Bayesian Estimation of Dynamic Discrete Choice Models*

Susumu Imai
Concordia University and CIREQ

Neelam Jain
Northern Illinois University

and
Andrew Ching
Ohio State University

January, 2005

Abstract

We propose a new methodology for structural estimation of dynamic discrete choice models. We combine the Dynamic Programming (DP) solution algorithm with the Bayesian Markov Chain Monte Carlo algorithm into a single algorithm that solves the DP problem and estimates the parameters simultaneously. As a result, the computational burden of estimating a dynamic model becomes comparable to that of a static model. Another feature of our algorithm is that even though per solution-estimation iteration, the number of grid points on the state variable is small, the number of effective grid points increases with the number of estimation iterations. This is how we overcome the "Curse of Dimensionality". We simulate and estimate several versions of a simple model of entry and exit to illustrate our methodology. We also prove that under standard conditions, the parameters converge in probability to the true posterior distribution, regardless of the starting values.

*Please direct all correspondence to Susumu Imai, Department of Economics, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, QC. H3G 1M8, Canada, e-mail: simai@alcor.concordia.ca, phone: 514-848-2424-3907, or Neelam Jain,Department of Economics, Northern Illinois University, 508 Zulauf Hall, DeKalb, IL. 60115, e-mail: njain@niu.edu, phone: (815)-753-6964

1 Introduction

Structural estimation of Dynamic Discrete Choice (DDC) models has become increasingly popular in empirical economics. Examples include Keane and Wolpin (1997) in labor economics, Erdem and Keane (1995) in marketing and Rust (1987) in empirical industrial organization. Structural estimation is appealing for at least two reasons. First, it captures the dynamic forward-looking behavior of individuals, which is very important in understanding agents' behavior in various settings. For example, in labor market, individuals carefully consider future prospects when they switch occupations. Secondly, since the estimation is based on explicit solution of a structural model, it avoids the Lucas Critique. Hence, after the estimation, policy experiments can be relatively straightforwardly conducted by simply changing the estimated value of "policy" parameters and simulating the model to assess the change. However, one major obstacle in adopting the structural estimation method has been its computational burden. This is mainly due to the following two reasons.

First, in dynamic structural estimation, the likelihood or the moment conditions are based on the explicit solution of the dynamic model. In order to solve a dynamic model, we need to compute the Bellman equation repeatedly until the calculated expected value function converges. Second, in solving the Dynamic Programming (DP) Problem, the Bellman equation has to be solved at each possible point in the state space. The possible number of points in the state space increases exponentially with the increase in the dimensionality of the state space. This is commonly referred to as the "Curse of Dimensionality", and makes the estimation of the dynamic model infeasible even in a relatively simple setting.

In this paper, we propose an estimator that helps overcome the two computational difficulties of structural estimation. Our estimator is based on the Bayesian Markov Chain Monte Carlo (MCMC) estimation algorithm, where we simulate the posterior distribution by repeatedly drawing parameters from a Markov Chain until convergence. In contrast to the conventional MCMC estimation approach, we combine the Bellman equation step and the MCMC algorithm step into a single hybrid solution-estimation step, which we iterate until convergence. The key innovation in our algorithm is that for a given state space point, we need to solve the Bellman equation only once between each estimation step. Since evaluating a Bellman equation once is as computationally demanding as computing a static model, the computational burden of estimating a DP model is in order of magnitude comparable to that of estimating a static model.

Furthermore, since we move the parameters according to the MCMC algorithm after each Bellman step, we are "estimating" the model and solving for the DP problem at the same time. This is in contrast to conventional estimation methods that "estimate" the model only after solving the DP problem. In that sense, our estimation method is related to the algorithm advocated by Aguirreagabiria and Mira (2001), where they propose either to iterate the Bellman equation only a limited number of times before constructing the likelihood, or to solve the DP problem "roughly" at the initial stage of the Maximum Likelihood routine and increase the precision of the DP solution with the iteration of the Maximum Likelihood routine. The first estimation strategy, which is not based on the full solution of the model, cannot handle unobserved heterogeneity. This is because this estimation method essentially recovers the value function from the observed choices of people at each point of the state space. But if there are unobserved heterogeneities, the state space points are unobservable in the data. In the second strategy, they still compute the solution of the DP problem, whether exact or inexact, during each estimation step. Hence, the computational burden of solving for the DP problem at each estimation step, although diminished, still remains. In our algorithm, we only need to solve the Bellman equation once between each estimation step.[1]

Specifically, we start with an initial guess of the expected value function (emax function). We then evaluate the Bellman equation for each state space point, if the number of state space points is finite, or for a subset of the state space grid points if the state space is continuous. We then use Bayesian MCMC to update the parameter vector. We update the emax function for a state space point by averaging with those past iterations in which the parameter vector is "close" to the current parameter vector and the state variables are either exactly the same as the current state variables (if the state space is finite) or close to the

[1] In contrast to Ackerberg (2004), where the entire DP problem needs to be solved for each parameter simulation, in our algorithm, the Bellman equation needs to be computed only once for each parameter value. Furthermore, there is an additional computational gain because MCMC algorithm guarantees that except for the initial burn-in simulations, most of the parameter draws are from a distribution close to the true posterior distribution. In Ackerberg's case, the initial parameter simulation and therefore the DP solution would be inefficient because at the initial stage, true parameter distribution is not known.

current state variables (if the state space is continuous). This method of updating the emax function is similar to Pakes and McGuire (2001) except in the important respect that we also include the parameter vector in determining the set of iterations over which averaging occurs.

Our algorithm also addresses the problem of 'the Curse of Dimensionality'. In most DP solution exercises involving a continuous state variable, the state space grid points, once determined, are fixed over the entire algorithm, as in Rust (1997). In our Bayesian DP algorithm, the state space grid points do not have to be the same for each solution-estimation iteration. In fact, by varying the state space grid points at each solution-estimation iteration, our algorithm allows for an arbitrarily large number of state space grid points by increasing the number of iterations. This is how our estimation method overcomes the "Curse of Dimensionality".

The main reason behind the computational advantage of our estimation algorithm is the use of information obtained from past iterations. In the conventional solution-estimation algorithm, at iteration t, most of the information gained in all past estimation iterations remains unused, except for the iteration $t-1$ likelihood and its Jacobian and Hessian in Classical ML estimation, and MCMC transition function in Bayesian MCMC estimation. In contrast, we extensively use the vast amount of computational results obtained in past iterations, especially those that are helpful in solving the DP problem.

We demonstrate the performance of our algorithm by estimating a dynamic model of firm entry and exit choice with observed and unobserved heterogeneities. The unobserved random effects coefficients are assumed to have a continuous distribution function, and the observed characterisitcs are assumed to be continuous as well. It is well known that for a conventional Dynamic Programming Simulated Maximum Likelihood estimation strategy, this setup imposes an almost prohibitive computational burden. The computational burden is due to the fact that during each estimation step, the DP problem has to be solved for each firm hundreds of times. Because of the observed heterogeneity, each firm has a different parameter value, and furthermore, because the random effects term has to be integrated out numerically via Monte-Carlo integration, for each firm, one has to simulate the random effects parameter hundreds of times, and for each simulation, solve for the DP problem. This is why most practitioners of structural estimation follow Heckman and Singer (1984) and assume discrete distributions for random effects and only allow for discrete types as observed characteristics.

We show that using our algorithm, the above estimation exercise becomes one that is computationally quite similar in difficulty to the Bayesian estimation of a static discrete choice model with random effects (see McCullogh and Rossi (1994) for details), and thus is feasible. Indeed, the computing time for our estimation exercise (with 100 firms and 100 time periods) is about 13 hours, similar to the time required to estimate a reasonably complicated static random effects model. In contrast, even a single iteration of the conventional simulated maximum likelihood estimation routine of the same model took 6 hours and 20 minutes.

In addition to the experiments, we formally prove that under very mild conditions, the distribution of parameter estimates simulated from our solution-estimation algorithm converges to the true posterior distribution in probability as we increase the number of iterations. The proof relies on coupling theory (see Rosenthal (1995)) in addition to the standard asymptotic techniques such as the Law of Large Numbers.

Our algorithm shows that the Bayesian methods of estimation, suitably modified, can be used effectively to conduct full solution based estimation of structural dynamic discrete choice models. Thus far, application of Bayesian methods to estimate such models has been particularly difficult. The main reason is that the solution of the DP problem, i.e. the repeated calculation of the Bellman equation is computationally so demanding that the MCMC, which typically involves far more iterations than the standard Maximum Likelihood routine, becomes infeasible. One of the few examples of Bayesian estimation is Lancaster (1997). He successfully estimates the equilibrium search model where the Bellman equation can be transformed into an equation where all the information on optimal choice of the individual can be summarized in the reservation wage, and hence, there is no need for solving the value function. Another example is Geweke and Keane (1995) who estimate the DDC model without solving the DP problem. In contrast, our paper accomplishes Bayesian estimation based on full solution of the DP problem by simultaneously solving for the DP problem and iterating on the MCMC algorithm. The difference turns out to be important because the estimation algorithms that are not based on the full solution of the model can only accomodate limited specification of unobserved heterogeneities.

Our estimation method not only makes Bayesian application to DDC models computationally feasible, but possibly even superior to the existing (non-Bayesian) methods, by reducing the computational burden of estimating a dynamic model to that of estimating a static one. Furthermore, the usually cited advantages of Bayesian estimation over classical estimation methods apply here as well. That is, first, the conditions for the convergence of the MCMC algorithm are in general weaker than the conditions for the global maximum of the Maximum Likelihood (ML) estimator, as we show in this paper. Second, in MCMC, standard errors can be derived straightforwardly as a byproduct of the estimation routine, whereas in ML estimation, standard errors have to be computed usually either by inverting the numerically calculated Information Matrix, which is valid only in a large sample world, or by repeatedly bootstrapping and reestimating the model, which is computationally demanding.

The organization of the paper is as follows. In Section 2, we present a general version of the DDC model and discuss conventional estimation methods as well as our Bayesian DP algorithm. In Section 3, we prove convergence of our algorithm under some mild conditions. In Section 4, we present a simple model of entry and exit. In Section 5, we present the simulation and estimation results of several experiments applied to the model of entry and exit. Finally, in Section 6, we conclude and briefly discuss future direction of this research. The Appendix contains all proofs.

2 The Framework

Let θ be the J dimensional parameter vector. Let S be the set of state space points and let s be an element of S. We assume that S is finite. Let A be the set of all possible actions and let a be an element of A. We assume A to be finite to study discrete choice models.

The value of choice a at parameter θ and state vector s is,

$$\mathcal{V}(s, a, \epsilon_a, \theta) = U(s, a, \epsilon_a, \theta) + \beta E_{\epsilon'}\left[V(s', \epsilon', \theta)\right] \tag{1}$$

where s' is the next period's state variable, U is the current return function. ϵ is a vector whose a th element ϵ_a is a random shock to current returns to choice a. Finally, β is the discount factor. We assume that ϵ follows a multivariate distribution $F_{\epsilon}(\epsilon|\theta)$, which is independent over time. The expectation is taken with respect to the next period's shock ϵ'. We assume that the next period's state variable s' is a deterministic function of the current state variable s, current action a, and parameter θ [2]. That is,

$$s' = s'(s, a, \theta).$$

The value function is defined to be as follows.

$$V(s, \epsilon, \theta) = \max_{a \in A} \mathcal{V}(s, a, \epsilon_a, \theta)$$

We assume that the dataset for estimation includes variables which correspond to state vector s and choice a in our model but the choice shock ϵ is not observed. That is, the observed data is $Y_{N,T} \equiv \left\{ s^d_{i,\tau}, a^d_{i,\tau}, F^d_{i,\tau} \right\}_{i=1,\tau=1}^{N,T}$ [3], where N is the number of firms and T is the number of time periods. Furthermore,

$$a^d_{i,\tau} = \arg\max_{a \in A} \mathcal{V}(s^d_{i,\tau}, a, \epsilon_a, \theta)$$

$$\begin{aligned} F^d_{i,\tau} \quad = \quad & U(s^d_{i,\tau}, a^d_{i,\tau}, \epsilon_{a^d_{i,\tau}}, \theta) \quad \text{if } \left(s^d_{i,\tau}, a^d_{i,\tau}\right) \in \Psi \\ & 0 \qquad\qquad\qquad\qquad \text{otherwise.} \end{aligned}$$

The current period return is observable in the data only when the pair of state and choice variables belongs to the set Ψ. In the entry/exit problem of firms that we discuss later, profit of a firm is only observed when the incumbent firm stays in. In this case, Ψ is a set whose state variable is being an incumbent (and the capital stock) and the choice variable is staying in.

Let $\pi()$ be the prior distribution of θ. Furthermore, let $L(Y_{N,T}|\theta)$ be the likelihood of the model, given the parameter θ and the value function $V(.,.,\theta)$, which is the solution of the DP problem. Then, we have the following posterior distribution function of θ.

$$P(\theta|Y_{N,T}) \propto \pi(\theta) L(Y_{N,T}|\theta). \tag{2}$$

Let $\boldsymbol{\epsilon} \equiv \{\epsilon_{i,\tau}\}_{i=1,\tau=1}^{N,T}$. Because $\boldsymbol{\epsilon}$ is unobserved to the econometrician, the likelihood is an integral over it. That is, if we define $L(Y_{N,T}|\boldsymbol{\epsilon}, \theta)$ to be the likelihood conditional on $(\boldsymbol{\epsilon}, \theta)$,

[2] This is a simplifying assumption for now. Later in the paper, we study random dynamics as well.

[3] We denote any variables with d superscript to be the data.

then,

$$L(Y_{N,T}|\theta) = \int L(Y_{N,T}|\boldsymbol{\epsilon}, \theta) dF_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}|\theta).$$

The value function enters in the likelihood through choice probability, which is a component of the likelihood. That is,

$$P\left[a = a^d_{i,\tau} | s^d_{i,\tau}, V, \theta\right] = \Pr\left[\widehat{\epsilon}_{a,i,\tau} : a^d_{i,\tau} = \arg\max_{a \in A} \mathcal{V}(s^d_{i,\tau}, a, \widehat{\epsilon}_{a,i,\tau}, \theta)\right]$$

Below we briefly describe the conventional estimation approaches and then, the Bayesian dynamic programming algorithm we propose.

2.1 The Maximum Likelihood Estimation

The conventional ML estimation procedure of the dynamic programming problem consists of two main steps. First is the solution of the dynamic programming problem and the subsequent construction of the likelihood, which is called "the inner loop" and second is the estimation of the parameter vector, which is called "the outer loop".

1. **Dynamic Programming Step**: Given parameter vector θ, we solve the Bellman equation, given by equation 1. This typically involves several steps.

   (a) First, the random choice shock, ϵ is drawn a fixed number of times, say, M, generating $\epsilon^{(m)}, m = 1, ..., M$. At iteration 0, set initial guess of the value function to be, for example, zero. That is, $V^{(0)}(s, \epsilon^{(m)}, \theta) = 0$ for every $s \in S$, $\epsilon^{(m)}$. We also let the expected value function (Emax function) to be zero, i.e., $E_{\epsilon'}\left[V^{(0)}(s, \epsilon', \theta)\right] = 0$ for every $s \in S$.

   (b) Assume we are at iteration t of the DP algorithm. Given $s \in S$ and $\epsilon^{(m)}$, the value of every choice $a \in A$ is calculated. For the Emax function, we use the approximated expected value function $\widehat{E}_{\epsilon'}\left[V^{(t-1)}(s', \epsilon', \theta)\right]$ computed at the previous iteration $t-1$ for every $s' \in S$. Hence, the iteration t value of choice a is,

$$\mathcal{V}^{(t)}(s, a, \epsilon^{(m)}_a, \theta) = U(s, a, \epsilon^{(m)}_a, \theta) + \beta \widehat{E}_{\epsilon'}\left[V^{(t-1)}(s', \epsilon', \theta)\right].$$

Then, we compute the value function,

$$V^{(t)}(s, \epsilon^{(m)}, \theta) = \max_{a \in A} \mathcal{V}^{(t)}(s, a, \epsilon_a^{(m)}, \theta). \quad (3)$$

The above calculation is done for every $s \in S$ and $\epsilon^{(m)}$, $m = 1, ..., M$.

c. The approximation for the expected value function is computed by taking the average of value functions over simulated choice shocks as follows.

$$\widehat{E}_{\epsilon'} \left[V^{(t)}(s', \epsilon', \theta) \right] \equiv \frac{1}{M} \sum_{m=1}^{M} V^{(t)}(s', \epsilon^{(m)}, \theta) \quad (4)$$

Steps b) and c) have to be done repeatedly for every state space point $s \in S$. Furthermore, all three steps have to be repeated until the value function converges. That is, for a small $\delta > 0$,

$$\left| V^{(t)}(s, \epsilon^{(m)}, \theta) - V^{(t-1)}(s, \epsilon^{(m)}, \theta) \right| < \delta$$

for all $s \in S$ and $m = 1, .., M$.

2. **Likelihood Construction**

The important increment of the likelihood is the choice probability $P\left[a = a_{i,\tau}^d | s_{i,\tau}^d, V, \theta\right]$. For example, suppose that the per period return function is specified as follows.

$$U(s, a, \epsilon_a^{(m)}, \theta) = \widetilde{U}(s, a, \theta) + \epsilon_a^{(m)},$$

where $\widetilde{U}(s, a, \theta)$ is the deterministic component of the per period return function. Also, denote,

$$\widetilde{\mathcal{V}}^{(t)}(s, a, \theta) = \widetilde{U}(s, a, \theta) + \beta \widehat{E}_{\epsilon'} \left[V^{(t-1)}(s', \epsilon', \theta) \right]$$

to be the deterministic component of the value of choosing action a. Then,

$$P\left[a_{i,\tau}^d | s_{i,\tau}^d, V, \theta\right] = P\left[\epsilon_a - \epsilon_{a_{i,\tau}^d} \le \widetilde{\mathcal{V}}^{(t)}(s, a_{i,\tau}^d, \theta) - \widetilde{\mathcal{V}}^{(t)}(s, a, \theta); a \ne a_{i,\tau}^d | s_{i,\tau}^d, V, \theta\right]$$

which becomes a multinomial probit specification when the error term ϵ is assumed to follow a joint normal distribution.

3. **Maximization routine**

   Supppose we have K parameters to estimate. In a typical Maximum Likelihood estimation routine, where one uses Newton hill climbing algorithm, at iteration t, likelihhood is derived under the original parameter vector $\theta^{(t)}$ and under the perturbed parameter vector $\theta^{(t)} + \Delta\theta_j$, $j = 1, ..., K$. The perturbed likelihood is used together with the original likelihood to derive the new direction of the hill climbing algorithm. This is done to derive the parameters for the iteration $t+1$, $\theta^{(t+1)}$. That is, during a single ML estimation routine, the DP problem needs to be solved in full $K+1$ times. Furthermore, often the ML estimation routine has to be repeated many times until convergence is achieved. During a single iteration of the maximization routine, the inner loop algorithm needs to be executed at least as many times as the number of parameters plus one. Since the estimation requires many iterations of the maximization routine, the entire algorithm is usually computationally extremely burdensome.

2.2 The conventional Bayesian MCMC estimation

A major computational issue in Bayesian estimation method is that the posterior distribution, given by equation 2, is a high-dimensional and complex function of the parameters. Instead of directly simulating the posterior, we adopt the Markov Chain Monte Carlo (MCMC) strategy and construct a transition density from current parameter θ to the next iteration parameter θ', $f(\theta, \theta')$, which satisfies, among other more technical conditions, the following equality.

$$P(\theta|Y_{N,T}) = \int f(\theta, \theta') \, P(\theta'|Y_{N,T}) \, d\theta'$$

We simulate from the transition density the sequence of parameters $\left\{\theta^{(s)}\right\}_{s=1}^{t}$, which is known to converge to the correct posterior.

Gibbs Sampling is a popular way of implementing the MCMC strategy discussed above, due to its simplicity. Gibbs sampling algorithm decomposes the transition density $f(\theta, \theta')$ into small blocks, where simulation from each block is straightforward. During each MCMC iteration, we also fill in the missing $\boldsymbol{\epsilon}$ following the Data Augmentation strategy (See Tanner and Wong (1987) for more details of Data Augmentation).

The conventional Bayesian estimation method applied to the DDC model proceeds in the following three main steps.

Dynamic Programming Step: Given parameter vector $\theta^{(t)}$, the Bellman equation, given by equation 1, is iterated until convergence. This solution algorithm for the DP Step is similar to the Maximum Likelihood algorithm discussed above.

Data Augmentation Step: Since data is generated by a discrete choice model, the observed data is $Y_{N,T} \equiv \left\{s^d_{i,\tau}, a^d_{i,\tau}, F^d_{i,\tau}\right\}^{N,T}_{i=1,\tau=1}$, and does not include the latent shock $\boldsymbol{\epsilon} \equiv \{\epsilon_{i,\tau}\}^{N,T}_{i=1,\tau=1}$. In order to 'integrate out' the latent shock, we simulate $\boldsymbol{\epsilon}$ for the next iteration $t+1$. Since the optimal choice is given as $a^d_{i,\tau}$ in the data, we need to simulate $\boldsymbol{\epsilon}^{(t+1)}$ subject to the constraint that for every sample i, τ, given $s^d_{i,\tau}$, $a^d_{i,\tau}$ is the optimal choice. That is,

$$a^d_{i,\tau} = \arg\max_{a \in A} \mathcal{V}(s^d_{i,\tau}, a, \widehat{\epsilon}^{(t+1)}_{a,i,\tau}, \theta^{(t)})$$

where $\widehat{\epsilon}^{(t+1)}_{i,\tau}$ is the data augmented shock for sample i, τ.

Gibbs Sampling Step: Draw the new parameters $\theta^{(t+1)}$ as follows: Suppose the first $j-1$ parameters have been updated ($\theta_1 = \theta^{(t+1)}_1, ..., \theta_{j-1} = \theta^{(t+1)}_{j-1}$) but the remaining $J-j+1$ parameters have not ($\theta_j = \theta^{(t)}_j, ..., \theta_J = \theta^{(t)}_J$). Then, update j th parameter as follows. Let

$$\theta^{(t,-j)} \equiv \left(\theta^{(t+1)}_1, ..., \theta^{(t+1)}_{j-1}, \theta^{(t)}_{j+1}, ..., \theta^{(t)}_J\right).$$

Then,

$$\theta^{(t+1)}_j \sim p^{(t)}\left(\theta^{(t+1)}_j | \theta^{(t,-j)}\right),$$

where

$$p\left(\theta^{(t+1)}_j | \theta^{(t,-j)}\right) \equiv \frac{\pi(\theta^{(t,-j)}, \theta^{(t+1)}_j) L(Y_T | \widehat{\boldsymbol{\epsilon}}^{(t+1)}, \theta^{(t,-j)}, \theta^{(t+1)}_j)}{\int \pi(\theta^{(t,-j)}, \theta_j) L(Y_T | \widehat{\boldsymbol{\epsilon}}^{(t+1)}, \theta^{(t,-j)}, \theta_j) d\theta_j}, \tag{5}$$

and $\widehat{\boldsymbol{\epsilon}}^{(t+1)}$ is the data augmented shock. Let $f\left(\theta^{(t)}, \theta^{(t+1)}\right)$ be the transition function of a Markov chain from $\theta^{(t)}$ to $\theta^{(t+1)}$ at iteration t. Then, given $\theta^{(t)}$, the transition density for the MCMC is as follows.

$$f\left(\theta^{(t)}, \theta^{(t+1)}\right) = \prod_{j=1}^{J} p\left(\theta^{(t+1)}_j | \theta^{(t,-j)}\right) \tag{6}$$

Although MCMC technique overcomes the computational burden of high dimensionality of parameters, that of solving the DP problem still remains. Since the likelihood is a function of the value function, during the estimation algorithm, the DP problem needs to be solved and value function derived at each iteration of the MCMC algorithm. This is a similar problem as discussed in the application of the Maximum Likelihood method.

We now present our algorithm for estimating the parameter vector θ. We call it the Bayesian Dynamic Programming Algorithm. The key innovation of our algorithm is that we solve the dynamic programming problem and estimate the parameters simultaneously, rather than sequentially.

2.3 The Bayesian Dynamic Programming Estimation

Our method is similar to the conventional Bayesian algorithm in that we construct a transition density $f^{(t)}(\theta, \theta')$, from which we simulate the sequence of parameters $\left\{\theta^{(s)}\right\}_{s=1}^{t}$ such that it converges to the correct posterior. We use Gibbs Sampling strategy described above. We also fill in the missing $\boldsymbol{\epsilon}$ following the Data Augmentation strategy. The main difference between the Bayesian DP algorithm and the conventional algorithm is that during each MCMC step, we do not solve the DP problem in full. In fact, during each MCMC step, we iterate the DP algorithm only once. As a result of this, because the transition density at iteration t depends on the value function approximation derived at iteration t, $V^{(t)}$, in our algorithm, the transition density $f^{(t)}(\theta, \theta')$ changes with each iteration. Therefore, the invariant distribution for the transition density $f^{(t)}(\theta, \theta')$ is,

$$P^{(t)}(\theta|Y_{N,T}) = \pi(\theta)L^{(t)}(Y_{N,T}|\theta) = \pi(\theta)L(Y_{N,T}|\theta, V^{(t)}).$$

That is, the transition density at iteration t satisfies the following equation.

$$P^{(t)}(\theta|Y_{N,T}) = \int f^{(t)}(\theta, \theta')\, P^{(t)}(\theta'|Y_{N,T})\, d\theta'$$

We later prove that the transition density at iteration t converges to the true transition density in probability as $t \to \infty$. That is,

$$f^{(t)}(\theta, \theta') \to f(\theta, \theta')$$

for any $\theta, \theta' \in \Theta$. Furthermore, we prove that the parameter simulations based on the MCMC algorithm using the above sequence of transition densities converges in probability to the parameter simulation generated by the MCMC using the true transition density $f(.,.)$. In other words, the sequence of simulated parameters of the Bayesian DP algorithm converges to the true posterior distribution.

A key issue in solving the DP problem is the way the expected value function (or the Emax function) is approximated. In conventional methods, this approximation is given by equation 4. In contrast, we approximate the emax function by averaging over a subset of past iterations. Let $\Omega^{(t)} \equiv \left\{\epsilon^{(s)}, \theta^{(s)}, V^{(s)}\right\}_{s=1}^{t}$ be the history of shocks, parameters and value functions up to the current iteration t[4]. Let $\mathcal{V}^{(t)}(s, a, \epsilon_a^{(t)}, \theta^{(t)}, \Omega^{(t-1)})$ be the value of choice a and let $V^{(t)}(s, \epsilon^{(t)}, \theta^{(t)}, \Omega^{(t-1)})$ be the value function derived at iteration t of our solution/estimation algorithm. Then, the value function and the approximation $\hat{E}{}_{\epsilon'}^{(t)}\left[V(s', \epsilon', \theta)|\Omega^{(t-1)}\right]$ for the expected value function $E_{\epsilon'}\left[V(s', \epsilon', \theta)\right]$ at iteration t are defined recursively as follows.

$$\hat{E}{}_{\epsilon'}^{(t)}\left[V(s', \epsilon', \theta)|\Omega^{(t-1)}\right] \equiv \sum_{n=1}^{N(t)} V^{(t-n)}(s', \epsilon^{(t-n)}, \theta^{(t-n)} \mid \Omega^{(t-n-1)}) \frac{K_h(\theta^{(t)} - \theta^{(t-n)})}{\sum_{k=1}^{N(t)} K_h(\theta^{(t)} - \theta^{(t-k)})}, \quad (7)$$

and,

$$\mathcal{V}^{(t-n)}(s, a, \epsilon_a^{(t-n)}, \theta^{(t-n)}, \Omega^{(t-n-1)}) = U(s, a, \epsilon_a^{(t-n)}, \theta^{(t-n)}) + \beta \hat{E}{}_{\epsilon'}^{(t-n)}\left[V(s', \epsilon', \theta^{(t-n)})|\Omega^{(t-n-1)}\right],$$

$$V^{(t-n)}(s, \epsilon^{(t-n)}, \theta^{(t-n)} \mid \Omega^{(t-n-1)}) = Max_{a \in A} \mathcal{V}^{(t-n)}(s, a, \epsilon_a^{(t-n)}, \theta^{(t-n)} \mid \Omega^{(t-n-1)})$$

where $K_h()$ is a kernel with bandwidth $h > 0$. That is,

$$K_h(z) = \frac{1}{h^J} K(\frac{z}{h}).$$

K is a nonnegative, continuous, bounded real function which is symmetric around zero and integrates to one. i.e. $\int K(z)dz = 1$. Furthermore, we assume that $\int zK(z)du < \infty$.

The approximated expected value function given by equation 5 is the weighted average of value functions of $N(t)$ most recent iterations. The sample size of the average, $N(t)$, increases with t. Futhermore, we let $t - N(t) \to \infty$ as $t \to \infty$. The weights are high for the value functions at iterations with parameters close to the current parameter vector $\theta^{(t)}$. This is similar to the idea of Pakes and McGuire (2002), where the expected value function

[4] The simulated shocks $\epsilon^{(s)}$ are those used for calculating the value function, not those used for data augmentation.

is the average of the past N iterations. In their algorithm, averages are taken only over the value functions that have the same state variable as the current state variable s. In our case, averages are taken over the value functions that have the same state variable as the current state variable s as well as parameters that are close to the current parameter $\theta^{(t)}$. From now on, to simplify the notation, we omit $\Omega^{()}$ from the value functions and the expected value function.

We now describe the complete Bayesian Dynamic Programming algorithm at iteration t. Suppose that $\{\epsilon^{(l)}\}_{l=1}^{t}$, $\left\{\theta^{(l)}\right\}_{l=1}^{t}$ are given and for all discrete $s \in S$, $\left\{V^{(l)}(s, \epsilon^{(l)}, \theta^{(l)})\right\}_{l=1}^{t}$ is also given. Then, we update the value function and the parameters as follows.

1. **Bellman Equation Step**: For all $s \in S$, derive $\hat{E}{}^{(t)}_{\epsilon'}\left[V(s', \epsilon', \theta^{(t)})\right]$ defined above in equation 7. Also, simulate the value function by drawing $\epsilon^{(t)}$ to derive,

$$\mathcal{V}^{(t)}(s, a, \epsilon_a^{(t)}, \theta^{(t)}) = U(s, a, \epsilon_a^{(t)}, \theta^{(t)}) + \beta \hat{E}{}^{(t)}_{\epsilon'}\left[V(s', \epsilon', \theta^{(t)})\right],$$

$$V^{(t)}(s, \epsilon^{(t)}, \theta^{(t)}) = \max_{a \in A} \mathcal{V}^{(t)}(s, a, \epsilon_a^{(t)}, \theta^{(t)}).$$

2. **Data Augmentation Step**: We simulate $\boldsymbol{\epsilon}^{(t+1)}$ subject to the constraint that for every sample (i, τ), given $s_{i,\tau}^d$, $a_{i,\tau}^d$ is the optimal choice. That is,

$$a_{i,\tau}^d = \arg\max_{a \in A} \mathcal{V}^{(t)}(s_{i,\tau}^d, a, \widehat{\epsilon}_{a,i,\tau}^{(t+1)}, \theta^{(t)})$$

where $\widehat{\epsilon}_{i,\tau}^{(t+1)}$ is the data augmented shock for sample (i, τ). This step is the same as that of the conventional Bayesian estimation.

3. **Gibbs Sampling Step**: This step again is very similar to that of the conventional Bayesian estimation. Therefore, we adopt the notation used there. Draw the new parameters $\theta^{(t+1)}$ as follows:

Suppose the first $j-1$ parameters have been updated $(\theta_1 = \theta_1^{(t+1)}, ..., \theta_{j-1} = \theta_{j-1}^{(t+1)})$ but the remaining $J - j + 1$ parameters are not $(\theta_j = \theta_j^{(t)}, ..., \theta_J = \theta_J^{(t)})$. Then, update j th parameter as follows.

$$\theta_j^{(t+1)} \sim p^{(t)}\left(\theta_j^{(t+1)} | \theta^{(t,-j)}\right),$$

where,

$$p^{(t)}\left(\theta_j^{(t+1)}|\theta^{(t,-j)}\right) \equiv \frac{\pi(\theta^{(t,-j)},\theta_j^{(t+1)})L(Y_{N,T}|\widehat{\boldsymbol{\epsilon}}^{(t+1)},\theta^{(t,-j)},\theta_j^{(t+1)},V^{(t)})}{\int \pi(\theta^{(t,-j)},\theta_j)L(Y_{N,T}|\widehat{\boldsymbol{\epsilon}}^{(t+1)},\theta^{(t,-j)},\theta_j,V^{(t)})d\theta_j},$$

and $\widehat{\boldsymbol{\epsilon}}^{(t+1)}$ is the data augmented shock. Then, given $\theta^{(t)}$, the transition density for the MCMC is derived as follows.

$$f^{(t)}\left(\theta^{(t)},\theta^{(t+1)}\right) = \prod_{j=1}^{J} p^{(t)}\left(\theta_j^{(t+1)}|\theta^{(t,-j)}\right)$$

We repeat Steps 1 to 3 until the sequence of the parameter simulations converges to a stationary distribution. In our algorithm, in addition to the Dynamic Programming and Bayesian methods, nonparametric kernel techniques are also used to approximate the value function. Notice that the convergence of kernel based approximation is not based on the large sample size of the data, but based on the number of Bayesian DP iterations.

Note that that the Bellman equation step (Step 1) is only done once during a single estimation iteration. Hence, the Bayesian DP algorithm avoids the computational burden of solving for the DP problem during each estimation step, which involves repeated evaluation of the Bellman equation.

3 Theoretical Results

Next we show that under some mild assumptions, our algorithm generates a sequence of parameters $\theta^{(1)},\theta^{(2)},...$ which converges in probability to the correct posterior distribution.

Assumption 1: Parameter space $\Theta \subseteq R^J$ is compact, i.e. closed and bounded in the Euclidean space R^J.

This is a standard assumption used in proving the convergence of MCMC algorithm. See, for example, McCullogh and Rossi (1994). It is often not necessary but simplifies the proofs.

Assumption 2: For any $s \in S$, $a \in A$, and ϵ, $\theta \in \Theta$, $|U(s,a,\epsilon_a,\theta)| < M_U$ for some $M_U > 0$. Also, $U(s,a,.,.)$ is a continuously differentiable function of ϵ and θ.

Assumption 3: β is known and $\beta < 1$.

Assumption 4: For any $s \in S$, ϵ and $\theta \in \Theta$, $V^{(0)}(s,\epsilon,\theta) < M_I$ for some $M_I > 0$. Furthermore, $V^0(s,.,.)$ is a continuously differentiable function of ϵ and θ.

Assumptions 2 and 3, and 4 jointly make $V^{(t)}(s,\epsilon,\theta)$ and hence $\hat{E}{}^{(t)}_{\epsilon'}[V(s',\epsilon',\theta)]$, $t=1,...$ uniformly bounded, measurable and continuously differentiable function of θ.

Assumption 5: $\pi(\theta)$ is positive and bounded for any $\theta \in \Theta$. Similarly, for any given $\boldsymbol{\epsilon}$, $\theta \in \Theta$ and V uniformly bounded, $L(Y_T|\boldsymbol{\epsilon},\theta,V) > 0$ and bounded.

Assumption 6: The support of ϵ is compact.

Assumption 7: The bandwidth h is a function of N and as $N \to \infty$, $h(N) \to 0$ and $Nh(N)^{2J} \to \infty$.

Assumption 8: For any $\theta \in \Theta$, $\widehat{a} \in A$, $s \in S$, V uniformly bounded,

$$P\left[a=\widehat{a}|s,V,\theta\right] = \Pr\left[\widehat{\epsilon} : \widehat{a} = \arg\max_{a\in A} \mathcal{V}(s,a,\widehat{\epsilon}_a,\theta)\right] > 0.$$

Assumption 9: $N(t)$ is nondecreasing in t, increases at most by one for a unit increase in t, and $N(t) \to \infty$. Furthermore, $t - N(t) \to \infty$. Define the sequence $t(l)$, $\widetilde{N}(l)$ as follows. For some $t > 0$, define $t(1) = t$, and $\widetilde{N}(1) = N(t)$. Let $t(2)$ be such that $t(2) - N(t(2)) = t(1)$. Such $t(2)$ exists from the assumption on $N(t)$. Also, let $\widetilde{N}(2) = N(t(2))$. Similarly, for any $l > 2$, let $t(l+1)$ be such that $t(l+1) - N(t(l+1)) = t(l)$, and let $\widetilde{N}(l+1) = N(t(l+1))$. Assume that there exists a finite constant $A > 0$ such that $\widetilde{N}(l+1) < A\widetilde{N}(l)$.

An example of a sequence that satisfies Assumption 9 is:

$$t(l) \equiv \frac{l(l+1)}{2}, \widetilde{N}(l) = l$$

and,

$$N(t) = l \text{ for } t(l) \leq t < t(l+1).$$

Now, we state the theoretical results of the paper.

Theorem 1 Suppose Assumptions 1 to 9 are satisfied for $V^{(t)}$, π, L, ϵ and θ. Then, the sequence of approximated value functions $V^{(t)}(s,\epsilon,\theta)$ converges in probability uniformly over s, ϵ and θ to $V(s,\epsilon,\theta)$ as $t \to \infty$. Also, $\hat{E}{}^{(t)}_{\epsilon'}[V(s',\epsilon',\theta)]$ converges to $E_{\epsilon'}[V(s',\epsilon',\theta)]$ in probability uniformly over $s' \in S$ and $\theta \in \Theta$.

Proof: See the Appendix.

Corollary 1: Suppose Assumptions 1 to 9 are satisfied. Then Theorem 1 implies that $f^{(t)}(\theta,\theta')$ converges to $f(\theta,\theta')$ in probability uniformly.

Proof: Recall Equations 5 and 6. Since $V^{(t)} \to V$ in probability uniformly, by compactness of Θ and support of ϵ, the result follows.

Theorem 2: Suppose Assumptions 1 to 6 are satisfied for $V^{(t)}$, $t = 1, ...,$ π, L, ϵ and θ. Suppose $\theta^{(t)}, t = 1, ...$ is a Markov chain with the transition density function $f^{(t)}$ which converges to f in probability uniformly as $t \to \infty$. Then, $\theta^{(t)}$ converges to $\widetilde{\theta}^{(t)}$ in probability, where $\widetilde{\theta}^{(t)}$ is a Markov chain with transition density function f.

Proof: See the Appendix.

Corollary 2: The Markov chain with transition function f converges to the true posterior.

Proof of Corollary 2: We need to show that f satisfies the minorization condition: there exists a density function $g(\theta)$, such that $g(\theta) > 0$ for any $\theta \in \Theta$ and such that , $f(\theta, .) \geq \varepsilon_0 g(.)$ (See Rosenthal (1995), or Tierney (1994) for more details). This is very similar to the proof of Lemma 1, which is in the Appendix, and hence is omitted.

By Corollary 2, we can conclude that the distribution of the sequence of parameters $\theta^{(t)}$ generated by the Bayesian DP algorithm converges in probability to the true posterior distribution.

To understand the basic logic of the proofs, suppose that the parameter $\theta^{(t)}$ stays fixed at a value θ^* for all iterations t. Then, equation (5) reduces to,

$$\widehat{E}_{\epsilon'}\left[V(s', \epsilon', \theta^*)\right] = \frac{1}{N(t)} \sum_{n=1}^{N(t)} V^{(t-n)}(s', \epsilon^{(t-n)}, \theta^*).$$

Then, our algorithm boils down to a simple version of the machine learning algorithm discussed in Pakes and McGuire (2001) and Bertsekas and Tsitsiklis (1996). They approximate the expected value function by taking the average over all past value function iterations whose state space point is the same as the state space point $s\prime$. Bertsekas and Tsitsiklis (1996) discuss the convergence issues and show that under some assumptions the sequence of the value functions from the machine learning algorithm converges to the true value function almost surely. The difficulty of proofs lies in extending the logic of the convergence of the machine learning algorithm to the framework of estimation, that is, the case where the parameter vector moves around as well.

3.1 Continuous State Space

So far, we assumed a finite state space with states evolving deterministically. However, the Bayesian DP algorithm can be applied in a straightforward manner to other settings of dynamic discrete choice models, with minor modifications. One example is the Random grid approximation of Rust (1997). There, given continuous state space vector s, action a and parameter θ, the transition function from state vector s to the next period state vector $s\prime$ is defined to be $f(s'|a,s,\theta)$. Then, to estimate the model, the Dynamic Programming part of our algorithm can be modified as follows.

At iteration t, the value of choice a at parameter θ, state vector s, shock ϵ is defined to be as,

$$\mathcal{V}^{(t)}(s,a,\epsilon_a,\theta) = U(s,a,\epsilon_a,\theta) + \beta \hat{E}_{s',\epsilon'}\left[V(s',\epsilon',\theta)\right],$$

where s' is the next period state varible. $\hat{E}_{s',\epsilon'}\left[V(s',\epsilon',\theta)\right]$ is defined to be the approximation for the expected value function. The value function is defined to be as follows.

$$V^{(t)}(s,\epsilon,\theta) = \max_{a\in A} \mathcal{V}^{(t)}(s,a,\epsilon_a,\theta)$$

Conventionally, randomly generated state vector grid points are fixed throughout the solution/estimation algorithm. If we follow this procedure, and let s_m, $m = 1,...,M$ be the random grids that are generated before the start of the solution/estimation algorithm, then, given parameter θ, the expected value function approximation at iteration t of the DP solution algorithm using Rust random grids method would be,

$$\sum_{m=1}^{M} E_\epsilon V^{(t)}(s_m,\epsilon,\theta) \frac{f\left(s_m|a,s,\theta\right)}{\sum_{l=1}^{M} f\left(s_l|a,s,\theta\right)}.$$

Hence, if we were to apply Rust method in our solution-estimation algorithm, the Emax function (i.e., the expected value function) $\hat{E}_{s',\epsilon'}\left[V(s',\epsilon',\theta)\right]$ could be approximated as follows:

$$\begin{aligned}&\hat{E}_{s',\epsilon'}\left[V(s',\epsilon',\theta)\right] \\ \equiv\ &\sum_{m=1}^{M}\left[\sum_{n=1}^{N(t)} V^{(t-n)}(s_m,\epsilon^{(t-n)},\theta^{(t-n)}) \frac{K_h(\theta-\theta^{(t-n)})}{\sum_{k=1}^{N(t)} K_h(\theta-\theta^{(t-k)})}\right] \frac{f\left(s_m|a,s,\theta\right)}{\sum_{l=1}^{M} f\left(s_l|a,s,\theta\right)}.\end{aligned}$$

Notice that in this definition of Emax approximation, the number of grid points remains fixed in each iteration. In contrast, in our Bayesian DP algorithm, random grids can be changed at each solution/estimation iteration. Let $s^{(t)}$ be the random grid point generated at iteration t. Here, $s^{(\tau)}$, $\tau = 1, 2, ...$ are drawn independently from a distribution. Furthermore, let $K_h(.)$ be the kernel function with bandwidth h. Then, the expected value function can be approximated as follows.

$$\begin{aligned}
&\hat{E}_{s',\epsilon'}\left[V(s',\epsilon',\theta)\right] \\
\equiv\ &\sum_{n=1}^{N(t)} V^{(t-n)}(s^{(t-n)},\epsilon^{(t-n)},\theta^{(t-n)})\frac{K_h(\theta-\theta^{(t-n)})f\left(s^{(t-n)}|a,s,\theta^{(t-n)}\right)}{\sum_{k=1}^{N(t)} K_h(\theta-\theta^{(t-k)})f\left(s^{(t-k)}|a,s,\theta^{(t-k)}\right)}
\end{aligned}$$

Notice that unlike Rust (1997), we do not need to fix the random grid points of the state vector throughout the entire estimation exercise. In fact, we could draw a different state vector for each solution/estimation iteration.

In Rust (1997), if the total number of random grids is M, then the number of computations required for each Dynamic Programming iteration is M. Hence, at iteration τ, the number of Dynamic Programming computations that is required is $M\tau$. If a single DP solution step requires τ DP iterations, and each Newton ML step requires K DP solution steps, then, to iterate Newton ML algorithm once, we need to compute a single DP step $M\tau K$ times.

In contrast, in our Bayesian DP algorithm, at iteration t we only draw one state vector $s^{(t)}$ (so that $M = 1$) and only compute the Bellman equation on that state vector. Further, we solve the DP problem only once (so that $\tau = 1$ and $K = 1$). Still, at iteration t, the number of random grid points is $N(t)$, which can be made arbitrarily large when we increase the number of iterations. In other words, in contrast to the Rust method, the accuracy of the Dynamic Programming computation in our algorithm automatically increases with iterations.

Another issue that arises in application of the Rust random grid method is that Rust (1997) assumes that the transition density function $f(s'|a,s,\theta)$ is not degenerate. That is, we cannot use the random grid algorithm if the transition from s to $s\prime$, given a, θ is deterministic. Similarly, it is well known that the random grid algorithm becomes inaccurate if the transition density has a small variance. In these cases, several versions of polynomial based expected value function (emax function) approximation have been used. Keane and Wolpin (1994) approximate the emax function using polynomials of deterministic part of the value functions for each choice and state variable. Imai and Keane (2004) use Chebychev polynomials of state variables. It is known that in some cases, global approximation using polynomials can be numerically unstable and exhibit "wiggling". Here, we propose a kernel based local interpolation approach to Emax function approximation. The main problem behind the local approximation has been the computational burden of having a large number of grid points. As pointed our earlier, in our solution/estimation algorithm, we can make the number of grid points arbitrarily large by increasing the total number of iterations, even though the number of grid points per iteration is one.

The next period state variable, s' is assumed to be a deterministic function of s, a, and θ. That is,

$$s' = s'(s, a, \theta).$$

Let $K_{h_s}(.)$ be the kernel function with bandwidth h_s for the state variable and $K_{h_\theta}(.)$ for the parameter vector θ. Then, $\hat{E}_{s',\epsilon'}\left[V(s', \epsilon', \theta)\right]$ is defined to be as follows.

$$\begin{aligned} & \hat{E}_{\epsilon'}\left[V(s', \epsilon', \theta)\right] \\ \equiv \; & \sum_{n=1}^{N(t)} V^{(t-n)}(s^{(t-n)}, \epsilon^{(t-n)}, \theta^{(t-n)}) \frac{K_{h_s}\left(s' - s^{(t-n)}\right) K_{h_\theta}(\theta - \theta^{(t-n)})}{\sum_{k=1}^{N(t)} K_{h_s}\left(s' - s^{(t-k)}\right) K_{h_\theta}(\theta - \theta^{(t-k)})}. \end{aligned}$$

4 Examples

We estimate a simple dynamic discrete choice model of entry and exit, with firms in competitive environment.[5] The firm is either an incumbent (I) or a potential entrant (O). If the incumbent firm chooses to stay, its per period return is,

$$R_{I,IN}(K_t, \epsilon_t, \theta) = \alpha K_t + \epsilon_{1t},$$

[5]For an estimation exercise based on the model, see Roberts and Tybout (1997).

where K_t is the capital of the firm, $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t})$ is a vector of random shocks, and θ is the vector of parameter values. If it chooses to exit, its per period return is,

$$R_{I,OUT}(K_t, \epsilon_t, \theta) = \delta_x + \epsilon_{2t},$$

where δ_x is the exit value to the firm. Similarly, if the potential entrant chooses to enter, its per period return is,

$$R_{O,IN}(K_t, \epsilon_t, \theta) = -\delta_E + \epsilon_{1t},$$

and if it decides to stay out, its per period return is,

$$R_{O,OUT}(K_t, \epsilon_t, \theta) = \epsilon_{2t}.$$

We assume the random component of the current period returns to be distributed i.i.d normal as follows.

$$\epsilon_{lt} \sim N(0, \sigma_l), \; l = 1, 2$$

The level of capital K_t evolves as follows. If the incumbent firm stays in, then,

$$\ln K_{t+1} = b_1 + b_2 \ln K_t + u_{t+1},$$

where,

$$u_t \,\tilde{}\, N(0, \sigma_u),$$

and if the potential entrant enters,

$$\ln K_{t+1} = b_e + u_{t+1}.$$

Now, consider a firm who is an incumbent at the beginning of period t. Let $V_I(K_t, \epsilon_t, \theta)$ be the value function of the incumbent with capital stock K_t, and $V_O(K_t, \epsilon_t, \theta)$ be the value function of the outsider, who has capital stock 0. The Bellman equation for the optimal choice of the incumbent is:

$$V_I(K_t, \epsilon_t, \theta) = Max\{V_{I,IN}(K_t, \epsilon_t, \theta), V_{I,OUT}(K_t, \epsilon_t, \theta)\}.$$

where,

$$V_{I,IN}(K_t, \epsilon_t, \theta) = R_{I,IN}(K_t, \epsilon_{1t}, \theta) + \beta E_{t+1} V_I(K_{t+1}(K_t, u_{t+1}, \boldsymbol{\theta}), \epsilon_{t+1}, \theta)$$

is the value of staying in during period t. Similarly,

$$V_{I,OUT}(K_t, \epsilon_t, \theta) = R_{I,OUT}(K_t, \epsilon_{2t}, \theta) + \beta E_{t+1} V_O(0, \epsilon_{t+1}, \theta)$$

is the value of exiting during period t . The Bellman equation for the optimal choice of the outsider is:

$$V_O(0, \epsilon_t, \theta) = Max\{V_{O,IN}(0, \epsilon_t, \theta), V_{O,OUT}(0, \epsilon_t, \theta)\}.$$

where,

$$V_{O,IN}(0, \epsilon_t, \theta) = R_{O,IN}(0, \epsilon_{1t}, \theta) + \beta E_{t+1} V_I(K_{t+1}(0, u_{t+1}, \boldsymbol{\theta}), \epsilon_{t+1}, \theta),$$

is the value of entering during period t and,

$$V_{O,OUT}(0, \epsilon_t, \theta) = R_{O,OUT}(0, \epsilon_{2t}, \theta) + \beta E_{t+1} V_O(0, \epsilon_{t+1}, \theta),$$

is the value of staying out during period t. Notice that the capital stock of an outsider is always 0.

The parameter vector θ of the model is $(\delta_x, \delta_E, \alpha, \beta, \sigma_1, \sigma_2, \sigma_u, b_1, b_2, b_e)$.The state variables are the capital stock K, the parameter vector θ and the status of the firm, $\Gamma \in \{I, O\}$, that is, whether the firm is an incumbent or a potential entrant. Notice that capital stock is a continuous state variable with random transition, in contrast to the theoretical framework where the state space was assumed to be finite and the transition function deterministic.

We assume that for each firm, we only observe the capital stock, profit of the firm that stays in and the entry/exit status over T periods. That is, we know,

$$\{K_{i,t}^d, \pi_{i,t}^d, \Gamma_{i,t}^d\}_{i=1,N}^{t=1,T}$$

where,

$$\pi_{i,t}^d = \alpha K_{i,t}^d + \varepsilon_{1t},$$

if the firm stays in.

We assume the prior of the exit value and entry cost to be normally distributed as follows.

$$\delta_x \sim N(\underline{\delta}_x, A_x^{-1}),$$

$$\delta_E \sim N(\underline{\delta}_E, A_E^{-1}),$$

where $\underline{\delta}_x$, $\underline{\delta}_E$ are the prior means and A_x, A_E are the prior precisions (inverse of variance) of the exit value and the entry cost, respectively. Let $\underline{\delta} \equiv (\underline{\delta}_x, \underline{\delta}_E)$ and $A_\delta \equiv (A_x, A_E)$.

For parameters α, b_1, b_2 and b_e, we assume the priors to be uninformative. This implies that the prior precision of each of these parameters, say A_α, A_b, $b \equiv (b_1, b_2, b_e)$, is assumed to be zero. Let $\underline{\alpha}$ and $\underline{b}$ be the prior means of these parameters.

We also assume independent Chi squares prior for the precision of the shocks ϵ_1 and u. That is,

$$\underline{s_1}^2 h_{\epsilon_1} \sim \chi^2(\nu_{\epsilon_1}),$$

where $\underline{s_1}^2$ is a parameter and ν_{ϵ_1} is the degree of freedom. Similarly,

$$\underline{s_u}^2 h_u \sim \chi^2(\nu_u).$$

Finally,

$$\underline{s_\eta}^2 h_\eta \sim \chi^2(\nu_\eta).$$

where $\eta = \epsilon_1 - \epsilon_2$.

Below, we explain the estimation steps in detail.

Bellman Equation Step

In this step, we derive the value function, i.e., $V_\Gamma^{(s)}(K, \epsilon^{(s)}, \theta^{(s)})$ for iteration s.

1) Suppose we have already calculated the approximation for the expected value function, where the expectation is over the choice shock ϵ, that is, $\widehat{E}_\epsilon^{(s)} V_\Gamma(K'(K, u^{(s)}, \theta^{(s)}), \epsilon, \theta^{(s)})$. To further integrate the value function over the capital shock u, we can either use the random grid integration method of Rust (1997) which uses a fixed grid or let the grid size increase over the iterations. Here, we use the Rust method although we conduct experiments for both cases. That is, given that we have drawn M i.i.d. capital stock grids K_m, $m = 1, .., M$ from a given distribution, we take the weighted average as follows,

$$\widehat{E}^{(s)}\left[V_\Gamma(K'(K, u, \theta^{(s)}), \epsilon, \theta^{(s)})\right] = \sum_{m=1}^{M} \widehat{E}_\epsilon^{(s)}\left[V_\Gamma^{(s)}(K_m, \epsilon, \theta^{(s)})\right] \frac{f(K_m|K, \theta^{(s)})}{\sum_{m=1}^{M} f(K_m|K, \theta^{(s)})}.$$

where $f(K_m|K,\theta^{(s)})$ is the capital transition function from K to K_m. In this example, the random grids remain fixed throughout the estimation. Note that if the firm exits or stays out, $K' = 0$ with probability 1. Hence, the expected value function becomes $\widehat{E}_\epsilon^{(s)}\left[V_O(0,\epsilon,\theta^{(s)})\right]$.

2) Draw $\epsilon^{(s)} = (\epsilon_1^{(s)}, \epsilon_2^{(s)})$.

3) Given $\epsilon^{(s)}$ and $\widehat{E}^{(s)}V_\Gamma(K,\epsilon,\theta^{(s)})$, solve the Bellman equation, that is, solve the decision of the incumbent (whether to stay or exit) or of the entrant (whether to enter or stay out) and derive the value function corresponding to the optimal decisions:

$$\begin{aligned} V_\Gamma^{(s)}(K,\epsilon^{(s)},\theta^{(s)}) &= Max\{R_{\Gamma,IN}(K,\epsilon_1^{(s)},\theta^{(s)}) + \beta\widehat{E}^{(s)}\left[V_I(K'(K,u,\theta^{(s)}),\epsilon,\theta^{(s)})\right], \\ &\quad R_{\Gamma,OUT}(K,\epsilon_2^{(s)},\theta^{(s)}) + \beta\widehat{E}^{(s)}\left[V_O(K'(K,u,\theta^{(s)}),\epsilon,\theta^{(s)})\right]\} \end{aligned}$$

Gibbs Sampling and Data Augmentation Step

Here, we describe how the new parameter vector $\theta^{(s+1)}$ is drawn. Let the deterministic values for the incumbent be defined as follows:

$$\overline{V}_{I,IN}(K,\theta^{(s)}) = \alpha^{(s)}K + \beta\widehat{E}^{(s)}\left[V_I(K',\epsilon,\theta^{(s)})\right],$$

and,

$$\overline{V}_{I,OUT}(K,\theta^{(s)}) = \delta_x^{(s)} + \beta\widehat{E}^{(s)}\left[V_O(0,\epsilon,\theta^{(s)})\right].$$

Similarly, for the potential entrant, we define,

$$\overline{V}_{O,IN}(0,\theta^{(s)}) = -\delta_E^{(s)} + \beta\widehat{E}^{(s)}\left[V_I(K',\epsilon,\theta^{(s)})\right],$$

and,

$$\overline{V}_{O,OUT}(0,\theta^{(s)}) = \beta\widehat{E}^{s}\left[V_O(0,\epsilon,\theta^{(s)})\right].$$

Then, at iteration s, we go through the following two steps.

1) **Data Augmentation Step on Entry and Exit choice**: Define current revenue difference net of $\alpha^{(s)}K_{i,t}^d$ by

$$\begin{aligned} r_{i,t}^{(s+1)} &\equiv R_{\Gamma,OUT}(K_{i,t}^d, \epsilon_{2,i,t}, \theta^{(s)}) - R_{\Gamma,IN}(K_{i,t}^d, \epsilon_{1,i,t}, \theta^{(s)}) + \alpha^{(s)}K_{i,t}^d \\ &= \delta_E^{(s)} I(\Gamma_{i,t} = O) + \delta_x^{(s)} I(\Gamma_{i,t} = I) - \epsilon_{1,i,t} + \epsilon_{2,i,t}. \end{aligned}$$

The empirical economist does not observe $r_{i,t}^{s+1}$ directly because he can only obtain data on status of the firm, that is, whether it is an incumbent or not, and through it, the entry-exit choices and profits, not the current revenues themselves. Nonetheless, the empirical economist can indirectly recover $r_{i,t}^{s+1}$ by simulating and augmenting the shock $\eta_{i,t} = \epsilon_{1,i,t} - \epsilon_{2,i,t}$. But the simulation of $\eta_{i,t}$ has to be consistent with the actual choices that the firm makes. That is, if, in the data, the firm i at period t either stays in or enters, that is, $\Gamma_{i,t+1}^d = I$, then draw $\widehat{\eta}_{i,t} = \epsilon_{1,i,t} - \epsilon_{2,i,t}$ such that,

$$\widehat{\eta}_{i,t}^{(s+1)} \geq \overline{V}_{\Gamma,OUT}(K_{i,t}^d, \theta^{(s)}) - \overline{V}_{\Gamma,IN}(K_{i,t}^d, \theta^{(s)}).$$

If, in the data, the firm i either stays out or exits, that is, $\Gamma_{i,t}^d = O$, then draw $\eta_{i,t}$ such that

$$\widehat{\eta}_{i,t}^{(s+1)} < \overline{V}_{\Gamma,OUT}(K_{i,t}^d, \theta^{(s)}) - \overline{V}_{\Gamma,IN}(K_{i,t}^d, \theta^{(s)}).$$

Once the shock $\widehat{\eta}_{i,t}$ is generated, the econometrician can proceed to recover $r_{i,t}^{(s+1)}$ using the following linear relationship.

$$r_{i,t}^{(s+1)} = \delta_E^{(s)} I(\Gamma_{i,t} = O) + \delta_x^{(s)} I(\Gamma_{i,t} = I) - \widehat{\eta}_{i,t}.$$

Data Augmentation Step on Profit: If the firm stays out or exits, then its potential profit is not observable. In that case, we simulate the profit as follows:

$$\pi_{i,t}^{(s+1)} = \alpha^{(s)}K_{i,t}^d + \widehat{\epsilon}_{1,i,t}.$$

We draw $\widehat{\epsilon}_{1,i,t}$ conditional on $\widehat{\eta}_{i,t}$ as follows:

$$\widehat{\epsilon}_{1,i,t}^{(s+1)} = \gamma_1^{(s)} \widehat{\eta}_{i,t} + v_{i,t},$$

where,

$$v_{i,t} \sim N(0, \sigma_v^2),$$

$$\sigma_v^2 = \frac{\sigma_1^{(s)2}\sigma_2^{(s)2}}{\sigma_1^{(s)2} + \sigma_2^{(s)2}},$$

and,

$$\gamma_1 = \frac{\sigma_1^{(s)2}}{\sigma_1^{(s)2} + \sigma_2^{(s)2}}.$$

2) Draw the new parameter vector $\theta^{(s+1)}$ from the posterior distribution.

We denote the stacked matrix $\mathbf{I}$ with row $T(i-1)+t$ defined as follows:

$$\mathbf{I}_{T(i-1)+t} = [I_{i,t}^d(IN), I_{i,t}^d(OUT)],$$

where $I_{i,t}^d(IN) = 1$ if the firm either enters or decides to stay in, and 0 otherwise, and $I_{i,t}^d(OUT) = 1$ if the firm either exits or stays out and 0 otherwise. Similarly, we denote $\mathbf{r}^{(s+1)}$, $\boldsymbol{\pi}^{(s+1)}$ to be the stacked vector of $r_{i,t}^{(s+1)}$ and $\pi_{i,t}^{(s+1)}$.

We draw $\delta^{(s+1)} = \left[\delta_x^{(s+1)}, \delta_E^{(s+1)}\right]'$ conditional on $(\mathbf{r}^{(s+1)}, h_\eta^{(s)})$ as follows.

$$\delta^{(s+1)}|(\mathbf{r}^{(s+1)}, h_\eta^{(s)}) \sim N(\overline{\delta}, \overline{A}_\delta^{-1}),$$

where,

$$\overline{A}_\delta = A_\delta + h_\eta^{(s)}\mathbf{I}'\mathbf{I},$$

and,

$$\overline{\delta} = \left(\overline{A}_\delta\right)^{-1}(A_\delta\underline{\delta} + h_\eta^{(s)}\mathbf{I}'\mathbf{r}^{(s+1)}).$$

We draw the posterior distribution of h_η from the following χ^2 distribution. That is,

$$[\underline{s_\eta}^2 + \sum_{i,t}\widetilde{\eta}_{i,t}^2]h_\eta^{(s+1)}|(\mathbf{w}^{(s+1)}, \delta^{(s+1)}) \sim \chi^2(NT + \underline{\nu}),$$

where $\widetilde{\eta}_{i,t}$ is the “residual”, that is,

$$\widetilde{\eta}_{i,t} = -r_{i,t}^{(s+1)} + \delta_E^{(s+1)}I_{i,t}^d(OUT) + \delta_x^{(s+1)}I_{i,t}^d(IN).$$

The above Gibbs sampling data augmentation steps are an application of McCulloch and Rossi (1994).

Next, we draw $\alpha^{(s+1)}$ conditional on $(\boldsymbol{\pi}^{(s+1)}, h_a^{(s)})$. Denote $k_t = ln(K_t)$, and $\mathbf{k}$ to be the stacked vector of log capital. That is,

$$\mathbf{k} = [k_{11}, k_{12}, ..., k_{1T}, ..., k_{N_d 1}, k_{N_d 2}, ..., k_{N_d T}].$$

Also, let k_0 be the stacked vector without the intial period capital, i.e.

$$\mathbf{k}_0 = [k_{12}, k_{13}, ..., k_{1T}, ..., k_{N_d 2}, k_{N_d 3}, ..., k_{N_d T}],$$

and $\mathbf{k}_{-1}$be the lagged log capital. That is,

$$\mathbf{k}_{-1} = [k_{11}, k_{12}, ..., k_{1T-1}, ..., k_{N_d 1}, k_{N_d 2}, ..., k_{N_d T-1}].$$

- Then, draw $\alpha^{(s+1)}$ from the following normal distribution.

$$\alpha^{(s+1)}|(\boldsymbol{\pi}^{(s+1)}, h_{\epsilon_1}^{(s)}) \ \sim N(\overline{\alpha}, \overline{A}_{\alpha}^{-1}),$$

where,

$$\overline{A}_{\alpha} = A_{\alpha} + h_{\epsilon_1}^{(s)} \mathbf{k}'\mathbf{k},$$

and,

$$\overline{\alpha} = (A_{\alpha})^{-1} \left(A_{\alpha}\underline{\alpha} + h_{\epsilon_1}^{(s)} \mathbf{k}' \boldsymbol{\pi}^{(s+1)}\right).$$

We draw the posterior distribution of h_{ϵ_1} from the following χ^2 distribution. That is,

$$[\underline{s_{\epsilon_1}}^2 + \sum_{i,t} \widetilde{\epsilon}_{1,i,t}^2] h_{\epsilon_1}^{(s+1)} |(\boldsymbol{\pi}^{(s+1)}, \alpha^{(s+1)}) \sim \chi^2(N_d T + \underline{\nu}),$$

where $\widetilde{\epsilon}_{1,i,t}$ is the "residual", that is,

$$\widetilde{\epsilon}_{1i,t} = \pi_{i,t}^{(s+1)} - \alpha^{(s+1)} k_{i,t}.$$

Furthermore, $(\sigma_{\epsilon_2}^{(s+1)})^2$ or $h_{\epsilon_2}^{(s+1)} = (\sigma_{\epsilon_2}^{(s+1)})^{-2}$ can be recovered as follows:

$$(\sigma_{\epsilon_2}^{(s+1)})^2 = (h_{\eta}^{(s+1)})^{-1} - (h_{\epsilon_1}^{(s+1)})^{-1}$$

Next, we draw $b^{(s+1)} = [b_1^{(s+1)}, b_2^{(s+1)}]'$ conditional on $(\mathbf{k}, h_u^{(s)})$ as follows.

$$b^{(s+1)}|(\mathbf{k}, h_u^{(s)}) \sim N(\overline{b}, \overline{A}_b^{-1}),$$

where,

$$\overline{A}_b = A_b + h_u^{(s)} \mathbf{k}_{-1}' \mathbf{k}_{-1},$$

and,

$$\overline{b} = \overline{A}_b^{-1}(A_b \underline{b} + h_u^{(s)} \mathbf{k}_{-1}' \mathbf{k}_0).$$

We draw the posterior distribution of h_u from the following χ^2 distribution. That is,

$$[\underline{s_u}^2 + \sum_{i,t} \widetilde{u}_{i,t}^2] h_u^{(s+1)} | b^{(s+1)} \sim \chi^2(N_d T + \underline{\nu}),$$

where $\widetilde{u}_{i,t}$ is the "residual", that is,

$$\widetilde{u}_{i,t} = k_{i,t}^d - b_1^{(s+1)} - b_2^{(s+1)} k_{i,t-1}^d.$$

Expected Value Function Iteration Step

Next, we update the expected value function for iteration $s+1$.

First, we derive $E_\epsilon^{(s+1)} V_\Gamma(K, \epsilon, \theta^{(s+1)})$.

$$E_\epsilon^{(s+1)} \left[V_\Gamma(K, \epsilon, \theta^{(s+1)}) \right] = \frac{\sum_{l=Max\{s-N(s),1\}}^{s} \left[\frac{1}{M} \sum_{m=1}^{M} V_\Gamma^{(l)}(K, \epsilon_m^{(l)}, \theta^{(l)}) \right] K_h(\theta^{(s+1)} - \theta^{(l)})}{\sum_{l=Max\{s-N(s),1\}}^{s} K_h(\theta^{(s+1)} - \theta^{(l)})},$$

where $K()$ is the kernel function. We adopt the following Gaussian kernel:

$$K_h(\theta^{(s)} - \theta^{(l)}) = (2\pi)^{-L/2} \prod_{j=1}^{J} h_j^{-1} \exp[-\frac{1}{2}(\frac{\theta_j^{(s)} - \theta_j^{(l)}}{h_j})^2].$$

The expected value function is updated by taking the average over those past $N(s)$ iterations where the parameter vector $\theta^{(l)}$ was close to $\theta^{(s+1)}$.

Then, if the firm enters or stays in, the expected value function is as follows.

$$
\begin{aligned}
& \widehat{E}^{(s+1)}\left[V_I(K'(K,u,\theta^{(s+1)}),\epsilon,\theta^{(s+1)})\right] \\
= \; & \widehat{E}^{(s+1)}_{\epsilon,K}\left[V_I(K'(K,u,\theta^{(s+1)}),\epsilon,\theta^{(s+1)})\right] \\
= \; & \sum_{m=1}^{M}\widehat{E}^{(s+1)}_{\epsilon}\left[V_I(K_m,\epsilon,\theta^{(s+1)})\right]\frac{f(K_m|K,\theta^{(s+1)})}{\sum_{m=1}^{M}f(K_m|K,\theta^{(s+1)})}.
\end{aligned}
$$

As discussed before, in principle, only one simulation of ϵ is needed during each solution/estimation iteration. But that requires the number of past iterations for averaging, i.e. $N(s)$ to be large, which adds to computational burden. Instead, in our example, we draw ϵ ten times and take an average. Hence, when we derive the expected value function, instead of averaging past value functions, we average over past average value functions, i.e., $\frac{1}{M}\sum_{m=1}^{M}V_\Gamma(K_m,\epsilon_m^{(j)},\theta^{(j)})$, where $M=10$. This obviously increases the accuracy per iteration, and reduces the need to have a large $N(s)$. That is partly why in the examples below, to have $N(s)$ increase up to 2000 turned out to be sufficient for good estimation performance. Notice that if the firm stays out or exits, then its future capital stock is zero. Therefore, no averaging over capital grid points is required to derive the expected value function, i.e., the emax function is simply $E_\epsilon^{(s+1)}\left[V_O(0,\epsilon,\theta^{(s+1)})\right]$.

After the Bellman equation step, data augmentation step and the expected value function iteration step, we now have the parameter vector $\theta^{(s+1)}$ and the expected value function $E^{(s+1)}V(K,\epsilon,\theta^{(s+1)})$ for $s+1$ th iteration. We repeat these steps to derive iteration $s+2$ in the same way as described above for $s+1$ th iteration.

In the next section, we present the results of several Monte Carlo studies we conducted using our Bayesian DP method. The first experiment is the basic model using the Rust random grid method. The second experiment incorporates observed and unobserved heterogeneity. The third experiment uses the basic model but lets the capital stock grid size increase over iterations and finally, we conduct an experiment in which capital stock evolves deterministically.

5 Simulation and Estimation

Denote the true values of θ by θ^*. Thus $\theta^* = (\delta_E^*, \delta_x^*, \sigma_{_1}^*, \sigma_{_2}^*, \sigma_u^*, \alpha^*, b_1^*, b_2^*, b_e^*, \beta^*)$. We set the following parameters for the above model. $\delta_E^* = 0.4$, $\delta_x^* = 0.4$, $\sigma_{_1}^* = 0.4$, $\sigma_{_2}^* = 0.4$, $\sigma_u^* = 0.4$, $\alpha^* = 0.2$, $b_1^* = 0.2$, $b_2^* = 0.2$, $b_e^* = -1.0$, $\beta^* = 0.9$.

We first solve the DP problem numerically using conventional numerical methods. Next, we generate artificial data based on this DP solution. All estimation exercises are done on a Sun Blade 2000 workstation. Below, we briefly explain how we solved the DP problem to generate the data for the basic model. For the other three experiments, the data generation step is basically similar involving only minor variations. Notice that for data generation, we only need to solve the DP problem once, that is, for a fixed set of parameters. Hence, we took time and made sure that the DP solution is accurate.

We first set the M capital grid points to be equally spaced between 0 and $\overline{K}$, which we set to be 10.0. Assume that we already know the expected value function of the s^{th} iteration for all capital grid points.

$$E_\epsilon V_\Gamma^s(K_m, \epsilon, \theta^*),\ \Gamma \in \{I, O\},\ m = 1, 2, ..., M.$$

Here, K_m $(m = 1, ..., M)$ are grid points.

The following steps are taken to generate the expected value function for the $(s+1)^{th}$ iteration.

Step 1 Given capital stock K, derive,

$$E^{(s)} V_\Gamma(K'(K, u, \theta^*), \epsilon^{(s)}, \theta^*) = \sum_{m=1}^{M} E_\epsilon^{(s)} V_\Gamma(K_m, \epsilon^{(s)}, \theta^*) f(K_m | K, \theta^*).$$

Here $f(K_m|K, \theta^{(s)})$ is the transition probability from K to K_m.

Step 2 Draw the random shocks ϵ_l. Then, for a given capital stock K, calculate

$$\begin{aligned} V_\Gamma^{(s+1)}(K, \epsilon_l, \theta^*) \quad &= \quad Max\{R_{\Gamma,IN}(K, \epsilon_{1l}, \theta^*) + \beta E^{(s)} V_I(K', \epsilon, \theta^*), \\ &\quad\quad R_{\Gamma,OUT}(K, \epsilon_{2l}, \theta^*) + \beta E^{(s)} V_O(0, \epsilon, \theta^*)\} \end{aligned}$$

Step 3 Repeat Step 2, L times and take an average to derive the approximated expected value function, given K, for the next iteration.

$$E_\epsilon^{(s+1)}V_\Gamma(K,\epsilon,\theta^*) = \frac{1}{L}\sum_{l=1}^{L} V_\Gamma^{(s+1)}(K,\epsilon_l,\theta^*).$$

The above steps are taken for all possible capital grid points, $K = K_1,..,K_M$. In our simulation exercise, we set the simulation size L to be 1000. The total number of capital grid points is set to be $M = 100$.

Step 4 Repeat Step 1 to Step 3 until the Emax function converges. That is, for a small δ (in our case, $\delta = 0.00001$),

$$Max_{m=1,..,M}\{E_\epsilon^{(s+1)}V_\Gamma(K_m,\epsilon,\theta^*), E_\epsilon^{(s)}V_\Gamma(K_m,\epsilon,\theta^*)\} < \delta.$$

We simulate artificial data of capital stock, profit and entry/exit choice sequences using the expected value functions derived above. We then estimate the model using the simulated data with our Bayesian DP routine. We do not estimate the discount factor β. Instead, we set it at the true value $\beta^* = 0.9$.

5.1 Experiment 1: Basic Model

We first describe the prior distributions of parameters. The priors are set to be reasonably diffuse in order to keep the influence on the outcome of the estimation exercise to a minimum.

$$\delta_x \sim N(\underline{\delta}_x, A_x^{-1}),\quad \underline{\delta}_x = 0.4, \underline{A}_x = 0.1,$$

$$\delta_E \sim N(\underline{\delta}_E, A_E^{-1}),\quad \underline{\delta}_E = 0.4, \underline{A}_E = 0.1$$

$$\underline{s_{\epsilon_1}}^2 h_{\epsilon_1} \sim \chi^2(\nu_{\epsilon_1}),\ (\underline{s_{\epsilon_1}}^2)^{-1/2} = 0.4,\ v_{\epsilon_1} = 1.$$

$$\underline{s_\eta}^2 h_\eta \sim \chi^2(v_\eta),\ (\underline{s_\eta}^2)^{-1/2} = \sqrt{0.32},\ v_\eta = 1.$$

$$\underline{s_u}^2 h_u \sim \chi^2(\nu_u),\ (\underline{s_u}^2)^{-1/2} = 0.4,\ v_u = 1.$$

The priors for α, b_1, b_2 and b_e are set to be noninformative.

We set the initial guesses of the parameters to be the true parameter values given by θ^*, and the initial guess of the expected value function to be 0. We used the same 100 grid points in each iteration as used in generating the data. The Gibbs sampling was conducted $10,000$ times. The Gibbs sampler for the simulation with sample size $10,000$ ($N = 1, T = 10,000$) is shown in figures 1 to 9. In estimation experiments with this sample size as well as others, the Gibbs sampler converged around $4,000$ iterations. The posterior mean and standard errors from the $(5,001)^{th}$ iteration up to $(10,000)^{th}$ iteration are shown in Table 1. The posterior mean of δ_x and δ_E are estimated to be somewhat away from the true values if the sample size is 2000, but they are estimated to be reasonably close to the true values for the sample size $5,000$ and $10,000$. Overall, we can see that as the sample size increases, the estimated values become closer to the truth, even though there are some exceptions, such as σ_2, the standard error of the revenue shock of being out.

Figures 1 and 2 show the Gibbs sampler output of parameters δ_x and δ_E. Even though the initial guess is set to be the true value, at the start of the Gibbs sampling algorithm, both parameters immediately jump to values very close to zero. Notice that these values are the estimates we should expect to get when we estimate the data generated by a dynamic model using a static model. Because the expected value functions are set to zero initially, the future benefit of being in or out is zero. Hence, if either exit value or entry cost were big in value, then either entry or exit choice would dominate most of the time, and thus the model would not predict both choices to be observed in the data. Notice that with iterations, the estimates of the parameters directly affecting entry and exit choices, such as δ_x and δ_E converge to the true value (see Figures 1 and 2). This is because as we iterate our Bayesian DP algorithm, the expected value functions become closer to the true value functions. Because the future values of entry and exit choices converge to the truth, so do the parameters representing the current benefits and costs of the entry and exit choices, i.e., δ_x and δ_E. This illustrates that our algorithm solves the Dynamic Programming problem and estimates the parameters simulataneously, and not subsequently.

Figure 3 plots the Gibbs sampler output of σ_2, the standard error of the revenue shock of being out. We can see that the parameter estimate frequently wanders off from the truth. Here prior information on σ_2 would greatly help to pin down the posterior. The Gibbs sampler of parameter b_1 is reported in Figure 4. There, we see that it stays closely around the true value from the start.[6] We have also conducted experiments where we set the initial values of the parameters to half the true values and ran the Gibbs sampler. The posterior mean and standard error of the parameters are shown in column 7 of Table 1. As we can see, the results turn out to be hardly different from the original ones. These results confirm the theorems on convergence in section 1 that the estimation algorithm is not sensitive to the initial values.

Table 1: Posterior Means and Standard Errors (standard errors are in parenthesis)

parameter	estimate	estimate	estimate	true value
δ_x	0.7415 (0.0291)	0.3868 (0.0203)	0.3983 (0.0097)	0.4
δ_E	0.6946 (0.0206)	0.4262 (0.0127)	0.4133 (0.0084)	0.4
α	0.2109 (0.0165)	0.1959 (0.0119)	0.1868 (0.0100)	0.2
σ_1	0.4046 (0.0153)	0.4140 (0.0135)	0.4221 (0.0115)	0.4
σ_2	0.3217 (0.1123)	0.3552 (0.0905)	0.2931 (0.0706)	0.4
b_1	0.0980 (0.0216)	0.1081 (0.0127)	0.1010 (0.0090)	0.1
b_2	0.1050 (0.0420)	0.0965 (0.0244)	0.0977 (0.0168)	0.1
b_e	−0.9620 (0.0128)	−0.9857 (0.0087)	−0.9938 (0.0061)	−1.0
σ_u	0.3891 (0.0052)	0.4047 (0.0033)	0.4033 (0.0022)	0.4
sample size	2,000	5,000	10,000	
CPU time	17 min. 20 sec.	40 min.34 sec.	1 hr. 18 min. 19 sec.	

parameter	estimate[7]	true value
δ_x	0.4032 (0.0097)	0.4
δ_E	0.4182 (0.0083)	0.4
α	0.1868 (0.0100)	0.2
σ_1	0.4221 (0.0115)	0.4
σ_2	0.2932 (0.0707)	0.4
b_1	0.1010 (0.0090)	0.1
b_2	0.0977 (0.0168)	0.1
b_e	−0.9938 (0.0061)	−1.0
σ_u	0.4033 (0.0022)	0.4
sample size	10,000	
CPU time	1 hr. 16 min. 51 sec.	

[6] Other parameters also stay close around their true values from the start. To see those figures, please refer to the the following website: http://alcor.concordia.ca/~simai

[7] This column corresponds to different starting values.

5.2 Experiment 2: Random Effects

We now report estimation results of a model that includes observed and unobserved heterogeneities. We assume that the profit coefficient for each firm i, α_i is distributed normally with mean $\mu = 2.0$ and standard error $\sigma_\alpha = 0.04$. The transition equation for capital is,

$$\ln K_{i,t+1} = b_1 X_i^d + b_2 \ln K_{i,t} + u_{i,t+1},$$

where X_i^d is a firm characteristics observable to the econometrician. In our simulation sample, we simulate X_i^d from $N(1.0, 0.04)$. Notice that if we use the conventional simulated ML method to estimate the model, for each firm i we need to draw α_i many times, say M_α times, and for each draw, we need to solve the dynamic programming problem with the constant coefficient for capital transition equation being $b_1 X_i^d$. If the number of firms in the data is N_d, then for a single simulated likelihood evaluation, we need to solve the DP problem $N_d M_\alpha$ times. This process is computationally so demanding that most researchers use only a finite number of types, typically less than 10, as an approximation of the observed heterogeneity and the random effect. The only exceptions are economists who have access to supercomputers or large PC clusters. Since in our Bayesian DP estimation exercise, the computational burden of estimating the dynamic model is roughly equivalent to that of a static model, we can easily accomodate random effects estimation as is shown below.

We set the prior for α_i as follows.

$$\alpha_i | \mu, \tau^2 \sim \ N(\mu, \tau^2),$$

$$\mu \ \sim N(\underline{\mu}, h_a^{-1}),$$

$$\underline{s_\tau}^2 \tau^{-2} \sim \chi^2(\nu_\tau).$$

Then, if we let $\boldsymbol{\alpha}' = (\alpha_1, ..., \alpha_N)$ and e_N the N by 1 vector of $1's$, then, given τ^2, the prior distribution of $\boldsymbol{\alpha}$ is assumed to be as follows.

$$\boldsymbol{\alpha} \ \sim \ N\left(e_N \underline{\mu}, \tau^2 I_N + h_\alpha^{-1} e_N e_N^{'}\right).$$

The rest of the parameters have the same priors as those of the basic model. Let $\theta_{-\alpha}^{(s)}$ be defined as parameters not including α_i. Below, we briefly describe the differences between the earlier estimation routine and the one that involves random effects.

Data Augmentation Step on Entry and Exit choice: For data augmentation, we need to generate,

$$r_{i,t}^{(s+1)} = R_{\Gamma,OUT}(K_{i,t}^d, \epsilon_{2,i,t}, \theta_{-\alpha}^{(s)}, \alpha_i^{(s)}) - R_{\Gamma,IN}(K_{i,t}^d, \epsilon_{1,i,t}, \theta_{-\alpha}^{(s)}, \alpha_i^{(s)}) + \alpha_i^{(s)} K_{i,t}^d.$$

To draw $\eta_{i,t} = \epsilon_{1,i,t} - \epsilon_{2,i,t}$ we follow the same data augmentation steps as in the basic case except for the fact that to evaluate the entry and exit values, we use different α_i for each firm i.

Data Augmentation Step on Profit: If the firm stays out or exits, then its potential profit is not observable. In that case, we simulate the profit:

$$\pi_{i,t} = \alpha_i^{(s)} K_t + \epsilon_{1,i,t}.$$

The only difference from the estimation of the basic model is that the capital coefficient α_i is different for each firm i. We skip discussing the rest of the step because it is the same as before.

Draw the new parameter vector $\theta^{(s+1)}$ from the posterior distribution: The only difference in the estimation procedure is for drawing the posterior of α_i. Let $\boldsymbol{\pi}' = (\pi_{11}, \pi_{12}, ..., \pi_{1T}, ..., \pi_{N_d 1}, ..., \pi_{N_d T})$ and,

$$\mathcal{K} = \begin{bmatrix} K_1 & 0 & \cdots & 0 \\ 0 & K_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & K_N \end{bmatrix}$$

where $K_j = [K_{j1}, K_{j2}, ..., K_{jT}]$. The posterior draw of $\boldsymbol{\alpha}$ for iteration s, $\boldsymbol{\alpha}^{(s+1)}$, can be done from the following distribution.

$$\boldsymbol{\alpha}^{(s+1)} | \left(\boldsymbol{\pi}^{(s)}, K\right) \sim N\left(\overline{\boldsymbol{\alpha}}, \overline{H}_{\boldsymbol{\alpha}}^{-1}\right),$$

where,

$$\overline{H}_{\boldsymbol{\alpha}} = (\sigma_1^{(s)})^{-2} \mathcal{K}' \mathcal{K} + \left(\tau^2 I_N + h_\alpha^{-1} e_N e_N'\right)^{-1},$$

$$\overline{\boldsymbol{\alpha}} = \overline{H}_{\boldsymbol{\alpha}}^{-1} \left[\left(\tau^2 I_N + h_\alpha^{-1} e_N e_N'\right)^{-1} e_N \mu + (\sigma_1^{(s)})^{-2} \mathcal{K}' \boldsymbol{\pi}\right].$$

Initial Conditions Problem

As pointed out by Heckman (1981) and others, the missing initial state vector (that is, the status of the firm and initial capital) is likely to be correlated with the unobserved heterogeneity α_i, which would result in bias of the parameter estimates. To deal with this problem, for each firm i, given parameters $(\theta_{-\alpha}^{(s)}, \alpha_i)$, we simulate the model for 20 initial periods to derive the initial capital and the status of the firm.

One-Step Bellman Equation and Expected Value Function Iteration Step

In contrast to the solution/estimation of the basic model, we solve the one step Bellman equation for each firm i separately. For given K, $\widehat{E}_\epsilon^{(s+1)} V_\Gamma(K, \epsilon, \theta_{-\alpha}^{(s+1)}, \alpha_i^{(s+1)})$ is derived as follows.

$$\widehat{E}_\epsilon^{(s+1)} V_\Gamma(K, \epsilon, \theta_{-\alpha}^{(s+1)}, \alpha_i^{(s+1)})$$
$$= \frac{\sum_{j=Max\{s-N(s),1\}}^{s} \left[\frac{1}{M} \sum_{l=1}^{M} V_\Gamma^{(j)}(K, \epsilon_l^{(j)}, \theta^{(j)}) \right] K_h(\theta_{-\alpha}^{(s+1)} - \theta_{-\alpha}^{(j)}) K_h(\alpha_i^{(s+1)} - \alpha_i^{(j)})}{\sum_{j=Max\{s-N(s),1\}}^{s} K_h(\theta_{-\alpha}^{(s+1)} - \theta_{-\alpha}^{(j)}) K_h(\alpha_i^{(s+1)} - \alpha_i^{(j)})}.$$

The remaining step to derive the expected value function $\widehat{E}^{(s+1)} \left[V_\Gamma(K'(K, u, \theta^{(s+1)}), \epsilon, \theta^{(s+1)}) \right]$ is the same as in Experiment 1.

We set $N(s)$ to go up to 1000 iterations. The one-step Bellman equation is the part where we have an increase in computational burden. But it turns out that the additional burden is far lighter than that of computing the DP problem again for each firm i, and for each simulation draw of α_i as would be done in the Simulated ML estimation strategy.

We set the sample size to be 100 firms for 100 periods, and the Gibbs sampling was conducted $10,000$ times. The Gibbs sampling routine converged after $4,000$ iterations. Table 2 describes the posterior mean and standard errors from the $5,001$ th iteration up to $10,000$ th iteration.

Table 2: Posterior Means and Standard Errors (standard errors are in parenthesis)

parameter	estimate	true value
δ_x	0.3704 (0.0253)	0.4
δ_E	0.3833 (0.0157)	0.4
μ	0.2089 (0.0112)	0.2
τ	0.03763 (0.00364)	0.04
σ_1	0.4031 (0.0117)	0.4
σ_2	0.4019 (0.0811)	0.4
b_1	0.1007 (0.0136)	0.1
b_2	0.1009 (0.0266)	0.1
b_e	-0.9661 (0.0102)	-1.0
σ_u	0.4064 (0.0036)	0.4
sample size	100×100	
CPU time	13 hrs 26 min 29 sec	

Notice that most of the parameters are close to the true values. The computation time is about 13 hours, which roughly corresponds to that required for Bayesian estimation of a reasonably complicated static random effects model.

We also conducted an estimation exercise using the conventional simulated ML routine. For each firm, we simulated α_i a hundred times (i.e. $M_\alpha = 100$). We solved the DP problem using Monte-Carlo integration to integrate over the choice shock ϵ. We set the simulation size for ϵ to be 100. We set the number of capital grid points M_K to be 100. A single likelihood calculation took about 35 minues to compute. Since we took numerical derivatives, in addition to the likelihood evaluation under the original parameter θ, we calculated the likelihood for the 10 parameter perturbations $\theta + \Delta\theta_i$, $i = 1, ..., 10$. Therefore, a single step of the Newton-Raphson method took 11 likelihood calculations. After computing the search direction, we calculated the likelihood twice to derive the step size. The entire computation took us 6 hours and 20 minutes. In this time, Bayesian DP routine would have completed $6,063$ iterations. That is, by the time the conventional ML routine finished its first iteration, the Bayesian DP routine would have already converged.

Another estimation strategy for the simulated ML could be to expand the state variables of the DP problem to include both X and α. Then, we have to assign grid points for the three-dimensional state space points (K, X, α). If we assign 100 grid points per dimension, then we end up having $10,000$ times more grid points than before. Hence, the overall computational burden would be quite similar to the orginal simulated ML estimation strategy.

5.3 Experiment 3: Infinite Random Grids

In Experiment 1, we used the same capital grid points at every iteration. As discussed earlier, instead of fixing the grid points throughout the DP solution/estimation algorithm, we can draw different random grid points for each solution/estimation iteration. Hence, even though per iteration, we only draw a small number of state vector grid points $K_1^{(t)}, ..., K_{M_K}^{(t)}$ (in this example, $M_K = 10$)[8], the number of random grid points can be made arbitrarily large when we increase the number of iterations.

The formula for the expected value function for the firm who stays in or enters is as follows.

$$\begin{aligned}
&\hat{E}^{(t+1)} \left[V_I(K'\left(K, u, \theta^{(t+1)}\right), \epsilon, \theta^{(t+1)}) \right] \\
\equiv\ & \sum_{n=1}^{N(t)} \sum_{m=1}^{M_K} \left[\frac{1}{M_\epsilon} \sum_{j=1}^{M_\epsilon} V_{IN}^{(t-n)} (K_m^{(t-n)}, \epsilon_j^{(t-n)}, \theta^{(t-n)}) \right] \\
& \times \frac{f\left(K_m^{(t-n)} | a, K, \theta^{(t-n)}\right) K_h(\theta^{(t)} - \theta^{(t-n)})}{\sum_{k=1}^{N(t)} \sum_{m=1}^{M_K} f\left(K_m^{(t-k)} | a, K, \theta^{(t-k)}\right) K_h(\theta^{(t)} - \theta^{(t-k)})}
\end{aligned}$$

The formula for the expected value function for either the firm who stays out or the firm who exits is similar to that of example 1, because there is no uncertainty about the future capital stock.

$$\begin{aligned}
&\hat{E}_{\epsilon'}^{(t+1)} \left[V_O(0, \epsilon, \theta^{(t+1)}) \right] \\
\equiv\ & \sum_{n=1}^{N(t)} \left[\frac{1}{M_\epsilon} \sum_{j=1}^{M_\epsilon} V_{OUT}^{(t-n)} (0, \epsilon_j^{(t-n)}, \theta^{(t-n)}) \right] \frac{K_h(\theta^{(t)} - \theta^{(t-n)})}{\sum_{k=1}^{N(t)} K_h(\theta^{(t)} - \theta^{(t-k)})}
\end{aligned}$$

We increase the total number of grid points up to $20,000$ by letting $N(s)$ increase up to $2,000$. Table 3 shows the estimation results. We can see that the estimates parameters are close to the true ones. The entire exercise took about 10 hours.

Table 3: Posterior Means and Standard Errors

[8] In principle, only one random capital grid per iteration is needed. But again, that requires the number of past iterations for averaging, $N(s)$ to be large.

(Standard errors are in parenthesis)

parameter	estimate	true value
δ_x	0.3817 (0.0146)	0.4
δ_E	0.3923 (0.0117)	0.4
α	0.1998 (0.0108)	0.2
σ_1	0.4069 (0.0123)	0.4
σ_2	0.3697 (0.0782)	0.4
b_1	0.1001 (0.0091)	0.1
b_2	0.1028 (0.0174)	0.1
b_e	-0.9836 (0.0061)	-1.0
σ_u	0.4018 (0.0022)	0.4
sample size	$10,000$	
CPU time	9 hrs52 min 42 sec	

5.4 Experiment 4: Continuous State Space with Deterministic Transition

The framework is similar to the basic model in Experiment 1 except for the capital transition of the incumbent, which now is deterministic. Assume that if the incumbent decides to stay in, the next period capital is,

$$K_{t+1} = K_t.$$

If the firm decides to either exit or stay out, then the next period capital is 0, and if it enters, then the next period capital is,

$$\ln(K_{t+1}) = b_1 + u_{t+1},$$

where,

$$u_{t+1} \sim N(0, \sigma_u).$$

Since the state space is continuous, we use $K_1^{(t)}, ..., K_{M_K}^{(t)}$ as grid points. As in the previous experiment, we set $M_K = 10$ but let the grid points grow over iterations. Now, the formula for the expected value function for the incumbent who stays in is as follows.

$$\begin{aligned}
&\hat{E}\left[V_I(K,\epsilon',\theta)\right] \\
\equiv\ & \sum_{n=1}^{N(t)}\sum_{m=1}^{M_K}\left[\frac{1}{M_\epsilon}\sum_{j=1}^{M_\epsilon}V_I^{(t-n)}(K_m^{(t-n)},\epsilon_j^{(t-n)},\theta^{(t-n)})\right] \\
& \frac{K_{h_K}\left(K-K_m^{(t-n)}\right)K_{h_\theta}(\theta-\theta^{(t-n)})}{\sum_{k=1}^{N(t)}\sum_{m=1}^{M_K}K_{h_K}\left(K-K_m^{(t-k)}\right)K_{h_\theta}(\theta-\theta^{(t-k)})},
\end{aligned}$$

where K_{h_K} is the kernel for the capital stock with bandwidth h_K. The expected value function for the entrant is different because unlike the incumbent who stays in, the entrant faces uncertain future capital. Thus, the entrant's expected value function is,

$$\begin{aligned}
&\hat{E}_{K',\epsilon'}\left[V_I(K'(u),\epsilon,\theta)\right] \\
\equiv\ & \sum_{n=1}^{N(t)}\sum_{m=1}^{M_K}\left[\frac{1}{M_\epsilon}\sum_{j=1}^{M_\epsilon}V_I^{(t-n)}(K_m^{(t-n)},\epsilon_j^{(t-n)},\theta^{(t-n)})\right] \\
& \times\frac{f\left(K_m^{(t-n)}|\theta^{(t-n)}\right)K_h(\theta-\theta^{(t-n)})}{\sum_{k=1}^{N(t)}\sum_{m=1}^{M_K}f\left(K_m^{(t-k)}|\theta^{(t-k)}\right)K_h(\theta-\theta^{(t-k)})}.
\end{aligned}$$

The formula for the expected value function for either the firm who stays out or the firm who exits is the same as in the infinite random grids case:

$$\begin{aligned}
&\hat{E}_{\epsilon'}\left[V_O(0,\epsilon,\theta)\right] \\
\equiv\ & \sum_{n=1}^{N(t)}\left[\frac{1}{M_\epsilon}\sum_{j=1}^{M_\epsilon}V_O^{(t-n)}(0,\epsilon_j^{(t-n)},\theta^{(t-n)})\right]\frac{K_h(\theta-\theta^{(t-n)})}{\sum_{k=1}^{N(t)}K_h(\theta-\theta^{(t-k)})}
\end{aligned}$$

We let the number of grid points increase up to $20,000$ over the iterations.

Table 4 shows the estimation results. We can see that the estimates parameters are reasonably close to the truth, except for the standard error of the revenue shock σ_2. The entire exercise took about 6 hours 30 minutes.

Table 4: Posterior Means and Standard Errors

(Standard errors are in parenthesis)

parameter	estimate	true value
δ_x	0.4294 (0.0159)	0.4
δ_E	0.4890 (0.0178)	0.4
α	0.1313 (0.0026)	0.1
σ_1	0.3776 (0.0033)	0.4
σ_2	0.7214 (0.0381)	0.4
b_1	0.2178 (0.0054)	0.2
σ_u	0.3911 (0.0039)	0.4
sample size	$10,000$	
CPU time	6 hrs 28 min 30 sec	

6 Conclusion

In conventional estimation methods of Dynamic Discrete Choice models, such as GMM, Maximum Likelihood or Markov Chain Monte Carlo, at each iteration step, given a new set of parameter values, the researcher first solves the Bellman equation to derive the expected value function, and then uses it to construct the likelihood or moments. That is, during the DP iteration, the researcher fixes the parameter values and does not "estimate". We propose a Bayesian estimation algorithm where the DP problem is solved and parameters estimated at the same time. In other words, we move parameters during the DP solution. This dramatically increases the speed of estimation. We have demonstrated the effectiveness of our approach by estimating a simple dynamic model of discrete entry-exit choice. Even though we are estimating a dynamic model, the required computational time is in line with the time required for Bayesian estimation of static models. The reason for the speed is clear. The computational burden of estimating dynamic models has been high because the researcher has to repeatedly evaluate the Bellman equation during a single estimation routine, keeping the parameter values fixed. We move parameters, i.e. 'estimate' the model after each Bellman equation evaluation. Since a single Bellman equation evaluation is computationally no different from computing a static model, the speed of our estimation exercise, too, is no different from that of a static model.

Another computational obstacle in the estimation of a Dynamic Discrete Choice model is the Curse of Dimensionality. That is, the computational burden increases exponentially with the increase in the dimension of the state space. In our algorithm, even though at each iteration, the number of state space points on which we calculate the expected value function is small, the total number of 'effective' state space points over the entire solution/estimation iteration grows with the number of Bayesian DP iterations. This number can be made arbitrarily large without much additional computational cost. And it is the total number of 'effective' state space points that determines accuracy. Hence, our algorithm moves one step further in overcoming the Curse of Dimensionality. This also explains why our nonparametric approximation of the expected value function works well under the assumption of continuous state space with deterministic transition function of the state variable. In this case, as is discussed in the main body of the paper, Rust (1997) random grid method may face computational difficulties..

It is worth mentioning that since we are locally approximating the expected value function nonparametrically, as we increase the number of parameters, we may face the "Curse of Dimensionality" in terms of the number of parameters to be estimated. So far, in our examples, this issue does not seem to have made a difference. The reason is that most dynamic models specify per period return function and transition functions to be smooth and well-behaved. Hence, we know in advance that the value functions we need to approximate are smooth, hence well suited for nonparametric approximation. Furthermore, the simulation exercises in the above examples show that with a reasonably large sample size, the MCMC simulations are tightly centered around the posterior mean. Hence, the actual multidimensional area where we need to apply nonparametric approximation is small. But in empirical exercises that involve many more parameters, one probably needs to adopt an iterative MCMC strategy where only up to 4 or 5 parameters are moved at once, which is also commonly done in conventional ML estimation.

7 References

Aguirreagabiria, Victor and Pedro Mira (2002) "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models. " *Econometrica*, Vol. 70, pp. 1519 – 1543

Ackerberg, Daniel (2004) "A New Use of Importance Sampling to Reduce Computational

Burden in Simulation Estimation," mimeo. University of Arizona.

Arcidiacono, Peter and John B. Jones (2003) "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm", *Econometrica,* , Vol. 71. No. 3, pp. $933-946$

Bierens, Herman J. (1994) "Topics in Advanced Econometrics", Cambridge University Press.

Erdem, Tulin and Michael P. Keane (1996) "Decision Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets." *Marketing Science*, Vol. $15:1$, pp. $1-20$.

Geweke, John and Michael P. Keane (1995) "Bayesian Inference for Dynamic Discrete Choice Models without the Need for Dynamic Programming." in Mariano, Scheuermann and Weeks (eds.), *Simulation Based Inference and Econometrics: Methods and Applications*, Cambridge University Press.

Geweke, John, Daniel Houser and Michael P. Keane (1998) "Simulation Based Inference for Dynamic Multinomial Choice Models." working paper, University of Arizona.

Haerdle, Wolfgang (1989) "Applied Nonparametric Regression," Cambridge University Press.

Heckman, James, J. (1981) "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. C. F. Manski and D. McFadden. Cambridge, MA. MIT Press, $pp.179-195$

Houser, Daniel (2003) "Bayesian Analysis of Dynamic Stochastic Model of Labor Supply and Savings", *Journal of Econometrics*, Vo. $113:2$, pp. $289-335$.

Imai, Susumu and Michael P. Keane (2003) "Intertemporal Labor Supply and Human Capital Accumulation", forthcoming, *International Economic Review*.

Imai, Susumu and Kala Krishna (2001) "Employment, Dynamic Deterrence and Crime." NBER Working Paper, No. 8281

Keane, Michael P. and Kenneth 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*, Vol. 74 (4), pp. 648 – 72

Keane, Michael P. and Kenneth I. Wolpin (1997) "The Career Decisions of Young Men." *Journal of Political Economy*, Vol. 105, pp. 473 – 521.

Lancaster, Tony (1997) "Exact Structural Inference in Optimal Job Search Models." *Journal of Business and Economic Statistics*, Vol. 15 (2), pp. 165 – 179.

McCullogh, Robert, and Peter Rossi (1994) "An Exact Likelihood Analysis of the Multinomial Probit Model." *Journal of Econometrics,* Vol. 64 (1), pp. 207 – 240.

Pakes, Ariel and Paul McGuire (2001) "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the 'Curse' of Dimensionality.", forthcoming *Econometrica*

Roberts, Mark and James Tybout (1997) "The Decision to Export in Columbia: An Empirical Model of Entry with Sunk Costs." *American Economic Review.*

Rosenthal, Jeffrey, S. (1995) "Minorization Conditions and Convergence Rates for Markov Chain Monte Carlo", *JASA*, Vol. 90, No. 430, pp. 558 – 566

Rust, John (1997) "Using Randomization to Break the Curse of Dimensionality." *Econometrica*, Vol. 55, No. 5, pp. 999 – 1033

Tanner, Martin A. and Wing H. Wong (1987) "The Calculation of Posterior Distributions by Data Augmentation." *Journal of the American Statistical Association*, vol. 82, pp. 528 – 549.

Tierney, Luke (1994) "Markov Chains for Exploring Posterior Distributions", *The Annals of Statistics*, Vol. 22, No. 4, pp. 1701 – 176

Appendix

Proof of Theorem 1

We need to show that for any $s \in S$, ϵ, $\theta \in \Theta$,

$$V^{(t)}\left(s, \epsilon, \theta\right) \xrightarrow{P} V\left(s, \epsilon, \theta\right) \text{ uniformly, as } t \to \infty$$

But since,

$$V^{(t)}(s, \epsilon, \theta) = \max_{a \in A} \mathcal{V}^{(t)}(s, a, \epsilon, \theta), \quad V(s, \epsilon, \theta) = \max_{a \in A} \mathcal{V}(s, a, \epsilon, \theta),$$

it suffices to show that for any $s \in S$, $a \in A$, ϵ, $\theta \in \Theta$,

$$\mathcal{V}^{(t)}\left(s, a, \epsilon, \theta\right) \xrightarrow{P} \mathcal{V}\left(s, a, \epsilon, \theta\right) \text{ as } t \to \infty.$$

Define

$$W_{N(t),h}(\theta, \theta^{(t-n)}) \equiv \frac{K_h(\theta - \theta^{(t-n)})}{\sum_{k=1}^{N(t)} K_h(\theta - \theta^{(t-k)})}.$$

Then, the difference between the true value function of action a and that obtained by the Bayesian Dynamic Programming iteration can be decomposed into 3 parts as follows.

$$\mathcal{V}\left(s, a, \epsilon, \theta\right) - \mathcal{V}^{(t)}\left(s, a, \epsilon, \theta\right)$$

$$\begin{aligned}
&= \beta \left[\int V(s', \epsilon', \theta) dF_{\epsilon'}(\epsilon', \theta) - \sum_{n=1}^{N(t)} V^{(t-n)}(s', \epsilon^{(t-n)}, \theta^{(t-n)}) W_{N(t),h}(\theta, \theta^{(t-n)}) \right] \\
&= \beta \left[\int V(s', \epsilon', \theta) dF_{\epsilon'}(\epsilon', \theta) - \sum_{n=1}^{N(t)} V(s', \epsilon^{(t-n)}, \theta) W_{N(t),h}(\theta, \theta^{(t-n)}) \right] \\
&+ \beta \left[\sum_{n=1}^{N(t)} \left[V(s', \epsilon^{(t-n)}, \theta) - V(s', \epsilon^{(t-n)}, \theta^{(t-n)}) \right] W_{N(t),h}(\theta, \theta^{(t-n)}) \right] \\
&+ \beta \left[\sum_{n=1}^{N(t)} \left[V(s', \epsilon^{(t-n)}, \theta^{(t-n)}) - V^{(t-n)}(s', \epsilon^{(t-n)}, \theta^{(t-n)}) \right] W_{N(t),h}(\theta, \theta^{(t-n)}) \right] \\
&\equiv A_1 + A_2 + A_3
\end{aligned}$$

The kernel smoothing part is difficult to handle because the underlying distribution of $\theta^{(s)}$ has a conditional density function $f^{(s)}(\theta^{(s-1)}, \theta^{(s)})$ (conditional on $\theta^{(s-1)}$), which is a complicated nonlinear function of all the past value functions and the parameters. Therefore, instead of deriving the asymptotic value of $\frac{1}{N(t)}\sum_{k=1}^{N(t)} K_h(\theta - \theta^{(t-k)})$, as is done in standard nonparametric kernel asymptotics, we derive and use its asymptotic lower bound and upper bound. Lemma 1 below is used for the derivation of the asymptotic lower bound. Lemma 2 is used for the derivation of the asymptotic upper bound. Using the results of Lemma 1 and 2, in Lemma 3 we prove that $A_1 \to 0$ and in Lemma 4 $A_2 \to 0$.

Lemma 1**:**There exists a density function $g(\theta)$, such that $g(\theta) > 0$ for any $\theta \in \Theta$ and for any t and θ, there exists ε_0 such that $0 < \varepsilon_0 \leq 1$ and $f^{(t)}(\theta, .) \geq \varepsilon_0 g(.)$.

Proof:

Recall that

$$p^{(t)}\left(\theta_j'|\theta_{-j}\right) \equiv \frac{\pi(\theta_{-j}, \theta_j')L(Y_T|\widehat{\boldsymbol{\epsilon}}, \theta_{-j}, \theta_j', V^{(t)})}{\int \pi(\theta_{-j}, \theta_j')L(Y_T|\widehat{\boldsymbol{\epsilon}}, \theta_{-j}, \theta_j', V^{(t)})d\theta_j'}$$

By assumptions 1 (Compactness of parameter space), 5 (Strict Positivity and Boundedness of π and L), and 6 (Compactness of support of ϵ), and because utility function is uniformly bounded, there exist $\eta_1, \eta_2, M_1, M_2 > 0$, such that for any θ, $\widehat{\epsilon}$, V satisfying the assumptions,

$$\eta_1 < \pi(\theta)L(Y_T|\widehat{\boldsymbol{\epsilon}}, \theta, V) < M_1, \text{ and}$$

$$\eta_2 < \int \pi(\theta_{-j}, \theta_j')L(Y_T|\widehat{\boldsymbol{\epsilon}}, \theta_{-j}, \theta_j', V)d\theta_j' < M_2.$$

Therefore, for any θ_j',

$$\inf_{\widehat{\epsilon}, \theta_{-j}} p^{(t)}\left(\theta_j'|\theta_{-j}\right)$$

exists, is strictly positive and uniformly bounded below by η_1/M_2. Let

$$h\left(\theta_j'\right) \equiv \inf_{\widehat{\epsilon}, \theta_{-j}} p^{(t)}\left(\theta_j'|\theta_{-j}\right).$$

Notice that $h(.)$ is Lebesgue integrable. Now, define,

$$g(\theta) \equiv \prod_{j=1}^{J} \frac{h(\theta_j)}{\int h\left(\tilde{\theta}_j\right) d\tilde{\theta}_j}, \quad \varepsilon_0 = \prod_{j=1}^{J} \int h\left(\tilde{\theta}_j\right) d\tilde{\theta}_j.$$

By construction, $g(\theta)$ is positive and bounded and $\int g(\theta)\, d\theta = 1$. Hence, $g(\theta)$ is a density function. Also, by construction, ε_0 is a positive constant.

Furthermore,

$$\varepsilon_0 g(\theta') = \prod_{j=1}^{J} h\left(\theta'_j\right) \leq \prod_{j=1}^{J} p^{(t)}\left(\theta'_j|\theta_{-j}\right) = f^{(t)}(\theta', \theta).$$

Finally, since both $g()$ and $f^{(t)}(\theta, .)$ are densities and integrate to 1, $0 < \varepsilon_0 \leq 1$.

Lemma 1 implies that the transition density of the parameter process has an important property: regardless of the current parameter values or the number of iterations, every parameter value is visited with a strictly positive probability.

Lemma 2: There exists a density function $\widetilde{g}()$, $\varepsilon_1 \geq 1$ such that $\widetilde{g}(\theta) > 0$ and for any t, for any $\theta \in \Theta$, $\varepsilon_1 \widetilde{g}(.) \geq f^{(t)}(\theta, .)$.

Proof: Using similar logic as in Lemma 1, one can show that for any θ'_j,

$$\sup_{\widehat{\epsilon},\theta_{-j}} p\left(\theta'_j|\theta_{-j}\right)$$

exists and is bounded. Let

$$\widetilde{h}\left(\theta'_j\right) \equiv \sup_{\widehat{\epsilon},\theta_{-j}} p^{(t)}\left(\theta'_j|\theta_{-j}\right)$$

Now, let,

$$\widetilde{g}(\theta) \equiv \prod_{j=1}^{J} \frac{\widetilde{h}(\theta_j)}{\int \widetilde{h}\left(\tilde{\theta}_j\right) d\tilde{\theta}_j}, \quad \varepsilon_1 = \prod_{j=1}^{J} \int \widetilde{h}\left(\tilde{\theta}_j\right) d\tilde{\theta}_j.$$

Then, $\widetilde{g}(\theta)$ and ε_1 satisfy the conditions of the Lemma.

Lemma 2 implies that the transition density is bounded above, the bound being independent of the current parameter value or the number of iterations.

Lemma 3: $A_1 \rightarrow^P 0$, as $t \rightarrow \infty$.

Proof: Recall that,

$$\frac{A_1}{\beta} = \int V(s', \epsilon', \theta) dF_{\epsilon'}(\epsilon', \theta) - \sum_{n=1}^{N(t)} V(s', \epsilon^{(t-n)}, \theta) W_{N(t),h}(\theta, \theta^{(t-n)}).$$

Rewrite it as,

$$\frac{A_1}{\beta} = \frac{\frac{1}{N(t)}\sum_{n=1}^{N(t)}\left(\int V(s',\epsilon',\theta)dF_{\epsilon'}(\epsilon',\theta) - V(s',\epsilon^{(t-n)},\theta)\right)K_h(\theta-\theta^{(t-n)})}{\frac{1}{N(t)}\sum_{k=1}^{N(t)} K_h(\theta-\theta^{(t-k)})}.$$

We show that the numerator goes to zero in probabilty and the denominator is bounded below by a positive number with probability arbitrarily close to one as $t \to \infty$.

Let

$$X_{N(t)n} = \frac{1}{N(t)}\left[\int V(s',\epsilon',\theta)dF_{\epsilon'}(\epsilon',\theta) - V(s',\epsilon^{(t-n)},\theta)\right]K_h(\theta-\theta^{(t-n)}),$$

where $n = 1, ..., N(t)$. Then, because $\epsilon^{(t-n)}$'s are i.i.d. and $\epsilon^{(t-n)} \tilde{} F_{\epsilon'}(\epsilon',\theta)$,

$$E\left[X_{N(t)n}\right] = 0,\ E\left[X_{N(t)n}X_{N(t)m}\right] = 0 \text{ for } n \neq m.$$

Hence,

$$\begin{aligned} Var\left[\sum_{n=1}^{N(t)} X_{N(t)n}\right] &= \sum_{n=1}^{N(t)} Var\left[X_{N(t)n}\right] \\ &\leq \frac{[\sup|K|]^2}{N(t)h(N(t))^{2J}} \sup_{s',\theta} E\left[\int V(s',\epsilon',\theta)dF_{\epsilon'}(\epsilon',\theta) - V(s',\epsilon^{(t-n)},\theta)\right]^2. \end{aligned}$$

Because $N(t)h(N(t))^{2J} \to \infty$ as $t \to \infty$ and $V(s,\epsilon,\theta)$ is assumed to be uniformly bounded, the RHS of the inequality converges to zero. That is, for any $\gamma > 0$, $\delta > 0$, there is $\bar{t}_\gamma$ such that for any $t > \bar{t}_\gamma$, i.e., $N(t) \geq N(\bar{t}_\gamma)$,

$$\frac{[\sup|K|]^2}{\delta^2 N(t)h(N(t))^{2J}} \sup_{s',\theta} E_\epsilon\left[\int V(s',\epsilon',\theta)dF_{\epsilon'}(\epsilon',\theta) - V(s',\epsilon,\theta)\right]^2 < \gamma$$

Hence, from Chebychev Inequality, for any $\theta \in \Theta$,

$$\Pr\left\{\left|\sum_{n=1}^{N(t)} X_{N(t)n} - 0\right| \geq \delta\right\} \leq \gamma \qquad \text{(A1)}$$

This shows that the numerator in $\frac{A_1}{\beta}$ goes to zero in probability uniformly over $\theta \in \Theta$.

We next show that the denominator is bounded below with probability arbitrarily close to one as t goes to infinity. Let

$$R^{(t-n)} \equiv \varepsilon_0 \frac{g\left(\theta^{(t-n)}\right)}{f^{(t-n)}\left(\theta^{(t-n-1)}, \theta^{(t-n)}\right)}. \tag{A2}$$

Then, from Lemma 1, $0 \leq R^{(t-n)} \leq 1$ and $0 \leq \varepsilon_0 \leq 1$. Also, define a random variable $Y^{(t-n)}$ as follows.

$$Y^{(t-n)} = \begin{cases} K_h\left(\theta - \theta^{(t-n)}(f^{(t-n)})\right) & \text{with probability } R^{(t-n)} \\ 0 & \text{with probability } 1 - R^{(t-n)} \end{cases} \tag{A3}$$

where $\theta^{(t-n)}(f^{(t-n)})$ means that $\theta^{(t-n)}$ has density $f^{(t-n)}\left(\theta^{(t-n-1)}, \theta^{(t-n)}\right)$ conditional on $\theta^{(t-n-1)}$. Then, $Y^{(t-n)}$ is a mixture of 0 and $K_h\left(\theta - \theta^{(t-n)}(g)\right)$, with the mixing probability being $1 - \varepsilon_0$ and ε_0. That is,

$$Y^{(t-n)} = \begin{cases} K_h\left(\theta - \theta^{(t-n)}(g)\right) & \text{with probability } \varepsilon_0 \\ 0 & \text{with probability } 1 - \varepsilon_0 \end{cases} \tag{A4}$$

Further, from the construction of $Y^{(t-n)}$,

$$Y^{(t-n)} \leq K_h\left(\theta - \theta^{(t-n)}(f^{(t-n)})\right).$$

Now, because $\theta^{(t-n)}(g)$, $n = 1, ..., N(t)$ are i.i.d., from equation $A4$, following Bierens (1994), section 10.1, we derive,

$$\begin{aligned} E\left[Y^{(t-n)}\right] &= \varepsilon_0 E\left[K_h\left(\theta - \theta^{(t-n)}(g)\right)\right] = \varepsilon_0 \int \frac{1}{h^J} K\left(\frac{\theta - \theta^{(t-n)}}{h}\right) g\left(\theta^{(t-n)}\right) d\theta^{(t-n)} \\ &= \varepsilon_0 \int g\left(\theta - hz\right) K\left(z\right) dz \rightarrow \varepsilon_0 g\left(\theta\right) \text{ as } h \rightarrow 0 \end{aligned}$$

and

$$
\begin{aligned}
& N(t)h^J Var\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} Y^{(t-n)}\right] \\
= & \frac{h^J}{N(t)}\sum_{n=1}^{N(t)} Var\left[Y^{(t-n)}\right] \\
= & h^J E\left[Y^{(t-n)2}\right] - h^J E\left[Y^{(t-n)}\right]^2 \\
= & E\left[\frac{\varepsilon_0}{h^J} K\left(\frac{\theta-\theta^{(t-n)}(g)}{h}\right)^2\right] - h^J\varepsilon_0^2 E\left[\frac{1}{h^J}K\left(\frac{\theta-\theta^{(t-n)}(g)}{h}\right)\right]^2 \\
= & \varepsilon_0 \int g\left(\theta - hz\right) K\left(z\right)^2 dz - h^J\left[\varepsilon_0\int g\left(\theta-hz\right)K\left(z\right)dz\right]^2 \\
\leq & \varepsilon_0 \sup_{\theta\in\Theta}\int g\left(\theta-hz\right)K\left(z\right)^2 dz \to \varepsilon_0 \sup_{\theta\in\Theta} g\left(\theta\right)\int K\left(z\right)^2 dz
\end{aligned}
$$

as $h \to 0$. Since $N(t)h^J \to \infty$, $Var\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} Y^{(t-n)}\right] \to 0$ as $t \to \infty$. Then, by Chebyshev's inequality,

$$\frac{1}{N(t)}\sum_{n=1}^{N(t)} Y^{(t-n)} \xrightarrow{P} \varepsilon_0 g(\theta).$$

Therefore, for any $\eta_1 > 0$, $\eta_2 > 0$, there exists $\overline{t} > 0$, $\overline{N} \equiv N(\overline{t})$ such that for any $t > \overline{t}$, i.e, $N(t) > \overline{N}$,

$$\Pr\left[\left|\frac{1}{N(t)}\sum_{n=1}^{N(t)} Y^{(t-n)} - \varepsilon_0 g\left(\theta\right)\right| \leq \eta_1\right] > 1-\eta_2.$$

That is,

$$\Pr\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} Y^{(t-n)} + \eta_1 \geq \varepsilon_0 \inf_{\theta\in\Theta} g\left(\theta\right)\right] > 1-\eta_2$$

Now, choose $\eta_1 < \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g\left(\theta\right)$. Then,

$$\Pr\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} Y^{(t-n)} > \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g\left(\theta\right)\right] > 1-\eta_2. \tag{A5}$$

Since $\sum_{n=1}^{N} K_h\left(\theta-\theta^{(t-n)}(f^{(t-n)})\right) \geq \sum_{n=1}^{N} Y^{(t-n)}$, we conclude that for any $\eta_2 > 0$, there exists $\bar{t}_\eta > 0$, $\overline{N} \equiv N(\bar{t}_\eta)$ such that for any $t > \bar{t}_\eta$, i.e, $N(t) > \overline{N}$,

$$\Pr\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h\left(\theta-\theta^{(t-n)}(f^{(t-n)})\right) > \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)\right] > 1-\eta_2. \tag{A6}$$

for any $\theta \in \Theta$. From $A1$ and $A6$, we can see that for $\bar{t} = \max\{\bar{t}_\gamma, \bar{t}_\eta\} > 0$, $\overline{N} \equiv N(\bar{t})$, the following holds: for any $t > \bar{t}$, i.e, $N(t) > \overline{N}$

$$\begin{aligned}
&\Pr\left[\frac{\frac{1}{N(t)}\sum_{n=1}^{N(t)}\left[\int V(s',\epsilon',\theta)dF_{\epsilon'}(\epsilon',\theta) - V(s',\epsilon^{(t-n)},\theta)\right] K_h(\theta-\theta^{(t-n)})}{\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h\left(\theta-\theta^{(t-n)}\right)} \leq \frac{\delta}{\frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)}\right] \\
\geq\ & \Pr\left\{\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)}\left[\int V(s',\epsilon',\theta)dF_{\epsilon'}(\epsilon',\theta) - V(s',\epsilon^{(t-n)},\theta)\right] K_h(\theta-\theta^{(t-n)}) \leq \delta\right]\right. \\
& \left. \bigcap \left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h\left(\theta-\theta^{(t-n)}\right) > \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)\right]\right\} \\
\geq\ & 1-\Pr\left\{\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)}\left[\int V(s',\epsilon',\theta)dF_{\epsilon'}(\epsilon',\theta) - V(s',\epsilon^{(t-n)},\theta)\right] K_h(\theta-\theta^{(t-n)}) > \delta\right]\right. \\
& \left. \bigcup \left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h\left(\theta-\theta^{(t-n)}\right) \leq \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)\right]\right\} \\
\geq\ & 1-\Pr\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)}\left[\int V(s',\epsilon',\theta)dF_{\epsilon'}(\epsilon',\theta) - V(s',\epsilon^{(t-n)},\theta)\right] K_h(\theta-\theta^{(t-n)}) > \delta\right] \\
& -\Pr\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h\left(\theta-\theta^{(t-n)}\right) \leq \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)\right] \\
\geq\ & 1-\gamma-\eta_2
\end{aligned}$$

uniformly over Θ. Since $\delta/\left[\frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)\right]$ can be made arbitrarily small by choosing δ small enough, we have shown that

$$A_1 \xrightarrow{P} 0 \text{ as } N(t) \to \infty$$

uniformly over Θ.

Lemma 4: $A_2 \to^P 0$ as $t \to \infty$.

Proof

$$\begin{aligned}
\left|\frac{A_2}{\beta}\right| &\leq \sum_{n=1}^{N(t)} \left|V\left(s', \epsilon^{(t-n)}, \theta\right) - V\left(s', \epsilon^{(t-n)}, \theta^{(t-n)}\right)\right| W_{N(t),h}\left(\theta, \theta^{(t-n)}\right) \\
&= \sum_{n=1}^{N(t)} \left|V\left(s', \epsilon^{(t-n)}, \theta\right) - V\left(s', \epsilon^{(t-n)}, \theta^{(t-n)}\right)\right| W_{N(t),h}\left(\theta, \theta^{(t-n)}\right) \\
&\quad I\left(\left|\theta - \theta^{(t-n)}\right| \leq \delta\right) \\
&\quad + \sum_{n=1}^{N(t)} \left|V\left(s', \epsilon^{(t-n)}, \theta\right) - V\left(s', \epsilon^{(t-n)}, \theta^{(t-n)}\right)\right| W_{N(t),h}\left(\theta, \theta^{(t-n)}\right) \\
&\quad I\left(\left|\theta - \theta^{(t-n)}\right| > \delta\right) \\
&\equiv H_1 + H_2 \qquad \text{(A7)}
\end{aligned}$$

where $\left|\theta - \theta^{(t-n)}\right| \equiv \max_{j\in J}\left|\theta_j - \theta_j^{(t-n)}\right|$ and $\delta > 0$ is arbitrarily set.

Step 1 of Lemma 4: We show that $H_2 \overset{P}{\to} 0$ uniformly over Θ.

Note that

$$H_2 \leq 2\overline{V} \sum_{n=1}^{N(t)} W_{N(t),h}\left(\theta, \theta^{(t-n)}\right) I\left(\left|\theta - \theta^{(t-n)}\right| > \delta\right) \qquad \text{(A8)}$$

where $\overline{V} = \sup_{s,\epsilon,\theta}\left|V(s,\epsilon,\theta)\right|$. Then,

$$RHS \text{ of } (A8) = 2\overline{V} \frac{\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta - \theta^{(t-n)}) I\left(\left|\theta - \theta^{(t-n)}\right| > \delta\right)}{\frac{1}{N(t)}\sum_{k=1}^{N(t)} K_h(\theta - \theta^{(t-k)})}. \qquad \text{(A9)}$$

We first show that the numerator goes to 0 in probability as h goes to 0.

Note that $K_{h(N(t))}(\theta - \theta^{(t-n)}) I\left(\left|\theta - \theta^{(t-n)}\right| > \delta\right) \geq 0$. Hence, from Chebychev Inequality, for any $\eta > 0$,

$$
\begin{aligned}
&\Pr\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)}) I\left(\left|\theta-\theta^{(t-n)}\right|>\delta\right)\geq\eta\right]\\
\leq\ & \frac{1}{\eta} E\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)}) I\left(\left|\theta-\theta^{(t-n)}\right|>\delta\right)\right]. \qquad \text{(A10)}
\end{aligned}
$$

From Lemma 2, there exists $\varepsilon_1 > 0$ such that for any s, $\theta^{(s-1)}$, $\theta \in \Theta$

$$
\varepsilon_1 \widetilde{g}(\theta) \geq f^{(s)}\left(\theta^{(s-1)}, \theta\right).
$$

Hence,

$$
\begin{aligned}
&E\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)}(f^{(t-n)})) I\left(\left|\theta-\theta^{(t-n)}(f^{(t-n)})\right|>\delta\right)\right]\\
\leq\ & \varepsilon_1 E\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)}\left(\widetilde{g}\right)) I\left(\left|\theta-\theta^{(t-n)}\left(\widetilde{g}\right)\right|>\delta\right)\right].
\end{aligned}
$$

Since $\theta^{(t-n)}\left(\widetilde{g}\right)$, $n = 1, 2.., N(t)$, are i.i.d., we have,

$$
\begin{aligned}
&E\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)}(f^{(t-n)})) I\left(\left|\theta-\theta^{(t-n)}(f^{(t-n)})\right|>\delta\right)\right]\\
\leq\ & \varepsilon_1 E\left[\frac{1}{h^J} K\left(\frac{\theta-\theta^{(t-n)}\left(\widetilde{g}\right)}{h}\right) I\left(\left|\theta-\theta^{(t-n)}\left(\widetilde{g}\right)\right|>\delta\right)\right]\\
=\ & \varepsilon_1 \int_{\left|\theta-\widetilde{\theta}\right|>\delta} \frac{1}{h^J} K\left(\frac{\theta-\widetilde{\theta}}{h}\right) \widetilde{g}\left(\widetilde{\theta}\right) d\widetilde{\theta} \qquad \text{(A11)}
\end{aligned}
$$

Now, by change of variables,

$$
\begin{aligned}
\int_{\left|\theta-\widetilde{\theta}\right|>\delta} \frac{1}{h^J} K\left(\frac{\theta-\widetilde{\theta}}{h}\right) g\left(\widetilde{\theta}\right) d\widetilde{\theta} &= \int_{|z|>\frac{\delta}{h}} K(z)\, g(\theta-hz)\, dz\\
&\leq \sup_{\theta\in\Theta} g(\theta) \int_{|z|>\frac{\delta}{h}} K(z)\, dz \qquad \text{(A12)}
\end{aligned}
$$

Because $\int K(z)\,dz = 1$, $\int_{|z|>\frac{\delta}{h}} K(z)\,dz \to 0$ as $h \to 0$. Furthermore, $g()$ is bounded by construction. Hence, The RHS of $A12$ converges to zero as h goes to zero. Therefore, RHS of $A11$ converges to zero as h converges to zero uniformly over Θ, and thus,

$$\Pr\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)})I\left(\left|\theta-\theta^{(t-n)}\right|>\delta\right)\geq\eta\right]\to 0 \tag{A13}$$

as $h \to 0$ uniformly over Θ. Note that η can be made arbitrarily small.

From $(A6)$, we know that the denominator in $A9$ is bounded below uniformly over $\theta \in \Theta$ by $\frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)$ with probability arbitrarily close to 1 as t goes to infinity. Thus, using similar steps as in Lemma 3, the result follows. That is, $H_2 \xrightarrow{P} 0$ as $t \to \infty$ uniformly over Θ.

Step 2 of Lemma 4: Show that $H_1 \xrightarrow{P} 0$ as $t \to \infty$, uniformly over Θ.

Define $L \equiv \sup_{j\in J, s\in S,\epsilon,\theta\in\Theta}\left|\frac{\partial V(s,\epsilon,\theta)}{\partial\theta}\right|$. Then, from the Intermediate Value Theorem,

$$\begin{aligned}
&\sum_{n=1}^{N(t)}\left|V\left(s',\epsilon^{(t-n)},\theta\right)-V\left(s',\epsilon^{(t-n)},\theta^{(t-n)}\right)\right|W_{N(t),h}\left(\theta,\theta^{(t-n)}\right)I\left(\left|\theta-\theta^{(t-n)}\right|\leq\delta\right)\\
\leq\;&\sum_{n=1}^{N(t)}L\left|\theta-\theta^{(t-n)}\right|W_{N(t),h}\left(\theta,\theta^{(t-n)}\right)I\left(\left|\theta-\theta^{(t-n)}\right|\leq\delta\right)\\
\leq\;&L\delta\sum_{n=1}^{N(t)}W_{N(t),h}\left(\theta,\theta^{(t-n)}\right)I\left(\left|\theta-\theta^{(t-n)}\right|\leq\delta\right)\leq L\delta\sum_{n=1}^{N(t)}W_{N(t),h}\left(\theta,\theta^{(t-n)}\right)=L\delta
\end{aligned}$$

which can be made arbitrarily small by choosing small enough $\delta > 0$.

From Step 1 of Lemma 4, we already know that given arbitrary $\delta > 0$, $H_2 \xrightarrow{P} 0$ as $t \to \infty$ uniformly over Θ. Hence it follows that $A_2 \xrightarrow{P} 0$ as $t \to \infty$ uniformly over Θ.

Now, we return to the proof of Theorem 1. That is, we need to show that

$$\mathcal{V}(s,a,\epsilon,\theta)-\mathcal{V}^{(t)}(s,a,\epsilon,\theta)\xrightarrow{P}0 \quad \text{as } t\to\infty$$

Define $A^{(t)}$ to be as follows:

$$A^{(t)}(\theta)\equiv A_1+A_2$$

From Lemma 3 and Lemma 4, we know that,

$$A^{(t)}(\theta)\xrightarrow{P}0, as\ t\to\infty,$$

uniformly over Θ. Therefore,

$$A^{(t)}(\theta^{(t)}) \xrightarrow{P} 0 \; as \; t \to \infty$$

Now,

$$\mathcal{V}\left(s,a,\epsilon,\theta\right) - \mathcal{V}^{(t)}\left(s,a,\epsilon,\theta\right) = A^{(t)}(\theta)$$
$$+\beta\left[\sum_{n=1}^{N(t)}\left[V(s',\epsilon^{(t-n)},\theta^{(t-n)}) - V^{(t-n)}(s',\epsilon^{(t-n)},\theta^{(t-n)})\right]W_{N(t),h}(\theta,\theta^{(t-n)})\right] \quad \text{(A14)}$$

Notice that if $V\left(s,\epsilon,\theta\right) \geq V^{(t)}\left(s,\epsilon,\theta\right)$, then

$$\begin{aligned} 0 &\leq V\left(s,\epsilon,\theta\right) - V^{(t)}\left(s,\epsilon,\theta\right) = Max_{a\in A}\mathcal{V}\left(s,a,\epsilon,\theta\right) - Max_{a\in A}\mathcal{V}^{(t)}\left(s,a,\epsilon,\theta\right) \\ &\leq Max_{a\in A}\left[\mathcal{V}\left(s,a,\epsilon,\theta\right) - \mathcal{V}^{(t)}\left(s,a,\epsilon,\theta\right)\right] \leq Max_{a\in A}\left|\mathcal{V}\left(s,a,\epsilon,\theta\right) - \mathcal{V}^{(t)}\left(s,a,\epsilon,\theta\right)\right| \end{aligned}$$

Similarly, if $V\left(s,\epsilon,\theta\right) \leq V^{(t)}\left(s,\epsilon,\theta\right)$, then

$$\begin{aligned} 0 &\leq V^{(t)}\left(s,\epsilon,\theta\right) - V\left(s,\epsilon,\theta\right) = Max_{a\in A}\mathcal{V}^{(t)}\left(s,a,\epsilon,\theta\right) - Max_{a\in A}\mathcal{V}\left(s,a,\epsilon,\theta\right) \\ &\leq Max_{a\in A}\left[\mathcal{V}^{(t)}\left(s,a,\epsilon,\theta\right) - \mathcal{V}\left(s,a,\epsilon,\theta\right)\right] \leq Max_{a\in A}\left|\mathcal{V}\left(s,a,\epsilon,\theta\right) - \mathcal{V}^{(t)}\left(s,a,\epsilon,\theta\right)\right| \end{aligned}$$

Hence, taking supremum over $s\prime$ on the right hand side of $A14$and then taking absolute values on both sides, we obtain:

$$\begin{aligned} &\left|V\left(s,\epsilon,\theta\right) - V^{(t)}\left(s,\epsilon,\theta\right)\right| \leq Max_{a\in A}\left|\mathcal{V}\left(s,a,\epsilon,\theta\right) - \mathcal{V}^{(t)}\left(s,a,\epsilon,\theta\right)\right| \\ \leq\; &\sup_{s'\in S}\left|A^{(t)}\left(\theta\right)\right| \\ &+\beta\left[\sum_{n=1}^{N(t)}\sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)}) - V^{(t-n)}(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})\right|W_{N(t),h}(\theta,\theta^{(t-n)})\right] \text{(A14')} \end{aligned}$$

Now, $\left|V\left(s,\epsilon,\theta\right) - V^{(t)}\left(s,\epsilon,\theta\right)\right|$ appears on the LHS and $\left|V(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)}) - V^{(t-n)}(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})\right|$ appears on the RHS of equation $A14\prime$. Using this, we can recursively substitute away $\left|V(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)}) - V^{(t-n)}(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})\right|$. This logic is used in the following Lemma.

Lemma 5: For $\tau < t$, let

$$\widehat{W}(t,t,\tau) \equiv \widetilde{W}\left(t,\tau\right) \equiv \beta W_{N(t),h}(\theta^{(t)},\theta^{(\tau)}).$$

Now, for $\underline{N} \geq 1$ and for m such that $0 < m \leq \underline{N} + 1$, define

$$\Psi_m (t + \underline{N}, t, \tau) \equiv \{J_m = (t_m, t_{m-1}, ..., t_1, t_0) : t_m = t + \underline{N} > t_{m-1} > .... > t_2 > t_1 \geq t, t_0 = \tau\}$$

and,

$$\widehat{W}(t + \underline{N}, t, \tau) \equiv \sum_{m=1}^{\underline{N}+1} \left\{ \sum_{\Psi_m (t+\underline{N}, t, \tau)} \prod_{k=1}^{m} \widetilde{W} (t_k, t_{k-1}) \right\}.$$

Then, for any $\underline{N} \geq 1$, $t > 0$,

$$\begin{aligned} & \left| V (s, \epsilon, \theta) - V^{(t+\underline{N})} (s, \epsilon, \theta) \right| \\ \leq \; & \sup_{s' \in S} \left| A^{(t+\underline{N})} (\theta) \right| \\ & + \sum_{m=0}^{\underline{N}-1} \widehat{W} (t + \underline{N}, t + \underline{N} - m, t + \underline{N} - m - 1) \sup_{s' \in S} \left| A^{(t+\underline{N}-m-1)} \left(\theta^{(t+\underline{N}-m-1)} \right) \right| \\ & + \sum_{n=1}^{N(t)} \sup_{\widehat{s} \in S} \left| V(\widehat{s}, \epsilon^{(t-n)}, \theta^{(t-n)}) - V^{(t-n)}(\widehat{s}, \epsilon^{(t-n)}, \theta^{(t-n)}) \right| \widehat{W}(t + \underline{N}, t, t - n). \quad \text{(A15)} \end{aligned}$$

Furthermore,

$$\sum_{n=1}^{N(t)} \widehat{W}(t + \underline{N}, t, t - n) \leq \beta \quad \text{(A16)}$$

Proof of Lemma 5.

First, we show that inequality $A15$ and $A16$ hold for $\underline{N} = 1$. For iteration $t + 1$, we get

$$\begin{aligned} & \left| V \left(s, \epsilon, \theta^{(t+1)} \right) - V^{(t+1)} \left(s, \epsilon, \theta^{(t+1)} \right) \right| \\ \leq \; & \sup_{s' \in S} \left| A^{(t+1)} \left(\theta^{(t+1)} \right) \right| \\ & + \sum_{n=1}^{N(t+1)} \sup_{s' \in S} \left| V(s', \epsilon^{(t+1-n)}, \theta^{(t+1-n)}) - V^{(t+1-n)}(s', \epsilon^{(t+1-n)}, \theta^{(t+1-n)}) \right| \\ & \widetilde{W}(t + 1, t + 1 - n) \\ \leq \; & \sup_{s' \in S} \left| A^{(t+1)} \left(\theta^{(t+1)} \right) \right| + \sup_{s' \in S} \left| V(s', \epsilon^{(t)}, \theta^{(t)}) - V^{(t)}(s', \epsilon^{(t)}, \theta^{(t)}) \right| \widetilde{W} (t + 1, t) \\ & + \sum_{n=1}^{N(t+1)-1} \sup_{s' \in S} \left| V(s', \epsilon^{(t-n)}, \theta^{(t-n)}) - V^{(t-n)}(s', \epsilon^{(t-n)}, \theta^{(t-n)}) \right| \widetilde{W}(t + 1, t - n) \end{aligned}$$

Now, we substitute away $\left|V(s',\epsilon^{(t)},\theta^{(t)})-V^{(t)}(s',\epsilon^{(t)},\theta^{(t)})\right|$ by using $A14'$) and the fact that $N(t)\geq N(t+1)-1$,

$$\begin{aligned}
&\left|V\left(s,\epsilon,\theta^{(t+1)}\right)-V^{(t+1)}\left(s,\epsilon,\theta^{(t+1)}\right)\right| \\
\leq\ & \sup_{s'\in S}\left|A^{(t+1)}(\theta^{(t+1)})\right|+\sup_{s'\in S}\left|A^{(t)}(\theta^{(t)})\right|\widetilde{W}(t+1,t) \\
&+\sum_{n=1}^{N(t)}\sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})-V^{(t-n)}(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})\right| \\
&\{\widetilde{W}(t+1,t)\widetilde{W}(t,t-n)+\widetilde{W}(t+1,t-n)\} \\
=\ & \sup_{s'\in S}\left|A^{(t+1)}\left(\theta^{(t+1)}\right)\right|+\sup_{s'\in S}\left|A^{(t)}\left(\theta^{(t)}\right)\right|\widehat{W}(t+1,t+1,t) \\
&+\sum_{n=1}^{N(t)}\sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})-V^{(t-n)}(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})\right|\widehat{W}\left(t+1,t,t-n\right)
\end{aligned}$$

Hence, Inequality $A15$ holds for $\underline{N}=1$.

Furthermore, because $\sum_{n=1}^{N(t)}\widetilde{W}(t,t-n)/\beta=\sum_{n=1}^{N(t)}W_{N(t),h}(\theta^{(t)},\theta^{(t-n)})=1$,

$$\begin{aligned}
&\sum_{n=1}^{N(t)}\widehat{W}\left(t+1,t,t-n\right)=\sum_{n=1}^{N(t)}\widetilde{W}(t+1,t)\widetilde{W}(t,t-n)+\sum_{n=1}^{N(t)}\widetilde{W}(t+1,t-n) \\
=\ & \widetilde{W}(t+1,t)\sum_{n=1}^{N(t)}\widetilde{W}(t,t-n)+\sum_{n=1}^{N(t)}\widetilde{W}(t+1,t-n) \\
=\ & \beta\widetilde{W}(t+1,t)+\sum_{n=1}^{N(t)}\widetilde{W}(t+1,t-n)\leq\sum_{n=1}^{N(t)+1}\widetilde{W}(t+1,t+1-n)
\end{aligned}$$

Since $\widetilde{W}(t+1,t+1-n)=0$ for any $n>N(t+1)$,

$$\begin{aligned}
\sum_{n=1}^{N(t)+1}\widetilde{W}\left(t+1,t+1-n\right) &= \sum_{n=1}^{N(t+1)}\widetilde{W}(t+1,t+1-n) \\
&= \beta\sum_{n=1}^{N(t+1)}W_{N(t+1),h}(\theta^{(t+1)},\theta^{(t+1-n)})=\beta
\end{aligned}$$

Thus,

$$\sum_{n=1}^{N(t)} \widehat{W}\left(t+1, t, t-n\right) \leq \beta \tag{A17}$$

Hence, inequality $A16$ holds for $\underline{N}=1$.

Next, suppose that inequality $A15$ holds for $\underline{N}=M$. Then, using $t+1$ instead of t in inequality $A15$, we get

$$\begin{aligned}
& \left|V\left(s,\epsilon,\theta\right)-V^{(t+1+M)}\left(s,\epsilon,\theta\right)\right| \\
\leq\; & \sup_{s'\in S}\left|A^{(t+1+M)}\left(\theta\right)\right| \\
& +\sum_{m=0}^{M-1}\widehat{W}\left(t+1+M,t+1+M-m,t+M-m\right)\sup_{s'\in S}\left|A^{(t+M-m)}\left(\theta^{(t+M-m)}\right)\right| \\
& +\sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t)},\theta^{(t)})-V^{(t)}(\widehat{s},\epsilon^{(t)},\theta^{(t)})\right|\widehat{W}(t+1+M,t+1,t) \\
& +\sum_{n=2}^{N(t+1)}\sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t+1-n)},\theta^{(t+1-n)})-V^{(t+1-n)}(\widehat{s},\epsilon^{(t+1-n)},\theta^{(t+1-n)})\right| \\
& \widehat{W}(t+1+M,t+1,t+1-n).
\end{aligned}$$

Now, using $A14\prime$ to substitute away $\sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t)},\theta^{(t)})-V^{(t)}(\widehat{s},\epsilon^{(t)},\theta^{(t)})\right|$, we get

$$\begin{aligned}
& \left|V\left(s,\epsilon,\theta\right)-V^{(t+M+1)}\left(s,\epsilon,\theta\right)\right| \\
\leq\; & \sup_{s'\in S}\left|A^{(t+M+1)}\left(\theta\right)\right| \\
& +\sum_{m=0}^{M}\widehat{W}\left(t+M+1,t+M+1-m,t+M-m\right)\sup_{s'\in S}\left|A^{(t+M-m)}\left(\theta^{(t+M-m)}\right)\right| \\
& +\sum_{n=1}^{N(t)}\sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})-V^{(t-n)}(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})\right| \\
& \left[\widehat{W}(t+M+1,t+1,t)\widetilde{W}(t,t-n)+\widehat{W}(t+M+1,t+1,t-n)\right]
\end{aligned} \tag{A18}$$

Now, we claim that, for any $M \geq 1$,

$$\begin{aligned} & \widehat{W}\,(t+M,t+1,t)\,\widetilde{W}(t,t-n)+\widehat{W}(t+M,t+1,t-n) \\ = \;& \widehat{W}\,(t+M,t,t-n) \end{aligned} \tag{A19}$$

Proof of the Claim:

Let

$$\begin{aligned} & \Psi_{m,1}(t+M,t,\tau) \\ \equiv \;& \left\{J_m=(t_m,t_{m-1},...,t_1,t_0): t_m=t+M>t_{m-1}>....>t_2\geq t+1, t_1=t, t_0=\tau\right\}. \end{aligned}$$

Notice that

$$\begin{aligned} & \Psi_{m}(t+M,t+1,\tau) \\ \equiv \;& \left\{J_m=(t_m,t_{m-1},...,t_1,t_0): t_m=t+M>t_{m-1}>....>t_2>t_1\geq t+1, t_0=\tau\right\}. \end{aligned}$$

Then,

$$\Psi_m(t+M,t,\tau)=\Psi_{m,1}(t+M,t,\tau)\cup\Psi_m(t+M,t+1,\tau)$$

and

$$\Psi_{m,1}(t+M,t,\tau)\cap\Psi_m(t+M,t+1,\tau)=\varnothing.$$

Also,

$$\Psi_{M+1}(t+M,t+1,\tau)=\varnothing$$

Therefore,

$$\begin{aligned}
&\widehat{W}(t+M,t,\tau)\\
\equiv\ & \sum_{m=1}^{M+1}\left\{\sum_{\Psi_m(t+M,t,\tau)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\\
=\ & \sum_{m=1}^{M+1}\left\{\sum_{\Psi_{m,1}(t+M,t,\tau)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}+\sum_{m=1}^{M+1}\left\{\sum_{\Psi_m(t+M,t+1,\tau)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\\
=\ & \sum_{m=2}^{M+1}\left\{\sum_{\Psi_{m-1}(t+M,t+1,t)}\prod_{k=1}^{m-1}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\widetilde{W}\left(t,\tau\right)+\sum_{m=1}^{M}\left\{\sum_{\Psi_m(t+M,t+1,\tau)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\\
=\ & \sum_{m=1}^{M}\left\{\sum_{\Psi_m(t+M,t+1,t)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\widetilde{W}\left(t,\tau\right)+\sum_{m=1}^{M}\left\{\sum_{\Psi_m(t+M,t+1,\tau)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\\
=\ & \widehat{W}\left(t+M,t+1,t\right)\widetilde{W}(t,\tau)+\widehat{W}(t+M,t+1,\tau)
\end{aligned}$$

Hence, the claim holds. Substituting this into equation $A17$ yields the first part of the lemma by induction.

Next, suppose that $A16$ holds for $\underline{N}=M$. That is,

$$\sum_{n=1}^{N(t)}\widehat{W}(t+M,t,t-n)\leq\beta.$$

Then,

$$\begin{aligned}
&\sum_{n=1}^{N(t)}\widehat{W}(t+M+1,t,t-n)\\
=\ & \sum_{n=1}^{N(t)}\widehat{W}\left(t+M+1,t+1,t\right)\widetilde{W}(t,t-n)+\sum_{n=1}^{N(t)}\widehat{W}(t+M+1,t+1,t-n)\\
\leq\ & \widehat{W}\left(t'+M,t',t\right)+\sum_{n=1}^{N(t)}\widehat{W}(t'+M,t',t-n)\\
=\ & \sum_{n=1}^{N(t)}\widehat{W}(t'+M,t',t'-n)\leq\beta
\end{aligned}$$

where $t' = t + 1$. Hence, induction holds and for any $\underline{N} > 0$,

$$\sum_{n=1}^{N(t)} \widehat{W}(t + \underline{N}, t, t - n) \leq \beta$$

Therefore, from induction, Lemma 5 holds.

Now, for any $m = 1, ...\widetilde{N}(l)$, if we substitute $t(l) - m$ for $t + \underline{N}$, $t(l-1)$ for t, then equation $A15$ becomes

$$\begin{aligned}
& \left| V\left(s, \epsilon^{(t(l)-m)}, \theta^{(t(l)-m)}\right) - V^{(t(l)-m)}\left(s, \epsilon^{(t(l)-m)}, \theta^{(t(l)-m)}\right) \right| \\
\leq\ & \sup_{s' \in S} \left| A^{(t(l)-m)}\left(\theta^{(t(l)-m)}\right) \right| \\
& + \sum_{i=0}^{\widetilde{N}(l)-m-1} \widehat{W}\left(t(l) - m, t(l) - m - i, t(l) - m - i - 1\right) \sup_{s' \in S} \left| A^{(t(l)-m-i-1)} \right| \\
& + \sum_{n=1}^{\widetilde{N}(l-1)} \sup_{\widehat{s} \in S} \left| V(\widehat{s}, \epsilon^{(t(l-1)-n)}, \theta^{(t(l-1)-n)}) - V^{(t(l-1)-n)}(\widehat{s}, \epsilon^{(t(l-1)-n)}, \theta^{(t(l-1)-n)}) \right| \\
& \widehat{W}(t(l) - m, t(l-1), t(l-1) - n)
\end{aligned}$$

Now, we take weighted sum of $\left| V\left(s, \epsilon, \theta^{(t(l)-m)}\right) - V^{(t(l)-m)}\left(s, \epsilon, \theta^{(t(l)-m)}\right) \right|$, $m = 1, ...\widetilde{N}(l)$, where the weights are defined to be $W^{\#}\left(t(l), t(l) - m\right)$. These weights satisfy $W^{\#}\left(t(l), t_l\right) > 0$ for t_l such that $t(l-1) \leq t_l < t(l)$ and

$$\sum_{t(l-1) \leq t_l < t(l)} W^{\#}\left(t(l), t_l\right) = 1 \tag{A20}$$

Then,

$$
\begin{aligned}
& \sum_{m=1}^{\widetilde{N}(l)} \left| V\left(s, \epsilon^{(t(l)-m)}, \theta^{(t(l)-m)}\right) - V^{(t(l)-m)}\left(s, \epsilon^{(t(l)-m)}, \theta^{(t(l)-m)}\right) \right| W^{\#}\left(t(l), t(l)-m\right) \\
\leq & \sum_{m=1}^{\widetilde{N}(l)} \left\{ \sup_{s' \in S} \left| A^{(t(l)-m)}\left(\theta^{(t(l)-m)}\right) \right| \right. \\
& + \sum_{i=0}^{\widetilde{N}(l)-m-1} \widehat{W}\left(t(l)-m, t(l)-m-i, t(l)-m-i-1\right) \\
& \left. \sup_{s' \in S} \left| A^{(t(l)-m-i-1)} \right| \right\} W^{\#}\left(t(l), t(l)-m\right) \\
& + \sum_{m=1}^{\widetilde{N}(l)} \sum_{n=1}^{\widetilde{N}(l-1)} \sup_{\widehat{s} \in S} \left| V(\widehat{s}, \epsilon^{(t(l-1)-n)}, \theta^{(t(l-1)-n)}) - V^{(t(l-1)-n)}(\widehat{s}, \epsilon^{(t(l-1)-n)}, \theta^{(t(l-1)-n)}) \right| \\
& \widehat{W}(t(l)-m, t(l-1), t(l-1)-n) W^{\#}\left(t(l), t(l)-m\right) \qquad \text{(A21)}
\end{aligned}
$$

Now, let,

$$
B_1(l,l) = \sum_{m=1}^{\widetilde{N}(l)} \sup \left| A^{(t(l)-m)} \right| W^{\#}\left(t(l), t(l)-m\right),
$$

$$
\begin{aligned}
B_2(l,l) \equiv & \sum_{m=1}^{\widetilde{N}(l)} W^{\#}\left(t(l), t(l)-m\right) \\
& \times \sum_{j=0}^{\widetilde{N}(l)-m-1} \left\{ \widehat{W}\left(t(l)-m, t(l)-m-j, t(l)-m-j-1\right) \sup \left| A^{(t(l)-m-j-1)} \right| \right\}
\end{aligned}
$$

and,

$$
A(l,l) \equiv B_1(l,l) + B_2(l,l).
$$

Lemma 6

$$
A(l,l) \xrightarrow{P} 0 \quad \text{as} \quad l \to \infty.
$$

Proof: We first show that $B_1(l,l) \xrightarrow{P} 0$. Recall that

$$
A^{(t)}(\theta) = \beta \left[\int V(s', \epsilon', \theta) dF_{\epsilon'}(\epsilon', \theta) - \sum_{n=1}^{N(t)} V(s', \epsilon^{(t-n)}, \theta) W_{N(t),h}(\theta, \theta^{(t-n)}) \right]
$$

$$+\beta\left[\sum_{n=1}^{N(t)}\left[V(s',\epsilon^{(t-n)},\theta)-V(s',\epsilon^{(t-n)},\theta^{(t-n)})\right]W_{N(t),h}(\theta,\theta^{(t-n)})\right]$$

Because $\int V(s',\epsilon',\theta)dF_{\epsilon'}(\epsilon',\theta)$, and $V(s',\epsilon^{(t-n)},\theta)$ are uniformly bounded and the parameter space is compact, $A^{(t)}$ is uniformly bounded. Hence, there exists $\overline{A}>0$ such that $A^{(t)}\leq\overline{A}$ for any t. Because $A^{(t)}\xrightarrow{P}0$ uniformly, for any $\eta_1>0$, $\eta_2>0$, there exists T such that for any $t>T$,

$$\Pr\left[\sup_{s'\in S,\theta\in\Theta}\left|A^{(t)}(\theta)\right|<\eta_1\right]>1-\eta_2$$

Therefore,

$$\begin{aligned}E\left[\sup_{s'\in S,\theta\in\Theta}\left|A^{(t)}(\theta)\right|\right] &\leq \eta_1\Pr\left[\sup_{s'\in S,\theta\in\Theta}\left|A^{(t)}(\theta)\right|<\eta_1\right]+\overline{A}\Pr\left[\sup_{s'\in S,\theta\in\Theta}\left|A^{(t)}(\theta)\right|\geq\eta_1\right]\\ &\leq \eta_1(1-\eta_2)+\overline{A}\eta_2 \qquad \text{(A22)}\end{aligned}$$

Hence,

$$\begin{aligned}E\left[B_1(l,l)\right]=E&\left[\sum_{m=1}^{\widetilde{N}(l)}\sup\left|A^{(t(l)-m)}\right|W^{\#}\left(t(l),t(l)-m\right)\right]\\ \leq&\sum_{m=1}^{\widetilde{N}(l)}W^{\#}\left(t(l),t(l)-m\right)\left[\eta_1(1-\eta_2)+\overline{A}\eta_2\right]\\ =&\left[\eta_1(1-\eta_2)+\overline{A}\eta_2\right]\end{aligned}$$

Now, from Chebychev's Inequality,

$$\begin{aligned}&\Pr\left[\frac{1}{\widetilde{N}(l)}\sum_{m=1}^{\widetilde{N}(l)}W^{\#}(t(l),t(l)-m)\sup_{s',\theta^{(t(l)-m)}\in\Theta}\left|A^{(t(l)-m)}\right|>\delta\right]\\ \leq&\frac{\left[\eta_1(1-\eta_2)+\eta_2\overline{A}\right]}{\delta} \qquad \text{(A23)}\end{aligned}$$

For any given δ, the RHS can be made arbitrarily small by choosing $\eta_1$and η_2. Thus, $B_1(l,l)\xrightarrow{P}0$ as $t\rightarrow\infty$.

We now show that $B_2(l,l)\xrightarrow{P}0$ as $t\rightarrow\infty$. Recall that $B_2(l,l)=\sum_{m=1}^{\widetilde{N}(l)}W^{\#}\left(t(l),t(l)-m\right)$

$\times \sum_{j=0}^{\widetilde{N}(l)-m-1} \left\{ \widehat{W}\left(t(l)-m, t(l)-m-j, t(l)-m-j-1\right) \sup \left| A^{(t(l)-m-j-1)} \right| \right\}$

For any $t' > t > 0$, let,

$$\widetilde{K}\left(t', t\right) \equiv K_h\left(\theta^{(t')} - \theta^{(t)}\right)$$

For $t_1 > t_2 > t$, define $W^*\left(t_1, t_2, t, j\right)$ recursively to be as follows.

$$\begin{aligned}
W^*\left(t_1, t_2, t, 1\right) &\equiv \widetilde{W}(t_1, t) \\
W^*\left(t_1, t_2, t, 2\right) &\equiv \sum_{j=1}^{t_1-t_2} \widetilde{W}(t_1, t_1-j) W^*\left(t_1-j, t_2, t, 1\right) \\
&\vdots \\
W^*\left(t_1, t_2, t, k\right) &\equiv \sum_{j=1}^{t_1-t_2-(k-2)} \widetilde{W}(t_1, t_1-j) W^*\left(t_1-j, t_2, t, k-1\right)
\end{aligned}$$

Similarly,

$$\begin{aligned}
K^*\left(t_1, t_2, t, 1\right) &\equiv \frac{1}{N(t_1)} \widetilde{K}(t_1, t) \\
K^*\left(t_1, t_2, t, 2\right) &\equiv \sum_{j=1}^{t_1-t_2} \frac{1}{N(t_1)} \widetilde{K}(t_1, t_1-j) K^*\left(t_1-j, t_2, t, 1\right) \\
&\vdots \\
K^*\left(t_1, t_2, t, k\right) &\equiv \sum_{j=1}^{t_1-t_2-(k-2)} \frac{1}{N(t_1)} \widetilde{K}(t_1, t_1-j) K^*\left(t_1-j, t_2, t, k-1\right)
\end{aligned}$$

Then,

$$\widehat{W}(t(l), t(l-1), \tau) \equiv \sum_{m=1}^{\widetilde{N}(l)+1} \left\{ \sum_{\Psi_m(t(l), t(l-1), \tau)} \prod_{k=1}^{m} \widetilde{W}\left(t_k, t_{k-1}\right) \right\}$$

$$= \sum_{k=1}^{\widetilde{N}(l)+1} W^*(t(l), t(l-1), \tau, k) \tag{A24}$$

Hence,

$$
\begin{aligned}
& \sum_{i=0}^{\widetilde{N}(l)-m-1} \left\{ \widehat{W}\left(t(l), t(l)-m-i, t(l)-m-i-1\right) \sup_{s' \in S} \left| A^{(t(l)-m-i-1)} \right| \right\} \\
= & \sum_{i=0}^{\widetilde{N}(l)-m-1} \left\{ \sum_{k=1}^{m+i+1} W^{*}(t(l), t(l)-m-i, t(l)-m-i-1, k) \sup_{s' \in S} \left| A^{(t(l)-m-i-1)} \right| \right\} \\
= & \sum_{k=1}^{\widetilde{N}(l)} \left\{ \sum_{i=\max\{0,k-m-1\}}^{\widetilde{N}(l)-m-1} W^{*}(t(l), t(l)-m-i, t(l)-m-i-1, k) \sup_{s' \in S} \left| A^{(t(l)-m-i-1)} \right| \right\}
\end{aligned}
$$

Also, notice that,

$$
\begin{aligned}
& W^{*}(t(l), t(l)-i, t(l)-i-1, k) \\
= & \sum_{\Psi_k(t(l), t(l)-i, t(l)-i-1)} \prod_{j=1}^{k} \widetilde{W}\left(t_j, t_j-1\right) \\
= & \sum_{\Psi_k(t(l), t(l)-i, t(l)-i-1)} \prod_{j=1}^{k} \beta \frac{\widetilde{K}\left(t_j, t_j-1\right)}{\sum_{i=1}^{N(t_j)} \widetilde{K}\left(t_j, t_j-i\right)} \\
\leq & \beta^k \left[\inf_{t(l-1) \leq t \leq t(l)} \sum_{i=1}^{N(t)} \widetilde{K}\left(t, t-i\right) \right]^{-k} \sum_{\Psi_k(t(l), t(l)-i, t(l)-i-1)} \prod_{j=1}^{k} \widetilde{K}\left(t_j, t_{j-1}\right) \\
= & \beta^k \left[\frac{1}{\widetilde{N}(l)} \inf_{t(l-1) \leq t \leq t(l)} \sum_{i=1}^{N(t)} \widetilde{K}\left(t, t-i\right) \right]^{-k} \sum_{\Psi_k(t(l), t(l)-i, t(l)-i-1)} \prod_{j=1}^{k} \frac{\widetilde{K}\left(t_j, t_{j-1}\right)}{\widetilde{N}(l)} \\
= & \beta^k \left[\frac{1}{\widetilde{N}(l)} \inf_{t(l-1) \leq t \leq t(l)} \sum_{i=1}^{N(t)} \widetilde{K}\left(t, t-i\right) \right]^{-k} K^{*}(t(l), t(l)-i, t(l)-i-1, k)
\end{aligned}
$$

Hence, we get

$$\Pr\left[\sum_{k=1}^{\widetilde{N}(l)}\sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l),t(l)-m) \sum_{i=\max\{0,k-m-1\}}^{\widetilde{N}(l)-m-1} W^{*}(t(l),t(l)-m-i,t(l)-m-i-1,k)\right.$$

$$\left.\sup_{s'\in S}\left|A^{(t(l)-m-i-1)}\right| \geq \frac{\delta-\delta^{\widetilde{N}(l)+2}}{1-\delta}\right]$$

$$\leq \Pr\left[\bigcup_{k=1}^{\widetilde{N}(l)}\left\{\sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l),t(l)-m) \sum_{i=\max\{0.k-m-1\}}^{\widetilde{N}(l)-m-1} W^{*}(t(l),t(l)-m-i,t(l)-m-i-1,k)\right.\right.$$

$$\left.\left.\sup_{s'\in S}\left|A^{(t(l)-m-i-1)}\right| \geq \delta^{k}\right\}\right]$$

$$\leq \sum_{k=1}^{\widetilde{N}(l)}\Pr\left[\sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l),t(l)-m) \sum_{i=\max\{0,k-m-1\}}^{\widetilde{N}(l)-m-1} W^{*}(t(l),t(l)-m-i,t(l)-m-i-1,k)\right.$$

$$\left.\sup_{s'\in S}\left|A^{(t(l)-m-i-1)}\right| \geq \delta^{k}\right]$$

$$\leq \sum_{k=1}^{\widetilde{N}(l)}\Pr\left\{\left[\sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l),t(l)-m) \sum_{i=\max\{0,k-m-1\}}^{\widetilde{N}(l)-m-1} K^{*}(t(l),t(l)-m-i,t(l)-m-i-1,k)\right.\right.$$

$$\left.\sup_{s'\in S}\left|A^{(t(l)-m-i-1)}\right| \geq \left[\frac{\delta}{4A\beta}\varepsilon_0 \inf_{\theta} g(\theta)\right]^{k}\right]$$

$$\left.\bigcup\left[\inf_{t(l-1)\leq t\leq t(l)}\left[\frac{1}{\widetilde{N}(l)}\sum_{i=1}^{N(t)}\widetilde{K}(t,t-i)\right] < \frac{1}{4A}\varepsilon_0 \inf_{\theta} g(\theta)\right]\right\}$$

$$\leq \sum_{k=1}^{\widetilde{N}(l)}\Pr\left[\sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l),t(l)-m) \sum_{i=\max\{0,k-m-1\}}^{\widetilde{N}(l)-m-1} K^{*}(t(l),t(l)-m-i,t(l)-m-i-1,k)\right.$$

$$\left.\sup_{s'\in S}\left|A^{(t(l)-m-i-1)}\right| \geq \left[\frac{\delta}{4A\beta}\varepsilon_0 \inf_{\theta} g(\theta)\right]^{k}\right]$$

$$+\Pr\left[\inf_{t(l-1)\leq t\leq t(l)}\left[\frac{1}{\widetilde{N}(l)}\sum_{i=1}^{N(t)}\widetilde{K}(t,t-i)\right] < \frac{1}{4A}\varepsilon_0 \inf_{\theta} g(\theta)\right]$$

Claim 1: The following equation holds.

$$E\left\{\sum_{m=1}^{\widetilde{N}(l)} K^*(t(l), t(l)-m-i, t(l)-m-i-1, k)\right\}$$

$$\leq \epsilon_1^{k+1}\left\{\sup_{\theta'\in\Theta} E\left[K_h\left(\theta' - \theta(\widetilde{g})\right)\right]\right\}^k \frac{1}{(k-1)!} \tag{A25}$$

Proof: First, by definition of K^*, note that,

$$E\left\{\sum_{m=1}^{\widetilde{N}(l)} K^*(t(l), t(l)-m-i, t(l)-m-i-1, k)\right\}$$

$$= \frac{1}{\widetilde{N}(l)^k}\sum_{m=1}^{\widetilde{N}(l)}\sum_{j_1,\ldots,j_{k-1}} I(t(l)-m-i \leq j_1 < j_2 < \ldots < j_{k-1} < j_k = t(l))$$

$$E\left[\prod_{i=0}^{k-1}\left[K_h\left(\theta^{(j_{i+1})}(f) - \theta^{j_i}(f)\right)\right]\right] \tag{A26}$$

Because $\theta'(\widetilde{g})$ and $\theta(\widetilde{g})$ are assumed to be independent,

$$\begin{aligned} E_{\theta',\theta}\left[K_h\left(\theta'(\widetilde{g}) - \theta(\widetilde{g})\right)\right] &= E_{\theta'}\left[E_\theta\left\{K_h\left(\theta'(\widetilde{g}) - \theta(\widetilde{g})\right)\right\}\right] \\ &\leq E_{\theta'}\left[\sup_{\widetilde{\theta}\in\Theta} E_\theta\left\{K_h\left(\widetilde{\theta} - \theta(\widetilde{g})\right)\right\}\right] \\ &= \sup_{\widetilde{\theta}\in\Theta} E_\theta\left[K_h\left(\widetilde{\theta} - \theta(\widetilde{g})\right)\right] \end{aligned} \tag{A27}$$

Now, for $k \geq 1$, let $(j_0, j_1, \ldots, j_k)$ satisfy $t(l)-m-i-1 = j_0 < j_1 < j_2 < \ldots < j_{k-1} < j_k = t(l)$. Notice that from Lemma 2, for any l,

$$f^{(t(l))}\left(\theta^{(t(l))}, \theta^{(t(l)-1)}\right) f^{(t(l)-1)}\left(\theta^{(t(l)-1)}, \theta^{(t(l)-2)}\right) \ldots f^{(2)}\left(\theta^{(2)}, \theta^{(1)}\right)$$

$$= \prod_{l=k}^{1}\left\{f^{(j_l)}(\theta^{(j_l)}, \theta^{(j_l-1)})\left[\prod_{j_l-1<t<tj_l} f^{(t)}\left(\theta^{(t)}, \theta^{(t-1)}\right)\right]\right\}$$

$$f^{(j_0)}(\theta^{(j_0)}, \theta^{(j_0-1)}) f^{(j_0-1)}(\theta^{(j_0-1)}, \theta^{(j_0-2)}) \ldots f^{(2)}\left(\theta^{(2)}, \theta^{(1)}\right)$$

$$\leq \prod_{l=k}^{1}\left\{\varepsilon_1\widetilde{g}(\theta^{(j_l)})\left[\prod_{j_l-1<t<tj_l} f^{(t)}\left(\theta^{(t)}, \theta^{(t-1)}\right)\right]\right\}\varepsilon_1\widetilde{g}(\theta^{(j_0)}) f^{(j_0-1)}(\theta^{(j_0-1)}, \theta^{(j_0-2)}) \ldots f^{(2)}\left(\theta^{(2)}, \theta^{(1)}\right)$$

Because $K_h() \geq 0$,

$$E\left[K_h(\theta^{(t(l))} - \theta^{(t(l)-m)})\right] = E\left[K_h\left(\theta^{(t(l))}(f^{t(l)}) - \theta^{(t(l)-m)}(f^{(t(l)-m)})\right)\right]$$
$$\leq \varepsilon_1^2 E\left[K_h\left(\theta^{(t(l))}(\widetilde{g}) - \theta^{(t(l)-m)}(\widetilde{g})\right)\right].$$

By $A27$,

$$E\left[\prod_{i=0}^{k-1}\left[K_h\left(\theta^{(j_{i+1})}(f) - \theta^{j_i}(f)\right)\right]\right] \leq \varepsilon_1^{k+1} E\left[\prod_{i=0}^{k-1}\left[K_h\left(\theta^{(j_{i+1})}(\widetilde{g}) - \theta^{j_i}(\widetilde{g})\right)\right]\right]$$
$$\leq \varepsilon_1^{k+1} E\left[\prod_{i=0}^{k-1}\sup_{\theta'\in\Theta}\left[K_h\left(\theta' - \theta^{j_i}(\widetilde{g})\right)\right]\right] = \varepsilon_1^{k+1}\left\{\sup_{\theta'\in\Theta} E\left[K_h\left(\theta' - \theta(\widetilde{g})\right)\right]\right\}^k \tag{A28}$$

Furthermore, for any i, m such that $0 < m+i \leq \widetilde{N}(l)$ and for any $k > 1$ such that $k \leq m+i$,

$$\frac{1}{\widetilde{N}(l)^{k-1}}\sum_{j_1,\ldots,j_{k-1}} I(t(l)-m-i \leq j_1 < \ldots < j_{k-1} < t(l))$$
$$= \frac{1}{\widetilde{N}(l)^{k-1}}\left(\frac{[m+i]!}{(k-1)!(m+i-(k-1))!}\right)$$
$$\leq \frac{\left[[m+i]/\widetilde{N}(l)\right]^{k-1}}{(k-1)!} \leq \frac{1}{(k-1)!} \tag{A29}$$

Substituting $A28$ and $A29$ into $A26$, $A25$ follows and hence Claim 1 is proved.

Now, by $\sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l), t(l)-m) = 1$, the law of iterated expectations and the results obtained in $A22$ and $A25$,

$$
\begin{aligned}
& E\left[\sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l),t(l)-m) \sum_{i=\max\{0,k-m-1\}}^{\widetilde{N}(l)-m-1} K^{*}(t(l),t(l)-m-i,t(l)-m-i-1,k)\right. \\
& \left.\sup_{s'\in S}\left|A^{(t(l)-m-i-1)}\right|\right] \\
= \; & E\left\{\sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l),t(l)-m)\right. \\
& \left. E\left[\sum_{i=\max\{0,k-m-1\}}^{\widetilde{N}(l)-m-1} K^{*}(t(l),t(l)-m-i,t(l)-m-i-1,k)|\Omega^{(t(l)-m-i-1)}\right]\sup_{s'\in S}\left|A^{(t(l)-m-i-1)}\right|\right\} \\
\leq \; & \sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l),t(l)-m)\left[\varepsilon_1^{k+1}\sup_{\theta'\in\Theta} E\left[K_h\left(\theta'-\theta(\widetilde{g})\right)\right]^k \frac{1}{(k-1)!}\right]\left[\eta_1\left(1-\eta_2\right)+\eta_2\overline{A}\right] \\
= \; & \left[\varepsilon_1^{k+1}\sup_{\theta'\in\Theta} E\left[K_h\left(\theta'-\theta(\widetilde{g})\right)\right]^k \frac{1}{(k-1)!}\right]\left[\eta_1\left(1-\eta_2\right)+\eta_2\overline{A}\right]
\end{aligned}
$$

Chebychev Inequality implies,

$$
\begin{aligned}
& \Pr\left[\sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l),t(l)-m) \sum_{i=\max\{0,k-m-1\}}^{\widetilde{N}(l)-m-1} K^{*}(t(l),t(l)-m-i,t(l)-m-i-1,k)\right. \\
& \left.\sup_{s'\in S}\left|A^{(t(l)-m-i-1)}\right| > \left[\frac{\delta}{4A\beta}\varepsilon_0 \inf_{\theta} g\left(\theta\right)\right]^k\right] \\
\leq \; & \frac{\left[\eta_1\left(1-\eta_2\right)+\eta_2\overline{A}\right]\varepsilon_1^{k+1}\sup_{\theta'\in\Theta} E\left[K_h\left(\theta'-\theta(\widetilde{g})\right)\right]^k \frac{1}{(k-1)!}}{\left[\frac{\delta}{4A\beta}\varepsilon_0 \inf_{\theta} g\left(\theta\right)\right]^k} \qquad \text{(A30)}
\end{aligned}
$$

Claim 2: For any $t(l-1)\leq t\leq t(l)$, either $\left[t(l-1)-\widetilde{N}(l-1)/2, t(l-1)\right]\subseteq[t-N(t),t]$ or
$\left[t(l-1), t(l-1)+\widetilde{N}(l-1)/2\right]\subseteq[t-N(t),t]$ or both.

Proof: First, we show that for t satistying $t(l-1)\leq t\leq t(l-1)+\widetilde{N}(l-1)/2$,

$$
\left[t(l-1)-\widetilde{N}(l-1)/2, t(l-1)\right]\subseteq[t-N(t),t] \qquad \text{(A31)}
$$

Because $N()$ is a nondecreasing function, $N(t) \geq \widetilde{N}(l-1)$. Hence,

$$t - t(l-1) \leq \widetilde{N}(l-1)/2 = \widetilde{N}(l-1) - \widetilde{N}(l-1)/2 \leq N(t) - \widetilde{N}(l-1)/2$$

Thus,

$$t - N(t) \leq t(l-1) - \widetilde{N}(l-1)/2$$

Since $t(l-1) \leq t$, $A31$ holds.

Next, we show that for t satisfying $t(l-1) + \widetilde{N}(l-1)/2 \leq t \leq t(l)$,

$$\left[t(l-1), t(l-1) + \widetilde{N}(l-1)/2\right] \subseteq [t - N(t), t]\,. \tag{A32}$$

From the definition of $\widetilde{N}()$,

$$t(l) - \widetilde{N}(l) = t(l-1)$$

Furthermore, because $N(s)$ is increasing at most by one with unit increase in s, $s - N(s)$ is nondecreasing in s. Hence,

$$t - N(t) \leq t(l) - \widetilde{N}(l) = t(l-1).$$

Furthermore, $t \geq t(l-1) + \widetilde{N}(l-1)/2$. Therefore, $A32$ holds. Hence, Claim 2 is proved.

Now, from $A6$, we know that for any $\eta_3 > 0$, there exists L such that for any $l > L$, $t_1 = t(l-1)$ and for $t_2 = t(l-1) + \widetilde{N}(l-1)/2$,

$$\Pr\left[\frac{1}{\widetilde{N}(l)/2}\sum_{k=1}^{\widetilde{N}(l)/2} K_h(\theta - \theta^{(t_i-k)}(f^{(t_i-k)})) \leq \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)\right] \leq \eta_3,\ \ i = 1, 2$$

Therefore,

$$\begin{aligned}
& \Pr\left[\left\{\frac{1}{\widetilde{N}(l)/2}\sum_{k=1}^{\widetilde{N}(l)/2} K_h(\theta - \theta^{(t_1-k)}(f^{(t_1-k)})) \leq \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)\right\} \cup \right. \\
& \left. \left\{\frac{1}{\widetilde{N}(l)/2}\sum_{k=1}^{\widetilde{N}(l)/2} K_h(\theta - \theta^{(t_2-k)}(f^{(t_2-k)})) \leq \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)\right\}\right] \\
\leq\ & \Pr\left[\frac{1}{\widetilde{N}(l)/2}\sum_{k=1}^{\widetilde{N}(l)/2} K_h(\theta - \theta^{(t_1-k)}(f^{(t_1-k)})) \leq \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)\right] \\
& + \Pr\left[\frac{1}{\widetilde{N}(l)/2}\sum_{k=1}^{\widetilde{N}(l)/2} K_h(\theta - \theta^{(t_2-k)}(f^{(t_2-k)})) \leq \frac{1}{2}\varepsilon_0 \inf_{\theta\in\Theta} g(\theta)\right] \\
\leq\ & 2\eta_3
\end{aligned}$$

Therefore,

$$\Pr\left[\left\{\frac{1}{\widetilde{N}(l)/2}\sum_{k=1}^{\widetilde{N}(l)/2}K_h(\theta-\theta^{(t_1-k)}(f^{(t_1-k)}))>\frac{1}{2}\varepsilon_0\inf_{\theta\in\Theta}g\left(\theta\right)\right\}\bigcap\right.$$
$$\left.\left\{\frac{1}{\widetilde{N}(l)/2}\sum_{k=1}^{\widetilde{N}(l)/2}K_h(\theta-\theta^{(t_2-k)}(f^{(t_2-k)}))>\frac{1}{2}\varepsilon_0\inf_{\theta\in\Theta}g\left(\theta\right)\right\}\right]>1-2\eta_3$$

Hence, for any t such that $t(l-1)\leq t\leq t(l)$,

$$\frac{1}{\widetilde{N}(l)}\sum_{n=1}^{N(t)}K_h(\theta-\theta^{(t-k)}(f^{(t-k)}))\geq\frac{\widetilde{N}(l-1)/2}{\widetilde{N}(l)}\frac{1}{\widetilde{N}(l-1)/2}\sum_{k=1}^{\widetilde{N}(l-1)/2}K_h(\theta-\theta^{(s-k)}(f^{(s-k)}))\tag{A33}$$

where either $s=t_1=t(l-1)$ or $s=t_2=t(l-1)+\widetilde{N}(l-1)/2$ or both. Furthermore, notice that $\frac{\widetilde{N}(l-1)/2}{\widetilde{N}(l)}\geq\frac{1}{2A}$. Therefore,

$$\begin{aligned}&\Pr\left[\inf_{t(l-1)\leq t\leq t(l)}\frac{1}{\widetilde{N}(l)}\sum_{n=1}^{N(t)}K_h(\theta-\theta^{(t-n)}(f^{(t-n)}))\geq\frac{1}{4A}\varepsilon_0\inf_{\theta\in\Theta}g\left(\theta\right)\right]\\&>\ 1-2\eta_3\end{aligned}$$

Thus,

$$\Pr\left[\inf_{t(l-1)\leq t\leq t(l)}\left[\frac{1}{\widetilde{N}(l)}\sum_{n=1}^{N(t)}\widetilde{K}(t,t-n)\right]<\frac{1}{4A}\varepsilon_0\inf_{\theta}g\left(\theta\right)\right]\leq 2\eta_3\tag{A34}$$

By $A30$ and $A34$,

$$\begin{aligned}&RHS\text{ of }A24\\\leq\ &\sum_{k=1}^{\widetilde{N}(l)+1}\frac{\left[\eta_1\left(1-\eta_2\right)+\eta_2\overline{A}\right]\varepsilon_1^{k+1}\sup_{\theta'\in\Theta}E_\theta\left[K_h\left(\theta'-\theta(g)\right)\right]^k\frac{1}{(k-1)!}}{\left[\frac{\delta}{4A\beta}\varepsilon_0\inf_\theta g\left(\theta\right)\right]^k}+2\eta_3\\=\ &\varepsilon_1\left[\eta_1\left(1-\eta_2\right)+\eta_2\overline{A}\right]e^\lambda\lambda\sum_{k=1}^{\widetilde{N}(l)+1}\left[e^{-\lambda}\frac{\lambda^{(k-1)}}{(k-1)!}\right]+2\eta_3\end{aligned}$$

where,

$$\lambda=\frac{4A\beta\varepsilon_1\sup_{\theta'\in\Theta}E_\theta\left[K_h\left(\theta'-\theta(g)\right)\right]}{\delta\varepsilon_0\inf_\theta g\left(\theta\right)}>0$$

Notice that $e^{-\lambda}\frac{\lambda^k}{k!}$ is the formula for the distribution function of the Poisson distribution. Hence,

$$\sum_{k=1}^{\widetilde{N}(l)+1} e^{-\lambda}\frac{\lambda^{(k-1)}}{(k-1)!} \leq \sum_{k=1}^{\infty} e^{-\lambda}\frac{\lambda^{(k-1)}}{(k-1)!} = 1$$

Together, we have shown that,

$$\begin{aligned} & LHS \text{ of } A24 \\ \leq \; & \varepsilon_1 \left[\eta_1 (1-\eta_2) + \eta_2 \overline{A}\right] \lambda \exp(\lambda) + 2\eta_3 \end{aligned} \tag{A35}$$

Now,

$$E_\theta \left\{K_h(\theta', \theta(g))\right\} \to g(\theta') \text{ as } h \to 0.$$

Hence, for any $B > \sup_{\theta\in\Theta} g(\theta)$, there exists $H > 0$ such that for any positive $h < H$,

$$E_\theta \left\{K_h(\theta', \theta(g))\right\} < B$$

Furthermore, for h satisfying $H \leq h \leq h(\widetilde{N}(1))$, $E_\theta \left\{K_h(\theta', \theta(g))\right\}$ is bounded. Therefore, supremum of this expectation over θ' is uniformly bounded. Therefore, λ also is uniformly bounded. Hence, RHS of $A35$ can be made arbitrarily small by choosing η_1, η_2 and η_3 small enough.

Thus, Lemma 6 is proved. That is, we have shown that

$$A(l,l) \to 0 \text{ as } l \to \infty$$

Let

$$\begin{aligned} & \Xi(l, l_1+1) \\ \equiv \; & \left\{(t_l, t_{l-1}, ..., t_{l_1+1}) : t(l_1) \leq t_{l_1+1} < t(l_1+1), ..., t_{l-1} \leq t(l-1) \leq t_l < t(l)\right\}. \end{aligned}$$

Now, define, $\overrightarrow{W}(t(l), t(l_1), t_{l_1})$ as follows: For $l_1 = l$,

$$\overrightarrow{W}(t(l), t(l), t_l) \equiv W^{\#}(t(l), t_l).$$

For $l_1 = l-1$,

$$\begin{aligned} & \overrightarrow{W}(t(l), t(l-1), t_{l-1}) \\ = \; & \sum_{m=1}^{\widetilde{N}(l)} W^{\#}(t(l), t(l)-m)\, \widehat{W}(t(l)-m, t(l-1), t_{l-1}). \end{aligned}$$

For $l_1 \leq l-2$,

$$\begin{aligned}
&\overrightarrow{W}(t(l), t(l_1), t_{l_1}) \\
\equiv &\sum_{\left(t_l, t_{t-1}, \ldots, t_{l_1+1}\right) \in \Xi(l, l_1+1)} W^{\#}\left(t(l), t_l\right)\left\{\prod_{j=l_1+1}^{l-1} \widehat{W}\left(t_{j+1}, t(j), t_j\right)\right\} \widehat{W}\left(t_{l_1+1}, t(l_1), t_{l_1}\right)
\end{aligned}$$

Hence, $A21$ can be written as follows.

$$\begin{aligned}
&\sum_{m=1}^{\widetilde{N}(l)}\left|V\left(s, \epsilon^{(t(l)-m)}, \theta^{(t(l)-m)}\right)-V^{(t(l)-m)}\left(s, \epsilon^{(t(l)-m)}, \theta^{(t(l)-m)}\right)\right| \overrightarrow{W}(t(l), t(l), t(l)-m) \\
\leq &\sum_{m=1}^{\widetilde{N}(l)} \overrightarrow{W}(t(l), t(l), t(l)-m) \sup_{s' \in S}\left|A^{(t(l)-m)}\left(\theta^{(t(l)-m)}\right)\right| \\
&+\sum_{m=1}^{\widetilde{N}(l)} \overrightarrow{W}(t(l), t(l), t(l)-m) \\
&\times \sum_{i=0}^{N(l)-m-1} \widehat{W}(t(l)-m, t(l)-m-i, t(l)-m-i-1) \sup_{s' \in S}\left|A^{(t(l)-m-i-1)}\right| \\
&+\sum_{m=1}^{\widetilde{N}(l-1)} \sup_{\widehat{s} \in S}\left|V(\widehat{s}, \epsilon^{(t(l-1)-m)}, \theta^{(t(l-1)-m)})-V^{(t(l-1)-m)}(\widehat{s}, \epsilon^{(t(l-1)-m)}, \theta^{(t(l-1)-m)})\right| \\
&\overrightarrow{W}(t(l), t(l-1), t(l-1)-m) \qquad \text{(A36)}
\end{aligned}$$

Now, let

$$A(l, l_1) \equiv B_1(l, l_1)+B_2(l, l_1)$$

where,

$$B_1(l, l_1) \equiv \sum_{m=1}^{\widetilde{N}(l_1)} \overrightarrow{W}(t(l), t(l_1), t(l_1)-m) \sup_{s' \in S}\left|A^{(t(l_1)-m)}\right|$$

and

$$\begin{aligned}
B_2(l, l_1) \equiv &\sum_{m=1}^{\widetilde{N}(l_1)} \overrightarrow{W}(t(l), t(l_1), t(l_1)-m) \sum_{j=0}^{N(l_1)-m-1} \\
&\left\{\widehat{W}(t(l_1)-m, t(l_1)-m-j, t(l_1)-m-j-1) \sup_{s' \in S}\left|A^{(t(l_1)-m-j-1)}\right|\right\}
\end{aligned}$$

Then,

$$
\begin{aligned}
& \sum_{m=1}^{\widetilde{N}(l_1)} \left| V\left(s, \epsilon^{(t(l_1)-m)}, \theta^{(t(l_1)-m)}\right) - V^{(t(l_1)-m)}\left(s, \epsilon^{(t(l_1)-m)}, \theta^{(t(l_1)-m)}\right) \right| \\
& \overrightarrow{W}\left(t(l), t(l_1), t(l_1)-m\right) \\
\leq \; & A(l, l_1) \\
& + \sum_{m=1}^{\widetilde{N}(l_1-1)} \left| V\left(s, \epsilon^{(t(l_1-1)-m)}, \theta^{(t(l_1-1)-m)}\right) - V^{(t(l_1-1)-m)}\left(s, \epsilon^{(t(l_1-1)-m)}, \theta^{(t(l_1-1)-m)}\right) \right| \\
& \overrightarrow{W}\left(t(l), t(l_1-1), t(l_1-1)-m\right) \qquad \text{(A37)}
\end{aligned}
$$

Lemma 7

Given $l > l_1$

$$A(l, l_1) \xrightarrow{P} 0 \;\text{ as }\; l \to \infty.$$

Proof: By definition of $\overrightarrow{W}$,

$$
\begin{aligned}
& \overrightarrow{W}(t(l), t(l_1), t(l_1)-m) \\
= \; & \left[\sum_{t(l-1) \leq t_l < t(l)} W^{\#}\left(t(l), t_l\right) \left\{ \sum_{t(l-2) \leq t_{l-1} < t(l-1)} \widehat{W}\left(t_l, t(l-1), t_{l-1}\right) \right.\right. \\
& \left.\left. \ldots \left\{ \sum_{t(l_1) \leq t_{l_1+1} < t(l_1+1)} \widehat{W}\left(t_{l_1+2}, t(l_1+1), t_{l_1+1}\right) \widehat{W}\left(t_{l_1+1}, t(l_1), t(l_1)-m\right) \right\} \right\} \right] \qquad \text{(A38)}
\end{aligned}
$$

By Lemma 5,

$$\sum_{m=1}^{\widetilde{N}(l_1)} \widehat{W}\left(t_{l_1+1}, t(l_1), t(l_1)-m\right) \leq \beta$$

and similarly, for any $k > l_1$,

$$\sum_{t(k-1) \leq t_k < t(k)} \widehat{W}\left(t_{k+1}, t(k), t_k\right) \leq \beta$$

Applying these inequalities to $A38$ yields,

$$\sum_{m=1}^{\widetilde{N}(l_1)} \overrightarrow{W}\left(t(l), t(l_1), t(l_1)-m\right) \leq \beta^{(l-l_1)} \qquad \text{(A39)}$$

By $A22$ and $A39$ (and using iterated expectations as earlier),

$$\begin{aligned} E\left[B_1(l,l_1)\right] &\equiv E\left[\sum_{m=1}^{\widetilde{N}(l_1)} \overrightarrow{W}\left(t(l),t(l_1),t(l_1)-m\right)\sup_{s'\in S}\left|A^{(t(l_1)-m)}\right|\right] \\ &\leq \beta^{l-l_1}\left[\eta_1\left(1-\eta_2\right)+\eta_2\overline{A}\right]. \end{aligned}$$

Hence, from Chebychev Inequality

$$\Pr\left[B_1(l,l_1)\geq\delta\right]\leq\frac{\beta^{l-l_1}\left[\eta_1\left(1-\eta_2\right)+\eta_2\overline{A}\right]}{\delta}$$

Since, η_1,η_2 can be made arbitrarily small by choosing l to be large enough, for any arbitarily positive δ, RHS can be made arbitarrily small by increasing l, while keeping $l-l_1$ constant,

$$B_1(l,l_1)\xrightarrow{P}0$$

as $l\to\infty$.

Next, we prove convergence of $B_2(l,l_1)$. We again use $A22$, $A25$, and $A39$ and the law of iterated expectations, to derive,

$$\begin{aligned} & E\left[\sum_{m=1}^{\widetilde{N}(l_1)} \overrightarrow{W}\left(t(l),t(l_1),t(l_1)-m\right)\right. \\ & \left.\left\{\sum_{j=\max\{0,k-m-1\}}^{\widetilde{N}(l_1)-m-1} K^*(t(l_1)-m,t(l_1)-m-j,t(l_1)-m-j-1,k)\sup_{s'\in S}\left|A^{(t(l_1)-m-j-1)}\right|\right\}\right] \\ \leq\ & \beta^{(l-l_1)}\left[\eta_1\left(1-\eta_2\right)+\eta_2\overline{A}\right]\epsilon_1^{k+1}\left\{\sup_{\theta'\in\Theta}E\left[K_h\left(\theta'-\theta(\widetilde{g})\right)\right]\right\}^k\frac{1}{(k-1)!} \end{aligned}$$

Hence, from Chebyshev's inequality, we get,

$$\begin{aligned}
& \Pr\left[\sum_{m=1}^{\widetilde{N}(l_1)} \overrightarrow{W}\left(t(l), t(l_1), t(l_1)-m\right)\right. \\
& \left\{\sum_{j=\max\{0,k-m-1\}}^{\widetilde{N}(l_1)-m-1} K^*(t(l_1)-m, t(l_1)-m-j, t(l_1)-m-j-1, k) \sup_{s' \in S}\left|A^{(t(l_1)-m-j-1)}\right|\right\} \\
& \left. > \left[\frac{\delta}{4A^{l+1-l_1}} \varepsilon_0 \inf_{\theta} g(\theta)\right]^k\right] \\
\leq & \frac{\beta^{(l-l_1)}\left[\eta_1(1-\eta_2)+\eta_2\overline{A}\right] \epsilon_1^{k+1} \sup_{\theta' \in \Theta} E\left[K_h\left(\theta' - \theta(\widetilde{g})\right)\right]^k \frac{1}{(k-1)!}}{\left[\frac{\delta}{4A^{l+1-l_1}} \varepsilon_0 \inf_\theta g(\theta)\right]^k}
\end{aligned}$$

Furthermore, let $t_1(l) \equiv t(l-1)$ and $t_2(l) = t(l-1) + \widetilde{N}(l-1)/2$. Then, arguments similar to ones used in deriving equation $A33$ can be used to derive the inequality below.

$$\begin{aligned}
& \inf_{t(l_1-1) \leq t \leq t(l)} \left[\frac{1}{\widetilde{N}(l)} \sum_{i=1}^{N(t)} \widetilde{K}(t, t-i)\right] \\
\geq & \min_{l_1-1 \leq \widetilde{l} < l} \left\{\frac{\widetilde{N}(\widetilde{l})/2}{\widetilde{N}(l)} \frac{1}{\widetilde{N}(\widetilde{l})/2} \min\left\{\sum_{k=1}^{\widetilde{N}(\widetilde{l})/2} K_h(\theta - \theta^{(t_1(\widetilde{l})-k)}(f^{(t_1(\widetilde{l})-k)})), \sum_{k=1}^{\widetilde{N}(\widetilde{l})/2} K_h(\theta - \theta^{(t_2(\widetilde{l})-k)}(f^{(t_2(\widetilde{l})-k)}))\right\}\right\} \\
\geq & \frac{1}{2A^{l+1-l_1}} \frac{1}{\widetilde{N}(l^*)/2} \min\left\{\sum_{k=1}^{\widetilde{N}(l^*)/2} K_h(\theta - \theta^{(t_1(l^*)-k)}(f^{(t_1(l^*)-k)})), \sum_{k=1}^{\widetilde{N}(l^*)/2} K_h(\theta - \theta^{(t_2(l^*)-k)}(f^{(t_2(l^*)-k)}))\right\}
\end{aligned}$$

where,

$$\begin{aligned}
l^* \equiv & \arg\min_{\widetilde{l}: l_1-1 \leq \widetilde{l} < l} \left\{\frac{1}{2A^{l+1-\widetilde{l}}} \frac{1}{\widetilde{N}(\widetilde{l})/2}\right. \\
& \left. \min\left\{\sum_{k=1}^{\widetilde{N}(\widetilde{l})/2} K_h(\theta - \theta^{(t_1(\widetilde{l})-k)}(f^{(t_1(\widetilde{l})-k)})), \sum_{k=1}^{\widetilde{N}(\widetilde{l})/2} K_h(\theta - \theta^{(t_2(\widetilde{l})-k)}(f^{(t_2(\widetilde{l})-k)}))\right\}\right\}
\end{aligned}$$

Hence, using $A24$,

$$
\begin{aligned}
& \Pr\left[\sum_{m=1}^{\widetilde{N}(l_1)} \overrightarrow{W}(t(l), t(l_1), t(l_1)-m)\right. \\
& \left.\sum_{j=\max\{0,k-m-1\}}^{\widetilde{N}(l_1)-m-1} \widehat{W}\left(t(l_1)-m, t(l_1)-m-j, t(l_1)-m-j-1\right) \sup_{s'\in S}\left|A^{(t(l_1)-m-j-1)}\right| > \frac{\delta-\delta^{\widetilde{N}(l)+1}}{1-\delta}\right] \\
\leq\ & \sum_{k=1}^{\widetilde{N}(l_1)} \Pr\left[\sum_{m=1}^{\widetilde{N}(l_1)} \overrightarrow{W}(t(l), t(l_1), t(l_1)-m)\right. \\
& \left.\sum_{j=\max\{0,k-m-1\}}^{\widetilde{N}(l_1)-m-1} W^*\left(t(l_1)-m, t(l_1)-m-j, t(l_1)-m-j-1, k\right) \sup_{s'\in S}\left|A^{(t(l_1)-m-j-1)}\right| \geq \delta^k\right] \\
\leq\ & \sum_{k=1}^{\widetilde{N}(l_1)} \Pr\left[\sum_{m=1}^{\widetilde{N}(l_1)} \overrightarrow{W}(t(l), t(l_1), t(l_1)-m)\right. \\
& \sum_{j=\max\{0,k-m-1\}}^{\widetilde{N}(l_1)-m-1} \beta^k K^*(t(l_1)-m, t(l_1)-m-j, t(l_1)-m-j-1, k) \sup_{s'\in S}\left|A^{(t(l_1)-m-j-1)}\right| \\
& \left.\geq \left[\frac{\delta}{4A^{l+1-l_1}}\varepsilon_0 \inf_\theta g(\theta)\right]^k\right] \\
& + \Pr\left[\inf_{l_1-1<\widetilde{l}\leq l} \frac{1}{\widetilde{N}(l)} \min\left\{\sum_{k=1}^{\widetilde{N}(\widetilde{l})/2} K_h(\theta-\theta^{(t_1(\widetilde{l})-k)}(f^{(t_1(\widetilde{l})-k)})), \sum_{k=1}^{\widetilde{N}(\widetilde{l})/2} K_h(\theta-\theta^{(t_2(\widetilde{l})-k)}(f^{(t_2(\widetilde{l})-k)}))\right\}\right. \\
& \left.< \frac{1}{4A^{l+1-l_1}}\varepsilon_0 \inf_\theta g(\theta)\right] \\
\leq\ & \beta^{l-l_1}\varepsilon_1\left[\eta_1(1-\eta_2)+\eta_2\overline{A}\right] e^\lambda \sum_{k=1}^{\widetilde{N}(l_1)}\left[e^{-\lambda}\frac{\lambda^k}{(k-1)!}\right] + 2(l+1-l_1)\eta_3
\end{aligned}
$$

where,

$$
\lambda = \frac{4\beta A^{l+1-l_1}\varepsilon_1 \sup_{\theta'\in\Theta} E\left[K_h\left(\theta'-\theta(g)\right)\right]}{\delta\varepsilon_0 \inf_\theta g(\theta)} > 0
$$

which can be made arbitrarily close to zero by increasing l while keeping $\Delta l \equiv l - l_1$ constant. Therefore,

$$
B_2(l, l-\Delta l) \xrightarrow{P} 0
$$

Hence, Lemma 7 holds.

Now, let,

$$\Delta V(m,n) \equiv \sup_{s\in S}\left|V(s,\epsilon^{(t(m)-n)},\theta^{(t(m)-n)}) - V^{(t(m)-n)}(s,\epsilon^{(t(m)-n)},\theta^{(t(m)-n)})\right|$$

$$\Delta V(m) \equiv \left[\Delta V(m,1),...,\Delta V(m,\widetilde{N}(m))\right]$$

$$\overline{W}(l,k) \equiv \left[\overrightarrow{W}(t(l),t(l+1-k),t(l+1-k)-m)\right]_{m=1}^{\widetilde{N}(l+1-k)}$$

Then, by $A39, \overline{W}(l,k)'\iota \leq \beta^{k-1}$ and from $A36$, we obtain the following.

$$\begin{aligned}\Delta V(l)'\overline{W}(l,1) &\leq A(l,l) + \Delta V(l-1)'\overline{W}(l,2)\\ &\leq ... \leq \sum_{i=0}^{k-1} A(l,l-i) + \Delta V(l-k)'\,\overline{W}(l,k+1)\,.\end{aligned}$$

Given k, the first term on the RHS, $\sum_{i=0}^{k-1} A(l,l-i)$ converge to 0 in probability as $l\to\infty$, by Lemma 7 and since $\Delta V(l+1-k)$ is bounded and $\overline{W}(l,k)'\iota \leq \beta^{k-1}$, the second term can be made arbitrarily small by chosing a large enough k. Therefore, $\Delta V(l)'\,W(l,1)$ converges to zero in probability as $l\to\infty$.

Lemma 8:

$$\left|V(s,\epsilon^{(t)},\theta^{(t)}) - V^{(t)}(s,\epsilon^{(t)},\theta^{(t)})\right| \xrightarrow{P} 0 \text{ as } t\to\infty$$

Suppose not. Then, there exists a positive δ, η and a sequence $\{t_k\}$ such that

$$\Pr\left(\left|V(s,\epsilon^{(t_k)},\theta^{(t_k)}) - V^{(t_k)}(s,\epsilon^{(t_k)},\theta^{(t_k)})\right| \geq \delta\right) > \eta. \tag{A40}$$

Set the weights $W^{\#}$ be as follows: If there is t_k such that $t(l-1)\leq t_k < t(l)$, then, let

$$t^*(l) = \min_{t(l-1)\leq t_k < t(l)} \{t_k\}\,.$$

Otherwise, let

$$t^*(l) = t(l-1).$$

Let

$$W^{\#}\left(t(l), t_l\right) = I(t_l = t^*(l))$$

Then, because $\Delta V\left(l\right)' \overline{W}\left(l, 1\right) \xrightarrow{P} 0$ as $l \to \infty$,

$$\left| V(s, \epsilon^{(t^*(l))}, \theta^{(t^*(l))}) - V^{(t^*(l))}(s, \epsilon^{(t^*(l))}, \theta^{(t^*(l))}) \right| \xrightarrow{P} 0 \text{ as } l \to \infty$$

which contradicts $A40$. Hence, Lemma 8 holds, and thus we have proved Theorem 1.

Proof of Theorem 2

We are given a Markov chain with transition function $f^{(t)}\left(.,.\right)$ which converges to $f\left(.,.\right)$ in probability uniformly as $t \to \infty$. As in Lemma 1, we can construct a density $g(.)$ and a constant $\varepsilon > 0$ such that for any $\theta \in \Theta$,

$$\begin{aligned} f^{(t)}(\theta, .) &\geq \varepsilon g(.) \\ f(\theta, .) &\geq \varepsilon g(.) \end{aligned}$$

Define $\nu^{(t)}$ as follows.

$$v^{(t)}(\theta) = \min \left\{ \inf_{\theta'} \left\{ \frac{f^{(t)}\left(\theta, \theta'\right)}{f\left(\theta, \theta'\right)} \right\}, 1 \right\}$$

Then,

$$\begin{aligned} f^{(t)}(\theta, .) &\geq v^{(t)} f\left(\theta, .\right) \\ f(\theta, .) &\geq v^{(t)} f\left(\theta, .\right) \end{aligned}$$

Now, construct the following coupling scheme. Let $X^{(t)}$ be a random variable that follows the transition probability $f^{(t)}(x, .)$ given $X^{(t-1)} = x$, and $Y^{(t)}$ be a Markov process that follows the transition probability $f(y, .)$, given $Y^{(t-1)} = y$. Suppose $X^{(t)} \neq Y^{(t)}$. With probability $\varepsilon > 0$, let

$$X^{(t+1)} = Y^{(t+1)} = Z^{(t+1)} \sim g(.)$$

and with probability $1 - \varepsilon$,

$$X^{(t+1)} \sim \frac{1}{1-\varepsilon} \left[f^{(t)}\left(X^{(t)}, .\right) - \varepsilon g(.) \right]$$

$$Y^{(t+1)}\sim\frac{1}{1-\varepsilon}\left[f\left(Y^{(t)},.\right)-\varepsilon g(.)\right]$$

Suppose $X^{(t)}=Y^{(t)}=Z^{(t)}$. With probability $v^{(t)}$,

$$X^{(t+1)}=Y^{(t+1)}\sim f(Z^{(t)},.)$$

and with probaiblity $1-v^{(t)}$,

$$X^{(t+1)}\sim\frac{1}{1-v^{(t)}}\left[f^{(t)}\left(X^{(t)},.\right)-v^{(t)}f(Z^{(t)},.)\right]$$

$$Y^{(t+1)}\sim\frac{1}{1-v^{(t)}}\left[f\left(Y^{(t)},.\right)-v^{(t)}f(Z^{(t)},.)\right]$$

As $f^{(t)}(x,.)\xrightarrow{P}f(x,.)$ uniformly over the compact parameter set Θ, $v^{(t)}$ converges to 1 in probability. Let $w^{(t)}=1-v^{(t)}$. Then, $w^{(t)}\xrightarrow{P}0$ as $t\to\infty$. Let $S^{(t)}\in\{1,2\}$ be the state at iteration t, where state 1 is assumed to be the state in which $X^{(t)}=Y^{(t)}$, and state 2 the state in which $X^{(t)}\neq Y^{(t)}$. Then, $S^{(t)}$ follows the Markov process with the following transition matrix.

$$P=\begin{bmatrix}1-w^{(t)} & w^{(t)}\\ \varepsilon & 1-\varepsilon\end{bmatrix}$$

Denote the unconditional probability of state 1 at time t as $\pi^{(t)}$. Then,

$$\left[\pi^{(t+1)},1-\pi^{(t+1)}\right]=\left[\pi^{(t)},1-\pi^{(t)}\right]\begin{bmatrix}1-w^{(t)} & w^{(t)}\\ \varepsilon & 1-\varepsilon\end{bmatrix}$$

Hence,

$$\begin{aligned}\pi^{(t+1)} &= \pi^{(t)}\left(1-w^{(t)}-\varepsilon\right)+\varepsilon\\ &\geq \pi^{(t)}\left(1-\varepsilon\right)+\varepsilon-w^{(t)}\\ &\geq \pi^{(t-m)}\left(1-\varepsilon\right)^{m+1}+1-\left(1-\varepsilon\right)^{m+1}-\left[w^{(t)}+\left(1-\varepsilon\right)w^{(t-1)}+...+\left(1-\varepsilon\right)^{m}w^{(t-m)}\right]\end{aligned}$$

We now prove that $\pi^{(t)}\xrightarrow{P}1$.

Define W_{tm} to be

$$W_{tm}=w^{(t)}+\left(1-\varepsilon\right)w^{(t-1)}+...+\left(1-\varepsilon\right)^{m}w^{(t-m)}$$

Because $w^{(t)} \xrightarrow{P} 0$, for any $\delta_1 > 0$, $\delta_2 > 0$, there exists $N > 0$ such that for any $t \geq N$,

$$\Pr\left[\left|w^{(t)} - 0\right| < \delta_1\right] > 1 - \delta_2$$

Now, given any $\overline{\delta}_1 > 0$, $\overline{\delta}_2 > 0$, let m be such that

$$\max\left\{(1-\varepsilon)^m, \varepsilon^{m+1}\right\} < \frac{\overline{\delta}_1}{5}$$

Also, let δ_1 satisfy $\delta_1 < \frac{\overline{\delta}_1}{5(m+1)}$, and δ_2 satisfy $\delta_2 < \frac{\overline{\delta}_2}{m+1}$. Then,

$$\begin{aligned}
\Pr\left\{|W_{tm} - 0| < \frac{\overline{\delta}_1}{5}\right\} &\geq \Pr\left\{\bigcap_{j=t-m}^{t} \left|w^{(j)} - 0\right| < \delta_1\right\} \\
&= 1 - \Pr\left\{\bigcup_{j=t-m}^{t} \left|w^{(j)} - 0\right| \geq \delta_1\right\} \\
&\geq 1 - \sum_{j=t-m}^{t} \Pr\left\{\left|w^{(j)} - 0\right| \geq \delta_1\right\} \geq 1 - \overline{\delta}_2 \qquad \text{(A47)}
\end{aligned}$$

Now, let $\overline{N}$ be defined as $\overline{N} = \max\{N, m\}$. Then, for each $k > \overline{N}$,

$$\begin{aligned}
\Pr\left[\left|\pi^{(t+1)} - 1\right| < \overline{\delta}_1\right] &\geq \Pr\left[\left|\pi^{(t-m)}(1-\varepsilon)^m - (1-\varepsilon)^{m+1} + W_{tm}\right| < \overline{\delta}_1\right] \\
&\geq \Pr\left[\left|\pi^{(t-m)}(1-\varepsilon)^m - (1-\varepsilon)^{m+1}\right| < \frac{2\overline{\delta}_1}{5}, |W_{tm}| < \frac{\overline{\delta}_1}{5}\right] \\
&= \Pr\left[|W_{tm}| < \frac{\overline{\delta}_1}{5}\right] \qquad \text{(A48)}
\end{aligned}$$

Last equality holds because $0 \leq \pi^{(t-m)} \leq 1$ and thus,

$$\left|\pi^{(t-m)}(1-\varepsilon)^m - (1-\varepsilon)^{m+1}\right| \leq \left|(1-\varepsilon)^m - (1-\varepsilon)^{m+1}\right| \leq |(1-\varepsilon)^m| < \frac{\overline{\delta}_1}{5}$$

From $(A47)$ and (A48), we conclude that

$$\Pr\left[\left|\pi^{(t+1)} - 1\right| < \overline{\delta}_1\right] \geq 1 - \overline{\delta}_2$$

Therefore, π_k converges to 1 in probability.

Therefore, for any $\delta > 0$, there exists M such that for any $t > M$,

$$\Pr\left[X^{(t)} = Y^{(t)}\right] > 1 - \delta$$

Since $Y^{(t)}$follows a stationary distribution, $X^{(t)}$ converges to a stationary process in probability.

Figure 1: Gibbs Sampler Output of Exit Value (True Value:0.4)

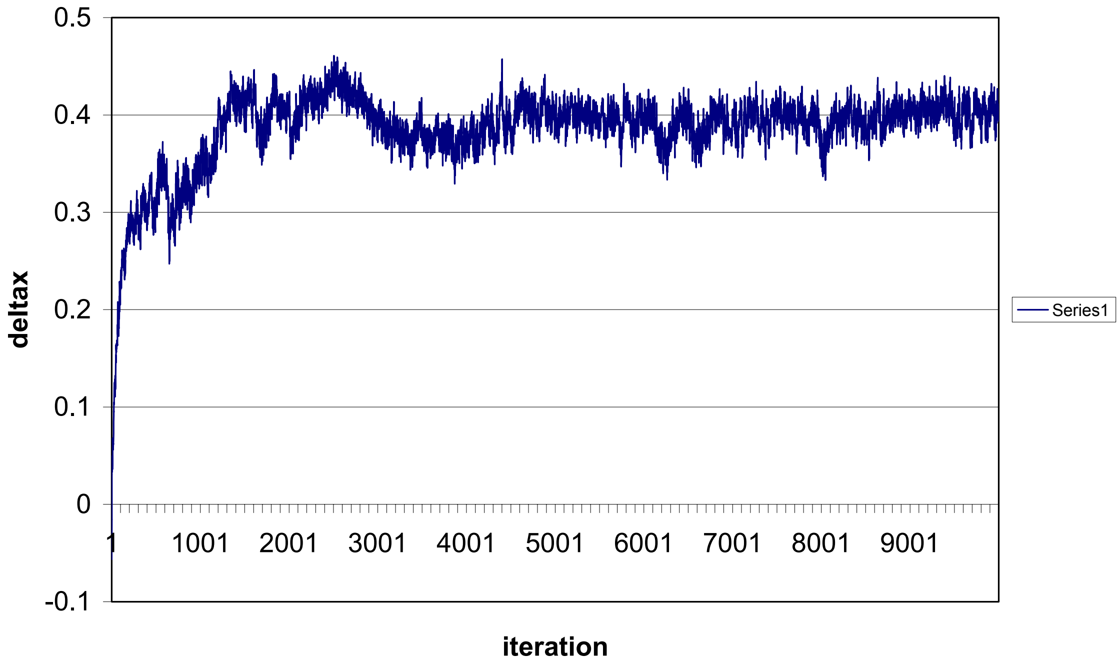

Figure 2: Gibbs Sampler Output of Entry Cost (True Value:0.4)

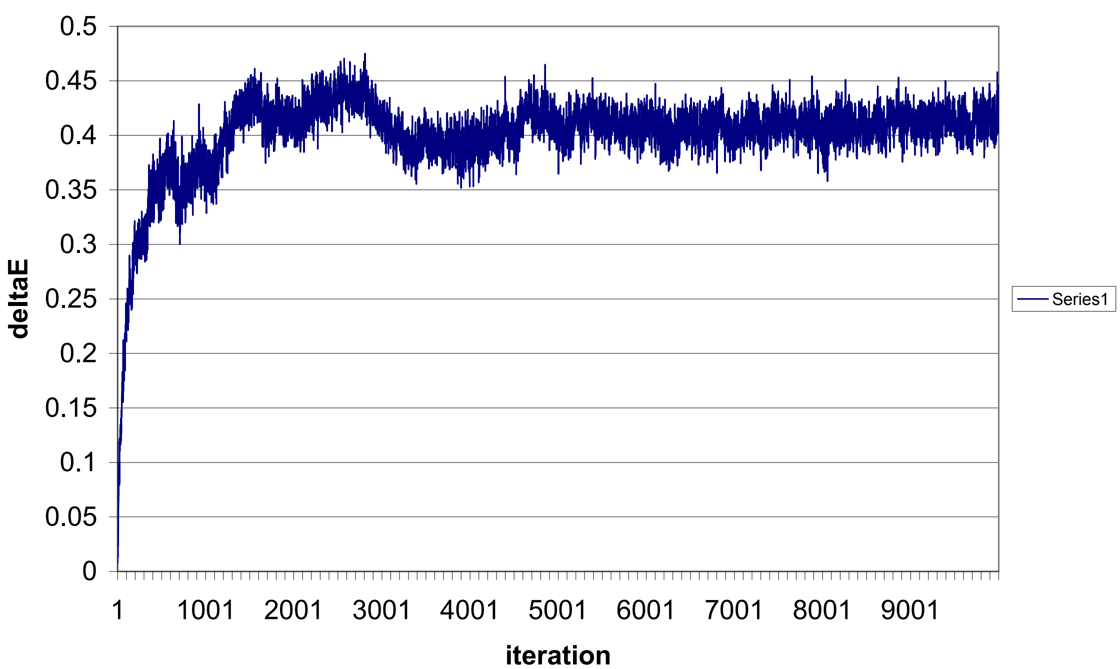

Figure 3: Gibbs Sampler Output of the Entry and Exit Shock Standard Error (True Value: 0.4)

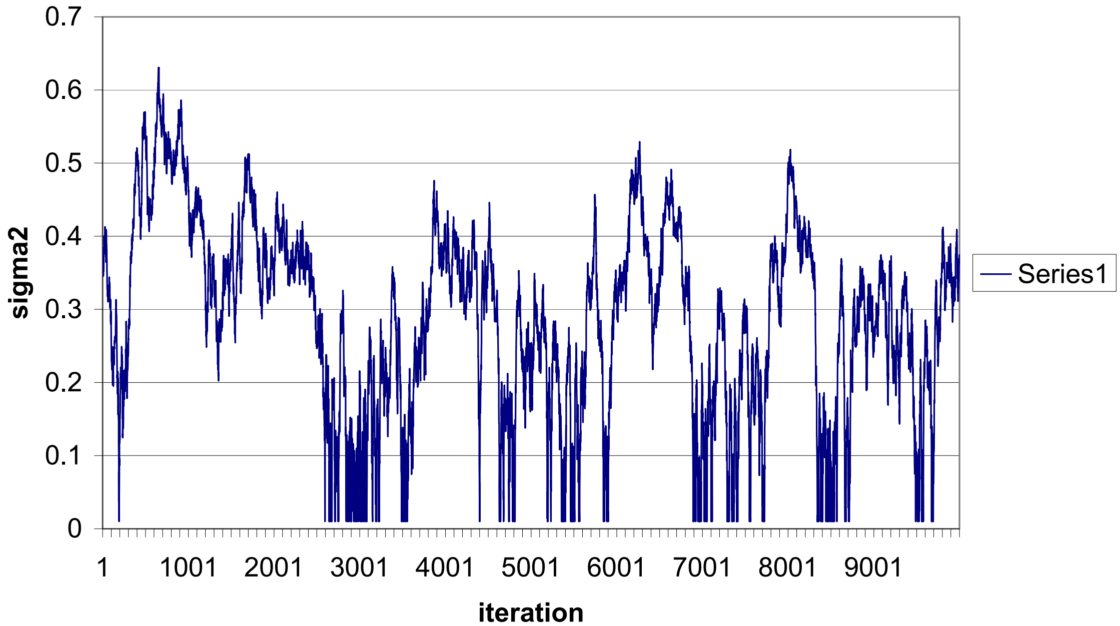

Figure 4: Gibbs Sampler Output of the Capital Stock Transition Parameter b1 (True Value: 0.1)