t \neq j \\ \zeta_t(j) & \text{, if } a_t = j \end{cases}$$

Where the probability density function of ζ is $P(\zeta|\lambda) = \lambda \exp(-\lambda\zeta)$.

Given the current mileage states x, the agent chooses to set at most one engine milage to zero in the current period. Contemporaneous utility is defined as $u(x, a_t) + \varepsilon_t(a_t)$, where ε is a random utility shock and mean flow utility is defined as

$$u(x_t, a_t) = \begin{cases} \alpha \sum_{j=1}^{J} x_t(j)/J & \text{, if } a_t = 0 \\ \alpha \sum_{j=1}^{J} x_t(j)/J - \alpha x_t(a_t)/J + RC & \text{, if } a_t = \{1, \ldots, J\} \end{cases}$$

RC is the replacement cost of a bus engine.

The agent's problem is to make an optimal sequence of purchasing decisions to maximize the expected discounted flow of future per-period pay-offs. With discount factor β, we can formulate the value function using the recursive Bellman representation:

$$V(x_t, \varepsilon_t) = \max_{a_t \in \{0,1,\ldots,J\}} \{u(x_t, a_t) + \varepsilon_t(a_t) + \beta E[V(x_{t+1}, \varepsilon_{t+1})|x_t, a_t]\}$$

Where the expectation is taken with respect to all future milage shocks ζ and utility shocks ϵ.

The choice specific value functions (excluding the shocks) for choice $a \in \{0, 1, \ldots, J\}$ are

$$v(x_t, a) = u(x_t, a) + \beta E[V(x_{t+1}, \varepsilon_{t+1})|x_t, a].$$

The agent's optimal decision in period t is then to choose a_t^* such that:

$$a_t^* = \arg \max_{a \in \{0,1,\ldots,J\}} \{v(x_t, a) + \varepsilon_t(a)\}.$$

6.1 Approximation

Let θ_n denote the sieve approximation of the ex-ante value function, i.e. $\theta_n(x_t) \approx E[\max_{a \in \{0,1,\ldots,J\}}(v(x_t, a) + \varepsilon_t(a))|x_t]$ where the expectation is over the random pay-off shock ε. Assuming the utility shocks are distributed type-I extreme value, we have a known closed form representation of the expectation:

$$E[V(x_t, \varepsilon_t)|x_t] = \ln\left(\sum\nolimits_{a \in \{0,1,\ldots,J\}} \exp(v(x_t, a))\right) + \gamma, \tag{6.1}$$

Where γ is Euler's constant.

Therefore, for a chosen sieve space Θ_n, we seek an approximation of the expression in Eq. (6.1), $\theta_n \in \Theta_n$, such that:

$$\begin{aligned}
\theta_n(x_t) &\approx \ln\left(\sum\nolimits_{a\in\{0,1,\ldots,J\}} \exp(v(x_t,a))\right) + \gamma \\
&= \ln\left(\sum\nolimits_{a\in\{0,1,\ldots,J\}} \exp\left(u(x_t,a) + \beta E\left[V(x_{t+1},\varepsilon_{t+1})|x_t,a\right]\right)\right) + \gamma \\
&\approx \ln\left(\sum\nolimits_{a\in\{0,1,\ldots,J\}} \exp\left(u(x_t,a) + \beta E\left[\theta_n(x_{t+1})|x_t,a\right]\right)\right) + \gamma. \qquad (6.2)
\end{aligned}$$

The right hand side of Eq. (6.2) represents one contraction on the approximation, θ_n, defined as $\Gamma^1\theta_n(x_{t+1}|x_t)$. Section 3.5 demonstrates that we can actually reduce the upper-bound of the approximation error by increasing the number of iterations of the contraction operator on the approximation. Rather than immediately plugging the approximation into the expected future value term, we could alternatively write this term using Eq. (6.1) as a function of next periods contemporaneous profit functions and a two-period away expected future value term. If we substitute the approximation into the two-period away expected future value function, this corresponds to two iterations of the contraction ($k = 2$). We could continue in this way for $k > 2$, though the computational burden increases exponentially with k.

We construct the sieve approximation of order n using ordinary polynomials of the state variables and their interactions, where n denotes the highest order of the polynomial, i.e. $n = 1$ includes only linear terms, $n = 2$ includes linear terms, the squares, and all two-way interactions, etc. Given a value of the state vector, x, our n^{th} order polynomial is generated by the function $\mathrm{W}_n(x)$. The sieve approximation is then,

$$\theta_n(x) = \mathrm{W}_n(x)\rho$$

Where ρ are the parameters that approximate the value function.

For any state variable x, the sieve approximation and its contraction will satisfy,

$$\mathrm{W}_n(x)\rho \approx \ln\left(\sum\nolimits_{a\in\{0,1,\ldots,J\}} \exp\left(u(x,a) + \beta E\left[\mathrm{W}_n(x')\rho|x,a\right]\right)\right) + \gamma \qquad (6.3)$$

A key convenience of constructing the approximation with a linear function is that the parameters ρ can be taken out of the expected value, with the expectation only applying to the state transitions, so we can replace $E[\mathrm{W}_n(x')\rho|x,a] = E[\mathrm{W}_n(x')|x,a]\rho$.

Given a sample of state vectors x_s for $s = 1,\ldots,S$, the parameters ρ in (6.3) are found through the iterative Algorithm 3.1 using the sum of squared errors as the distance function. At iteration

Table 2: Sieve Approximation ($J = 1$) [a]

Max. Poly	Within-Approximation Sample Fit			Comparison to Full-Solution
	$\|\|\theta - \Gamma\theta\|\|_2$	$\|\|\theta - \Gamma^1\theta\|\|_\infty$	R-squared	$\max\{\|\Pr^{sieve} - \Pr^{fullsol}\|\}$
1^{st}	90.03315	2.11202	0.72411	0.43999
2^{nd}	32.81664	0.85165	0.88525	0.12709
3^{rd}	6.71985	0.23755	0.97456	0.03073
4^{th}	2.59063	0.07202	0.99010	0.00914
5^{th}	1.71290	0.05776	0.99345	0.01741
6^{th}	0.41342	0.01338	0.99842	0.00740
7^{th}	0.38959	0.01210	0.99851	0.00587

[a] Sieve approximation uses 10,000 draws on the interval (0,1). The parameters are set to $\alpha = -5$, $RC = -10$, $\lambda = 1/.05$ and $\beta = .99$

m, the parameters are updated for $m+1$ by solving:

$$\rho^{m+1} = \arg\max_{\rho} \sum_{s=1}^{S} \left[\mathrm{W}_n(x_s)\rho - \ln\left(\sum\nolimits_{a\in\{0,1,\ldots,J\}} \exp\left(u(x_s,a) + \beta E\left[\mathrm{W}_n(x_s')|x_s,a\right]\rho^m\right)\right) + \gamma \right]^2, \tag{6.4}$$

Repeating until convergence.

Let X represent a matrix whose s^{th} row contains $\mathrm{W}_n(x_s)$. Similarly, $\overline{X}_a$ is a matrix who's s^{th} row contains $E\left[\mathrm{W}_n(x_s')|x_s,a\right]$. The maximum of eq. (6.4) has a closed form solution, so the iterative procedure finds a fixed point to the equation

$$\rho^{m+1} = (X'X)^{-1}X'\left[\ln\left(\sum\nolimits_{a\in\{0,1,\ldots,J\}} \exp\left(u(X,a) + \beta\overline{X}_a\rho^m\right)\right) + \gamma\right] \tag{6.5}$$

The element inside the brackets on the right-hand side of eq. (6.5) can be abbreviated as $\Gamma^1\theta(\rho^m)$ since it represents one contraction on the sieve function with parameters ρ^m.

Equation (6.5) has two important features. First, most of it's components, including $(X'X)^{-1}X'$ and $\overline{X}_a$ can be computed outside of the algorithm and do not need to be recomputed at each iteration. Second, since the sieve is a simple linear function, the elements of $\overline{X}_a$, the expectations of the sieve given the state transitions, are simply the raw moments of the random variables. In many cases, these moments have a closed form expression or can be approximated to any precision with simulation methods, which holds true for both continuous or discrete state variables.[12]

[12] In our engine replacement problem, for example, if an element of the approximating function is the interaction of mileages between bus 1 and bus 2 (i.e. $x(1) \times x(2)$), and neither engine is replaced, then the expectation for the state transition next period is defined as,

$$E[(x(1)+\zeta(1))(x(2)+\zeta(2))] = x(1)x(2) + x(1)E[\zeta(2)] + x(2)E[\zeta(1)] + E[\zeta(1)\zeta(2)]$$

Given the assumption that ζ is an exponential random variable with parameter λ, $E[\zeta] = 1/\lambda$.

Table 2 illustrates the performance of the sieve approximation for the single bus case. In this example we consider up to a seventh order polynomial which, including the constant, contains eight terms. The benefit of looking at the single state variable case is that we can also solve the value function using other methods to compare to the results of the sieve approximation. Once we have the approximating function, we can also compute three measures of fit within the approximation sample to evaluate the performance of the sieve. These include,

1. Sum of squared errors: $||\theta - \Gamma^1\theta||_2$

$$\sum_{s=1}^{S} \Big(\theta(x_s) - \Gamma^1\theta(x'_s|x_s)\Big)^2$$

2. Supremum norm: $||\theta - \Gamma^1\theta||_\infty$

$$\max\{|\theta(x_s) - \Gamma^1\theta(x'_s|x_s)|\}_{s=1}^{S}$$

3. R-squared

$$1 - \frac{||\theta - \Gamma^1\theta||_2}{||\Gamma^1\theta - E(\Gamma^1\theta)||_2}$$

The column to the far right of Table 2 uses the sieve approximation to compute choice probabilities for a given state vector and compares them to the actual choice probabilities from a full solution method. This shows that for the low order polynomials, in particular $n = 1$, the sieve does a very poor job approximating the value function. This is also apparent looking at the within approximation fit, for example the R-squared of 0.72. However, as suggested by the theoretical results, by increasing the sieve complexity we approach a near perfect approximation to the value function, such that with a seventh order polynomial the *maximum* difference between the choice probability predicted by the sieve and the actual choice probability is close to one-half of one percent.

Table 3 displays the results of the sieve approximation method for a much larger problem containing 10 continuous state variables. Similar to the smaller problem, the measures of within approximation fit become successively better as the complexity of the sieve is increased. Although we have no method to compare the predicted probabilities with the true values as we could with the smaller model, the fact that we can achieve similar measures of fit within the sample gives some indication that we are close to the true value function.

6.2 Estimation

We now apply sieve value function approximation to estimate dynamic models, where we simultaneously solve for the structural parameters and the associated sieve value function approximation that maximizes the likelihood of observed choice data. We consider the single bus replacement case

Table 3: Sieve Approximation ($J = 10$) [a]

Max. Poly	Num. Parms.	Within-Approximation Sample Fit $\|\|\theta - \Gamma\theta\|\|_2$	$\|\|\theta - \Gamma^1\theta\|\|_\infty$	R-squared
1^{st}	11	101.06255	3.66225	0.92834
2^{nd}	66	7.00686	0.27396	0.99177
3^{rd}	286	2.04118	0.08259	0.99768
4^{th}	1001	0.64141	0.03279	0.99928
5^{th}	3003	0.34694	0.01685	0.99961

[a] Sieve approximation uses 10,000 draws on the interval (0,2). The parameters are set to $\alpha = -5$, $RC = -10$, $\lambda = 1/.05$ and $\beta = .99$

so that data can be generated from the true choice probabilities. In general, the data will contain fewer observations than are necessary to approximate the sieve function. This is not a problem because we are free to draw as many state vectors as we want from the state space to supplement the states that are actually observed. Since the observed sample contains the most relevant part of the state space, for example we will never observe buses with extraordinarily high mileages because they will have been replaced long before then, rather than randomly drawing from the state space for the approximation, we may be able to improve our approximation (and thus our structural parameter estimates) by focusing on the state space near the generated data. For these exercises we will draw 10,000 random points which will be perturbations around the observed data points.

Our model contains two unknown structural parameters, α and RC. Estimation is implemented by modifying the iterative algorithm used for approximation described in Section 3.3. The parameter of interest is denoted by $\pi \equiv (\alpha, RC)$. Let the observations be indexed by $i = 1, \ldots, I$, where I denotes the sample size, a_i denotes the observed choice for observation i, and x_i denotes the observed state for observation i.

The estimator of π is denoted by $\hat{\pi}_I$ and is computed according to the following procedure. First, we choose arbitrary initial values of $(\hat{\pi}_I^0, \hat{\theta}_n^0) \in \Pi \times \tilde{\Theta}_n$. Then, we set $m = 1$ and use the following iterative algorithm:

1. Choose $\hat{\pi}_I^m$ to maximize the (approximate) likelihood:

$$\hat{\pi}_I^m = \arg\max_{\pi \in \Pi} \sum_{i=1}^{I} \ln(\Pr(a_i|\pi, x_i, \hat{\theta}_n^{m-1}).$$

2. Choose $\hat{\theta}_n^m$ to solve the fixed point in eq. (6.5) taking $\hat{\pi}_I^m$ as given.

3. If $||\hat{\pi}_I^m - \hat{\pi}_I^{m-1}|| < 10^{-6}$, then stop the algorithm and set $\hat{\pi}_I = \hat{\pi}_I^m$. Otherwise, set $m = m + 1$ and return to the first step.

Table 4: Sieve Estimation $(J = 1)^a$

	α	RC
Parameters	-5.000	-10.000
Max. Poly		
1^{st}	-2.3369	-5.7587
	(0.1200)	(0.1551)
2^{nd}	-8.6813	-16.6561
	(1.1337)	(1.7981)
3^{rd}	-5.0001	-10.0168
	(0.2724)	(0.4787)
4^{th}	-5.0248	-10.0595
	(0.2803)	(0.5008)

[a] Results from 2000 replications, each with 5,000 observations. Initial values of structural parameters and sieve approximation were set to zero.

This algorithm parallels Aguirregabiria and Mira (2002) in the sense that it swaps the calculation of the value function to outside of the maximization of the likelihood over the structural parameters. The outer iterations of this procedure will converge linearly, however, this approach requires an inner iterative routine to maximize the structural utility parameters conditional on the approximation and also to resolve the approximating function conditional on the new structural parameters. Although this process is sped up by using the previous estimates as warm starts, these inner iterative routines are undesirable. A computationally efficient estimation algorithm is outlined in James (2012), which requires no inner iterative routines and whose outer iterations converge super-linearly.

Results of the estimation exercise for different values of sieve complexity are presented in Table 4. The results show that for the first and second order polynomial the parameter estimates are extremely biased, which is not surprising given that these functions performed very poorly in approximating the value function in Table 2. However this bias completely disappears for the third order polynomial and above.

7 Conclusion

This paper proposes a methodology to approximate the value function in single agent dynamic problems where a large state space makes value function iteration unfeasible. Our method is based on non-parametric sieve estimation, and we refer to it as Sieve Value Function Iteration (SVFI). We provide a formal framework to analyze the approximation error. In particular, we show that

the SVFI approximation converges to the value function as the complexity of the sieve increases, and we characterize the rate of this convergence. Furthermore, we provide a concrete upper bound on the error of approximation which can be used to analyze its contributing factors.

A Monte Carlo analysis reveals that the SVFI approximation is very successful in estimating the value function. These results suggest that our approximation can successfully be used to solve models that would otherwise be computationally infeasible, implying that these techniques may substantially broaden the class of models that can be solved and estimated.

Given the standard challenges with large state space problems, we expect SVFI to open up a wide variety of avenues of theoretical and empirical exploration of complex dynamic single agent and equilibrium problems. For example, in Arcidiacono, Bayer, Bugni, and James (2012), we consider sequential move dynamic games. Estimation of these games can be done via standard two-step procedures. Through the use of sieves, it is possible to calculate the conditional choice probabilities of the finite horizon game which has a unique equilibrium, the limit of which is also an equilibrium in the infinite horizon game. It is then possible to compare the conditional choice probabilities from the finite horizon equilibrium to those observed in the data, testing to see whether this equilibrium was played.

References

Aguirregabiria, V., and P. Mira (2002): "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70(2), 1519–1543.

Amemiya, T. (1985): *Advanced Econometrics*. Harvard University Press.

Arcidiacono, P., P. Bayer, F. Bugni, and J. James (2012): "Sieve Value Function Iteration for Large State Space Dynamic Games," Mimeo: Duke University and Federal Reserve Bank of Cleveland.

Bénitez-Silva, H., G. Hall, G. Hitsch, G. Pauletto, and J. Rust (2000): "A comparison of discrete and parametric approximation methods for continuous-state dynamic programming problems," Mimeo: Yale University, S.U.N.Y. Stony Brook and University Geneva.

Bertsekas, D. P. (2001): *Dynamic Programming and Optimal Control, 2nd edition*, vol. 2.

——— (2012): *Dynamic Programming and Optimal Control, 3nd edition*, vol. 2.

Bertsekas, D. P., and J. N. Tsitsiklis (1996): *Neuro-Dynamic Programming*, vol. 2.

Blackwell, D. (1965): "Discounted Dynamic Programming," *The Annals of Mathematical Statistics*, 36(1), 226–235.

Chen, X. (2007): "Large Sample Sieve Estimation of Semi-Nonparametric Models," *Handbook of Econometrics*, 6B, 5550–5588.

Crawford, G., and M. Shum (2005): "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73(4), 1137–1173.

Hendel, I., and A. Nevo (2006): "Measuring the Implications of Sales and Consumer Inventory Behavior," *Econometrica*, 74(6), 1637–1673.

James, J. (2012): Mimeo: Federal Reserve Bank of Cleveland.

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

——— (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

Lorentz, G. (1966): *Approximation of functions*. Holt.

McFadden, D., and W. Newey (1994): "Large sample estimation and hypothesis testing," vol. 4 of *Handbook of Econometrics*, pp. 2111–2245.

Powell, M. (1981): *Approximation theory and methods*. Cambridge University Press.

Powell, W. (2008): *Approximate Dynamic Programming: Breaking the Curses of Dimensionality*. Wiley Series in Probability and Statistics.

Royden, H. L. (1988): *Real Analysis*. Prentice-Hall.

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

——— (1988): "Maximum Likelihood Estimation of Discrete Control Processes," *SIAM J. Control and Optimization*, 26, 1006–1024.

Rust, J. (2000): "Parametric Policy Iteration: An Efficient Algorithm for Solving Multidimensional DP problems," Mimeo: University of Maryland.

Stokey, N., and R. Lucas (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press.

Sweeting, A. (2011): "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry," Mimeo: Duke University.

Appendix A Technical appendix

Lemma A.1. *Assume Assumptions A.1-A.2 and let $\theta_n \in \Theta_n$ be the result of a convergence in Algorithm 3.1. Then, θ_n satisfies Eq. (3.5) with $\eta_{2,n} = O(\max\{\varepsilon_n, \eta_{1,n}\})$.*

Proof. By definition, the algorithm stops when: $d(f, \theta_n) \leq \varepsilon_n$ and $d_n(\theta_n, \Gamma^k f) \leq \varepsilon_n$ for some $\theta_n, f \in \Theta_n$. Based on this, consider the following argument:

$$\begin{aligned}
d_n(\theta_n, \Gamma^k \theta_n) &\leq d_n(\theta_n, \Gamma^k f) + d_n(\Gamma^k f, \Gamma^k \theta_n) \\
&\leq d_n(\theta_n, \Gamma^k f) + K_1^{-1} d(\Gamma^k f, \Gamma^k \theta_n) + K_1^{-1}\eta_{1,n} \\
&\leq d_n(\theta_n, \Gamma^k f) + K_1^{-1}\beta^k d(f, \theta_n) + K_1^{-1}\eta_{1,n} \\
&\leq d_n(\theta_n, \Gamma^k f) + K_1^{-1}\beta^k K_2 d_n(f, \theta_n) + K_1^{-1}\beta^k K_2\eta_{1,n} + K_1^{-1}\eta_{1,n} \\
&\leq \varepsilon_n(1 + K_1^{-1}\beta^k K_2) + (K_1^{-1}\beta^k K_2 + K_1^{-1})\eta_{1,n} \\
&\leq \inf_{\theta \in \Theta_n} d_n(\theta, \Gamma^k \theta) + \eta_{2,n},
\end{aligned}$$

where $\eta_{2,n} \equiv \varepsilon_n(1 + K_1^{-1}\beta^k K_2) + (K_1^{-1}\beta^k K_2 + K_1^{-1})\eta_{1,n}$ and, thus, $\eta_{2,n} = O(\max\{\varepsilon_n, \eta_{1,n}\})$, as required. The first inequality follows from the triangular inequality (applied to the pseudo-metric d_n), the second and fourth inequalities follow from Assumption A.2, the third inequality follows from the fact that Γ is a contraction mapping with modulus β, the fifth inequality follows from the stopping rule in the algorithm, the final inequality follows from the definition of $\eta_{2,n}$ and the fact that the pseudo-metric d_n is positive. □

Proof of Lemma 3.1. We begin by showing that or any $\theta \in \mathcal{F}$ and $m \in \mathbb{N}$:

$$d(\theta, V) \leq d(\theta, \Gamma^m \theta)/(1 - \beta^m),$$

To see this, consider the following derivation:

$$\begin{aligned}
d(\theta, V) &\leq d(\theta, \Gamma^m \theta) + d(\Gamma^m \theta, \Gamma^m V) + d(\Gamma^m V, V), \\
&= d(\theta, \Gamma^m \theta) + d(\Gamma^m \theta, \Gamma^m V), \\
&\leq d(\theta, \Gamma^m \theta) + \beta^m \; d(\theta, V),
\end{aligned}$$

where the first inequality follows from the Triangle Inequality, the next equality follows from the fact that V is a fixed point, and the final inequality follows from the fact that Γ is a contraction mapping. Eq. (A) is a straight-forward consequence of this result.

Let $\hat{\theta}_n \in \Theta_n \subseteq \mathcal{F}$ be the SVFI approximation in Definition 3.3. On the one hand, consider the following derivation:

$$\begin{aligned}
d(\hat{\theta}_n, V)(1 - \beta^k) &\leq d(\hat{\theta}_n, \Gamma^k \hat{\theta}_n), \\
&\leq K_2 \; d_n(\hat{\theta}_n, \Gamma^k \hat{\theta}_n) + \eta_{1,n}, \\
&\leq K_2 \inf_{\theta \in \Theta_n} d_n(\theta, \Gamma^k \theta) + \eta_{1,n} + K_2\eta_{2,n}, \\
&\leq K_1^{-1} K_2 \inf_{\theta \in \Theta_n} d(\theta, \Gamma^k \theta) + (1 + K_2 K_1^{-1})\eta_{1,n} + K_2\eta_{2,n},
\end{aligned}$$

where the first inequality follows from Eq. (A), the second and fourth inequalities follow from Assumption A.2, and the third inequality follows from Eq. (3.7).

On the other hand, for any $\theta \in \mathcal{F}$, consider the following derivation:

$$\begin{aligned} d(\theta, \Gamma^k\theta) &\leq d(\theta, V) + d(V, \Gamma^k V) + d(\Gamma^k V, \Gamma^k\theta), \\ &= d(\theta, V) + d(\Gamma^k V, \Gamma^k\theta), \\ &\leq (1+\beta^k) d(\theta, V), \end{aligned}$$

where the first inequality follows from the triangle inequality, the next equality follows from the fact that V is a fixed point, and the final inequality follows from the fact that Γ is a contraction mapping in (d, Γ). If we take infimum of $\theta \in \Theta_n \subseteq \mathcal{F}$ on both sides:

$$\inf_{\theta\in\Theta_n} d(\theta, \Gamma^k\theta) \leq (1+\beta^k) \inf_{\theta\in\Theta_n} d(\theta, V) = (1+\beta^k)\eta_{3,n}(V).$$

The result follows directly from combining the previous results. □

Proof of Theorem 3.1. By combining Lemma 3.1 with Assumption A.4, it follows that:

$$\begin{aligned} d(\hat{\theta}_n, V) &\leq \left\{K_1^{-1}K_2(1+\beta^k)\eta_{3,n}(V) + (1+K_2K_1^{-1})\eta_{1,n} + K_2\eta_{2,n}\right\}(1-\beta^k)^{-1}, \\ &= O_p(\max\{\gamma_{1,n}, \gamma_{2,n}, \gamma_{3,n}(V)\}), \end{aligned}$$

where $\max\{\gamma_{1,n}, \gamma_{2,n}, \gamma_{3,n}(V)\} = o(1)$ as $n \to \infty$. Using elementary arguments, this result implies that: (1) $d(\hat{\theta}_n, V) = o_p(1)$ as $n \to \infty$ and (2) $d(\hat{\theta}_n, V)$ converges in probability to zero at a rate of $\max\{\gamma_{1,n}, \gamma_{2,n}, \gamma_{3,n}(V)\}^{-1} = \min\{\gamma_{1,n}^{-1}, \gamma_{2,n}^{-1}, \gamma_{3,n}(V)^{-1}\}$. □

Lemma A.2. *Under Assumption B.1, Assumption B.5 is satisfied for all the formulations of the dynamic decision problem described in Appendix B.*

Proof. We verify the result for each formulation described in Appendix B. We begin with the *conditional value function* formulation. In this case: $s = (x, \varepsilon)$ and $F(V, s') = V(s')$. Then:

$$E[F(f_1, s') - F(f_2, s')|s, a] = E(f_1(s') - f_2(s')|s, a) \leq d(f_1, f_2),$$

where the equality holds by definition of F and the inequality holds because d is the sup-norm metric.

We now consider the *social surplus function* formulation. In this case: $s = (x, a)$, $F(V, s') = G(V(x')|x')$, and $A(s) = \{a\}$. Then:

$$\begin{aligned} &E[F(f_1, s') - F(f_2, s')|s, a] = E[G(f_1(s')|s') - G(f_2(s')|s')|s, a], \\ &= E[E[\max_{a'\in A(x')}(f_1(x', a') + \varepsilon(a')) - \max_{a'\in A(x')}(f_2(x', a') + \varepsilon(a'))|x']|x, a], \\ &\leq E[|E[\max_{a'\in A(x')}(f_1(x', a') + \varepsilon(a')) - \max_{a'\in A(x')}(f_2(x', a') + \varepsilon(a'))|x']||x, a], \\ &\leq E[\max_{a'\in A(x')}|f_1(x', a') - f_2(x', a')||x, a] \leq d(f_1, f_2), \end{aligned}$$

where the first equality holds by definition of F, the second equality holds by definition of G, the first and second inequalities hold by elementary arguments, and the final inequality holds because d is the sup-norm metric.

We conclude with the *choice-specific value function* formulation. In this case: $s = (x, a)$, $F(V, s') = G(\beta^{-1}u(x') + V(x')|x')$, and $A(s) = \{a\}$. Using the same arguments as in the social surplus function

formulation, it is not hard to verify the result. □

Proof of Theorem 4.1. This proof proceeds by induction. Set $t = 1$. By definition, $\hat{\theta}_{T+2-t,n} = V_{T+1}$ and, thus, $d(V_{T+2-t}, \hat{\theta}_{T+2-t,n}) = 0$. We now prove the inductive step. Suppose that for some $t \geq 1$, $d(V_{T+2-t}, \hat{\theta}_{T+2-t,n}) = O_p(\max\{\upsilon_{1,n}, \upsilon_{2,n}, \upsilon_{3,n}\})$. We now show that: $d(V_{T+1-t}, \hat{\theta}_{T+1-t,n}) = o_p(1)$, as $n \to \infty$. By the triangular inequality:

$$d(V_{T+1-t}, \hat{\theta}_{T+1-t,n}) = d(V_{T+1-t}, \hat{\theta}^*_{T+1-t,n}) + d(\hat{\theta}^*_{T+1-t,n}, \hat{\theta}_{T+1-t,n}).$$

We now analyze each of the terms on the right hand side satisfies the desired property.

We begin with $d(V_{T+1-t}, \hat{\theta}^*_{T+1-t,n})$. Fix $s \in S$ arbitrarily. Let $a_{1,n}(s) \in A(s)$ be such that:

$$\hat{\theta}^*_{T+1-t,n}(s) = u(s, a_{1,n}(s)) + \beta E(F(\hat{\theta}_{T+2-t,n}, s')|s, a_{1,n}(s)),$$

i.e., $a_{1,n}(s)$ is the maximizer, which exists due to the fact that $A(s)$ is a finite set. Then, consider the following derivation:

$$\begin{aligned}
\hat{\theta}^*_{T+1-t,n}(s) &= u(s, a_{1,n}(s)) + \beta E(F(\hat{\theta}_{T+2-t,n}, s')|s, a_{1,n}(s)), \\
&\leq u(s, a_{1,n}(s)) + \beta E(F(V_{T+2-t}, s')|s, a_{1,n}(s)) + \beta d(\hat{\theta}_{T+2-t,n}, V_{T+2-t}), \\
&\leq V_{T+1-t}(s) + \beta d(\hat{\theta}_{T+2-t,n}, V_{T+2-t}),
\end{aligned}$$

where the first equality holds by Eq. (A), the next inequality holds by Assumption B.5, and the final inequality holds by Eq. (4.1). As a consequence, it follows that: $\sup_{s\in S}(\hat{\theta}^*_{T+1-t,n}(s) - V_{T+1-t}(s)) \leq \beta d(\hat{\theta}_{T+2-t,n}, V_{T+2-t})$. By a similar argument, one can show that $\sup_{s\in S}(V_{T+1-t}(s) - \hat{\theta}^*_{T+1-t,n}(s)) \leq \beta d(\hat{\theta}_{T+2-t,n}, V_{T+2-t})$. By combining both inequalities with Assumption B.1, it follows that:

$$d(\hat{\theta}^*_{T+1-t,n}, V_{T+1-t}) = \sup_{s\in S} |\hat{\theta}^*_{T+1-t,n}(s) - V_{T+1-t}(s)| \leq \beta d(\hat{\theta}_{T+2-t,n}, V_{T+2-t}).$$

By the inductive assumption, it follows that: $d(\hat{\theta}^*_{T+1-t,n}, V_{T+1-t}) = O_p(\max\{\upsilon_{1,n}, \upsilon_{2,n}, \upsilon_{3,n}\})$.

We continue with $d(\hat{\theta}^*_{T+1-t,n}, \hat{\theta}_{T+1-t,n})$. Consider the following derivation:

$$\begin{aligned}
d(\hat{\theta}^*_{T+1-t,n}, \hat{\theta}_{T+1-t,n}) &\leq K_2\, d_n(\hat{\theta}^*_{T+1-t,n}, \hat{\theta}_{T+1-t,n}) + \lambda_{1,n}, \\
&\leq K_2 \inf_{\theta\in\Theta_n} d_n(\theta, \hat{\theta}^*_{T+1-t,n}) + \lambda_{1,n} + \lambda_{2,n}(\hat{\theta}^*_{T+1-t,n}), \\
&\leq K_1^{-1} K_2 \inf_{\theta\in\Theta_n} d(\theta, \hat{\theta}^*_{T+1-t,n}) + \lambda_{1,n}(1 + K_1^{-1}) + \lambda_{2,n}, \\
&\leq K_1^{-1} K_2 \sup_{f\in\mathcal{F}} \lambda_{3,n}(f) + \lambda_{1,n}(1 + K_1^{-1}) + \lambda_{2,n}
\end{aligned}$$

where the first and third inequalities hold by Assumption B.2, the second inequality holds by Assumption B.3, and the final inequality holds by Assumption B.4. By the properties of $\lambda_{1,n}$, $\lambda_{2,n}$, and $\lambda_{3,n}$, it follows that $d(\hat{\theta}^*_{T+1-t,n}, \hat{\theta}_{T+1-t,n}) = O_p(\max\{\upsilon_{1,n}, \upsilon_{2,n}, \upsilon_{3,n}\})$. □

Proof of Lemma 4.1. Part 1. Consider a pair of functions $f_1, f_2 \in \tilde{\mathcal{F}}$. First, consider $t = T+1$. By definition, $\tilde{\Gamma} f_1(s,t) = \tilde{\Gamma} f_2(s,t) = 0\ \forall s \in S$, which implies that: $\sup_{s\in S} |\tilde{\Gamma} f_1(s, T+1) - \tilde{\Gamma} f_2(s, T+1)| = 0$. Next, consider

$t = 1, \ldots, T$. For any arbitrary $s \in S$:

$$\begin{aligned}
|\tilde{\Gamma} f_1(s,t) - \tilde{\Gamma} f_2(s,t)| &= \left\{ \begin{array}{c} \sup_{a \in A(s)} \{u(s,a) + \beta E(F(f_1,(s',t+1))|(s,t),a)\} \\ -\sup_{a \in A(s)} \{u(s,a) + \beta E(F(f_2,(s',t+1))|(s,t),a)\} \end{array} \right\} \\
&\leq \beta \sup_{a \in A(s)} E(F(f_1,(s',t+1)) - F(f_2,(s',t+1))|(s,t),a) \\
&= \beta \sup_{a \in A(s)} E(F(f_{1,t+1},s') - F(f_{2,t+1},s')|s,a) \\
&\leq \beta \sup_{a \in A(s)} E(f_{1,t+1}(s') - f_{2,t+1}(s')|s,a) \\
&\leq \beta \sup_{s \in S} |f_{1,t+1}(s) - f_{2,t+1}(s)| = \beta \sup_{s \in S} |f_1(s,t+1) - f_2(s,t+1)|,
\end{aligned}$$

where for any $f \in \tilde{\mathcal{F}}$ and $(s,t) \in S \times \{1, \ldots, T+1\}$, we use $f_t(s) \equiv f(s,t)$. Notice that we are also using the fact that conditional expectations are time invariant, but that is assumed for simplicity of notation, i.e., the assumption can be eliminated by indexing expectations with a time index. By reversing the roles of f_1 and f_2, we deduce that: $\forall t = 1, \ldots, T$:

$$\sup_{s \in S} |\tilde{\Gamma} f_1(s,t) - \tilde{\Gamma} f_2(s,t)| \leq \beta \sup_{s \in S} |f_1(s,t+1) - f_2(s,t+1)|.$$

By combining information from all values of $t = 1, \ldots, T+1$, it follows that:

$$\tilde{d}(\tilde{\Gamma} f_1, \tilde{\Gamma} f_2) = \max_{t=1,\ldots,T+1} \sup_{s \in S} |tilde\Gamma f_1(s,t) - tilde\Gamma f_2(s,t)| \leq \beta \max_{t=1,\ldots,T+1} \sup_{s \in S} |f_1(s,t) - f_2(s,t)| = \beta \tilde{d}(f_1, f_2).$$

Part 2. By the Contraction Mapping Theorem (see, e.g., Stokey and Lucas (1989, page 50)) the mapping

$tilde\Gamma$ has a unique fixed point in $\tilde{\mathcal{F}}$. It suffices to show that V is such fixed point, i.e., $\tilde{d}(V, \tilde{\Gamma} V) = 0$. For $t = T+1$, the definition of $\tilde{\Gamma} V$ is: $(\tilde{\Gamma} V)(s, T+1) = 0 = V(T+1, s) \equiv V_{T+1}(s)$. For any other $t = 1, \ldots, T$, it follows that:

$$\begin{aligned}
(\tilde{\Gamma} V)(s,t) &= \sup_{a \in A(s)} \{u(s,a) + \beta E(F(V,(s',t+1))|(s,t),a)\} \\
&= \sup_{a \in A(s)} \{u(s,a) + \beta E(F(V_{t+1},s'))|s,a)\} = V_t(s) = V(s,t),
\end{aligned}$$

which completes the proof. □

Proof of Theorem 4.2. By Lemma 4.1, the analogue of Assumptions A.2-A.4 hold for the state space $\tilde{S}$. Under these assumptions, the result is a corollary of Theorem 3.1. □

Proof of Theorem 5.1. Fix $\varepsilon > 0$ arbitrarily. Then, $\exists \delta > 0$ such that $\forall \pi \in \Pi$: $d_2(\pi, \pi^*) > \varepsilon \Rightarrow Q(V(\cdot|\pi)) - Q(V(\cdot|\pi^*)) > \delta$. This implies that $\exists \delta > 0$ such that:

$$P(d_2(\hat{\pi}_I, \pi^*) \leq \varepsilon) \geq P(Q(V(\cdot|\hat{\pi}_I)) - Q(V(\cdot|\pi^*)) \leq \delta).$$

The strategy of the proof is to show that the RHS converges to one as $I \to \infty$. To this end, for a fixed $\delta > 0$,

consider the following argument:

$$
\begin{aligned}
&P(Q(V(\cdot|\hat{\pi}_I)) - Q(V(\cdot|\pi^*) \leq \delta) \\
&= P\left(\begin{array}{c} Q(V(\cdot|\hat{\pi}_I)) - Q(\hat{\theta}_n(\cdot|\hat{\pi}_I)) + Q(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) \\ +Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q(V(\cdot|\pi^*)) \leq \delta \end{array} \right), \\
&\geq P\left(\begin{array}{c} \{Q(V(\cdot|\hat{\pi}_I)) - Q(\hat{\theta}_n(\cdot|\hat{\pi}_I)) \leq \delta/3\} \cap \{Q(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) \leq \delta/3\} \\ \cap\{Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q(V(\cdot|\pi^*)) \leq \delta/3\} \end{array} \right), \\
&\geq \left\{ \begin{array}{c} P(Q(V(\cdot|\hat{\pi}_I)) - Q(\hat{\theta}_n(\cdot|\hat{\pi}_I)) \leq \delta/3) + P(Q(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) \leq \delta/3) \\ +P(Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q(V(\cdot|\pi^*)) \leq \delta/3) - 2 \end{array} \right\}.
\end{aligned}
$$

The last expression on the RHS includes three probability expressions. The proof is completed by showing that these expressions converge to one as $I \to \infty$.

Consider the first expression. By the uniform continuity of Q, there is $\eta > 0$ such that:

$$\sup_{\pi \in \Pi} d_1(\hat{\theta}_n(\cdot|\pi), V(\cdot|\pi)) < \eta \Rightarrow d_1(\hat{\theta}_n(\cdot|\hat{\pi}_I), V(\cdot|\hat{\pi}_I)) < \eta \Rightarrow Q(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q(V(\cdot|\hat{\pi}_I)) \leq \delta/3.$$

It follows then that:

$$P(Q(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q(V(\cdot|\hat{\pi}_I)) \leq \delta/3) \to 1, \text{ as } n \to \infty,$$

or, equivalently, as $I \to \infty$ because $n(I) \to \infty$ as $I \to \infty$.

Consider the second expression. Notice that:

$$Q(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) \leq \sup_{f \in \Theta_n} |Q_I(f) - Q(f)| = \hat{c}(n, I),$$

and since $\hat{c}(n, I) = o_p(\delta(n(I), I)) = o_p(1)$ as $I \to \infty$, it follows that:

$$P(Q(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) \leq \delta/3) \to 1, \text{ as } I \to \infty.$$

Finally, consider the third expression. By definition of the estimation:

$$Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) \leq \inf_{\pi \in \Pi} Q_I(\theta_n(\cdot|\pi)) + o_p(1) \leq Q_I(\theta_n(\cdot|\pi^*)) + o_p(1),$$

as $I \to \infty$. This implies that:

$$
\begin{aligned}
&Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q(V(\cdot|\pi^*)) \leq Q_I(\hat{\theta}_n(\cdot|\pi^*)) - Q(V(\cdot|\pi^*)) + o_p(1), \\
&= Q_I(\hat{\theta}_n(\cdot|\pi^*)) - Q(\hat{\theta}_n(\cdot|\pi^*)) + Q(\hat{\theta}_n(\cdot|\pi^*)) - Q(V(\cdot|\pi^*)) + o_p(1), \\
&\leq \sup_{f \in \Theta_n} |Q_I(f) - Q(f)| + Q(\hat{\theta}_n(\cdot|\pi^*)) - Q(V(\cdot|\pi^*)) + o_p(1).
\end{aligned}
$$

Repeating arguments used in previous expressions, it follows that:

$$P(Q_I(\hat{\theta}_n(\cdot|\hat{\pi}_I)) - Q(V(\cdot|\pi^*)) \leq \delta/3) \to 1, \text{ as } I \to \infty,$$

completing the proof. □

Appendix B Alternative formulations of decision problem

As described in the main text of Section 2, there are three possible ways to formulate the dynamic decision problem of Eq. (2.1) as a recursive problem. The objective of this section is to describe each of these formulations and show that each of them is a special case of the unified formulation described in the main text.

B.1 Value function formulation

This formulation of the dynamic decision problem is the most obvious one but, as we will argue, the most computationally demanding in terms of the size of the state space. It is easy to verify that Eq. (2.1) can be represented as follows:

$$\begin{aligned} V(x,\varepsilon) &= \max_{a\in A(x)} \{u(x,a)+\varepsilon(a)+\beta \int V(x',\varepsilon')dP(x',\varepsilon'|x,\varepsilon,a)\}, \\ &= \max_{a\in A(x)} \{u(x,a)+\varepsilon(a)+\beta E[V(x',\varepsilon')|x,\varepsilon,a]\}. \end{aligned}$$

If we let $s \equiv (x,\varepsilon)$ and $S \equiv X \times E$, $V : S \to \mathbb{R}$ is the fixed point of the following mapping:

$$(\Gamma_1 m)(x,\varepsilon) = \max_{a\in A(x)} \{u(x,a)+\varepsilon(a)+\beta E[m(x',\varepsilon')|x,\varepsilon,a]\}, \ \forall (x,\varepsilon) \in S.$$

This corresponds to the conditional value function formulation of the problem (Rust (1987, Eq. (4.4))). Notice that our derivations so far did not require the Conditional Independence (CI) Assumption. It is straight-forward to verify that it is a contraction mapping (in the sup-norm) by using the Blackwell (1965) sufficient conditions for a contraction (see Stokey and Lucas (1989, Theorem 3.3)). This is an obvious object of interest, as it allows us to deduce optimal decision rules in the following way:

$$f(x,\varepsilon) = \arg\max_{a\in A(x)} \{u(x,a)+\varepsilon(a)+\beta E[V(x',\varepsilon')|x,\varepsilon,a]\},$$

with a related conditional choice probability (CCP):

$$P(a=\bar{a}|x) = \int_{\varepsilon} \left(\arg\max_{a\in A(x)} \{u(x,a)+\varepsilon(a)+\beta E[V(x',\varepsilon')|x,\varepsilon,a]\} = \bar{a} \right) dP(\varepsilon|x).$$

We now verify that this formulation is a particular case of the unified formulation. This is shown in Table 5, which explains how the objects in the unified notation are defined in terms of the primitive objects in Eq. (2.1). According to the table, F is defined as follows: $F(V,s') = V(s')$. It is straightforward to verify that this function satisfies all of the properties required for F.

Primitive object	V	(x,ε)	$X \times E$	$A(x)$	a	$u(x,a)+\varepsilon(a)$	$V(s')$
Unified formulation	V	s	S	$A(s)$	a	$u(s,a)$	$F(V,s')$

Table 5: Representation of the primitive object of the conditional value function formulation in terms of the unified formulation.

Notice that the state space in this formulation is given by $s = (x,\varepsilon)$, which is relatively large as it includes $\varepsilon = \{\varepsilon(a) : a \in A(x)\}$ (i.e., has the cardinality of $A(x)$). By the nature of the type of problems

under consideration it is reasonable to consider alternative formulations of the problem, where ε is not part of the state space.

B.2 Social surplus function formulation

This formulation of the problem is developed in Rust (1988). In order to explain the formulation, it is necessary to define certain objects. Given a vector $v(x) = \{v(x,a) : a \in A(x)\}$, define the social surplus function as follows:

$$G(v(x)|x) \equiv \int_{\varepsilon} \max_{a \in A(x)} \{v(x,a) + \varepsilon(a)\}\, dP(\varepsilon|x) = E\left(\max_{a \in A(x)} \{v(x,a) + \varepsilon(a)\} \middle| x \right).$$

If we let $s \equiv (x,a)$ and $S \equiv \{(x,a) : a \in A(x), x \in X\}$, Rust (1988, Theorem 3.2) indicates that:

$$V(x,\varepsilon) = \max_{a \in A(x)} [v(x,a) + \varepsilon(a)],$$

where $v : S \to \mathbb{R}$ is the unique fixed point of the following contraction mapping:

$$(\Gamma_2 m)(x,a) = u(x,a) + \beta E[G(m(x')|x'), a], \forall (x,a) \in S.$$

Notice that, in the formula, $m(x')$ is shorthand for the vector $\{m(x',i') : i' \in A(x')\}$. Our assumptions imply that the mapping is a contraction mapping (in the sup-norm) by using the Blackwell (1965) sufficient conditions for a contraction (in this case, it is important to establish that G satisfies the properties in Rust (1988, Theorem 3.2)). Notice that the fixed point v is the object of interest because it allows us to recover the original value function V (using Eq. (B.2)) and, with it, we can then construct the rest of the objects of interest such as optimal decision rules and CCPs.

The advantage of working with this formulation of the dynamic decision problem is that the state space is (x,a), whose dimension is smaller than for the condition value function formulation. Another advantage is that the CCPs can be written simply as (see Rust (1988, Theorem 3.4)):

$$P(a = \bar{a}|x) = G_{\bar{a}}(v(x)|x),$$

where $G_{\bar{a}}$ denotes the derivative of the social surplus function (i.e., Eq. (B.2)) with respect to $v(x,\bar{a})$.

We now verify that this formulation is a particular case of the unified formulation. This is shown in Table 6, which explains how the objects in the unified notation are defined in terms of the primitive objects in Eq. (2.1). According to the table, F is defined as follows: $F(V,s') = G(V(x')|x')$, where G is the social surplus function (i.e., Eq. (B.2)) with its own choice set $A(x')$. Theorem 3.1 in Rust (1988) shows that the function G satisfies all of the desired properties.

Unified formulation	v	s	S	$A(s)$	a	$u(s,a)$	$F(V,s')$
Primitive object	V	(x,a)	$\{(x,a) : a \in A(x), x \in X\}$	$\{a\}$	a	$u(x,a)$	$G(V(x')\|x')$

Table 6: Representation of the primitive object of the social surplus function formulation formulation in terms of the unified formulation.

B.3 Choice-specific value function formulation

This formulation is related to the social surplus function formulation and is explained in Rust (1988), page 1014. As in the previous formulation, let $s \equiv (x, a)$ and $S \equiv \{(x, a) : a \in A(x), x \in X\}$. Under our assumptions, Rust (1988) indicates that:

$$v(x, a) = u(x, a) + \beta EV(x, a),$$

where $v : S \to \mathbb{R}$ is the function in Eq. (B.2) and $EV : S \to \mathbb{R}$ is the fixed point in the following contraction mapping:

$$(\Gamma_3 m)(x, a) = E[G(u(x') + \beta m(x')|(x'))|x, a], \ \forall (x, a) \in S.$$

This formulation is intimately related to the social surplus function formulation and shares all its advantages relative to the conditional value function formulation, i.e., the state space is (x, a), which have a relatively smaller dimension, and the CCPs are relatively easy to compute using Eq. (B.2).

We now verify that this formulation is a particular case of the unified formulation. This is shown in Table 7, which explains how the objects in the unified notation are defined in terms of the primitive objects in Eq. (2.1). According to the table, F is defined as follows: $F(V, s') = G(\beta^{-1}u(x') + V(x')|x')$, where G is the social surplus function (i.e., Eq. (B.2)) with its own action set $A(x')$. Rust (1988, Theorem 3.1) shows that the function G satisfies all of the desired properties.

Unified formulation	V	s	S	$A(s)$	a	$u(s, i)$	$F(V, s')$
Primitive object	EV	(x, a)	$\{(x, a) : i \in A(x), x \in X\}$	$\{a\}$	a	0	$G(\beta^{-1}u(x') + V(x')\|x')$

Table 7: Representation of the primitive object of the choice-specific value function formulation formulation in terms of the unified formulation.