Bayesian Estimation of Dynamic Discrete Choice Models*

Susumu Imai
Concordia University

Neelam Jain
Northern Illinois University

and
Andrew Ching
Ohio State University

February 1, 2004

Abstract

We propose a new estimator for dynamic programming discrete choice models. Our estimation method combines the Dynamic Programming algorithm with a Bayesian Markov Chain Monte Carlo algorithm into one single Markov Chain algorithm that solves the dynamic programming problem and estimates the parameters at the same time.

Our key innovation is that during each solution-estimation iteration both the parameters and the expected value function are updated only once. This is in contrast to the conventional estimation methods where at each estimation iteration the dynamic programming problem needs to be fully solved. A single dynamic programming solution requires repeated updates of the expected value functions. As a result, in our algorithm the computational burden of estimating a dynamic model is of similar order of magnitude as that of a static model.

Another feature of our algorithm is that even though per estimation iteration, we keep the number of grid points on the state variable small, we can make the number of effective grid points as large as we want by simply increasing the number of estimation iterations. This is how our algorithm overcomes the "Curse of Dimensionality".

We prove that under mild conditions similar to those imposed in standard Bayesian literature, the parameters in our algorithm converge in probability to the true posterior distribution, regardless of the starting values. We show how our method can be applied to models with standard random effects where observed and unobserved heterogeneities are continuous. This is in contrast to most dynamic structural estimation models where only a small number of discrete types are allowed as heterogeneities.

*Very preliminary. Comments welcome. 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 both empirical economics and marketing. Examples include Keane and Wolpin (1997), Erdem and Keane (1995). Recently, it has also been applied in analyzing criminal behavior, as in Imai and Krishna (2001). 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' behaviors 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. There are mainly two reasons why this estimation method is computationally demanding. First, in structural estimation, the likelihood or the moment conditions we construct 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 (Emax function) converges. That is, given a parameter value, in order to evaluate the likelihood or the moment condition once, Bellman equation has to be computed many times until convergence. Secondly, in solving the Dynamic Programming 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 even in relatively simple setting infeasible.

In this paper, we propose an estimator that helps overcome the two computational difficulties of structural estimation. We adopt 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, we need to solve the Bellman equation only once between each estimation step. Since evaluating a single Bellman equation 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 equations only limited number of times before constructing the likelihood, or to solve the DP problems "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. In the second strategy, they still compute the solution of the DP problem, whether exact or inexact, during each estimation step. In our algorithm, we only need to solve the Bellman equation once between each estimation step.

Specifically, we start with some initial guess of the emax function. We then evaluate the Bellman equation for each state space point, or for a subset of state space grid points. That is, we solve the optimal policies and calculate the value function. 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 (in the case with finite state space points) or close to the current state variables (when 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 Dynamic Programming 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 the way how our estimation method overcomes the "Curse of Dimensionality".

The main difference between our estimation algorithm to the conventional structural estimation algorithms is in 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 remain unused, except for the likelihood and its Jacobian and Hessian in Classical case, and MCMC transition function in Bayesian case. In contrast, we extensively use the vast amounts 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 Programming model of firm entry and exit choice with observed and unobserved heterogeneity. 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 Maximum Likelihood estimation strategy, this setup imposes almost prohibitive computational burden because during each estimation step, the Dynamic Programming model has to be solved for each random effects parameter value and each value of observed firm characteristic. This is why most practitioners of structural estimation assume discrete distributions for random effects following Heckman and Singer (1984) and allow for only discrete types as observed characteristics. We show that using our algorithm, the estimation exercise becomes one that is computationally quite similar in difficulty to a Bayesian estimation of 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 10 hours, similar to the time required to estimate a reasonably complicated static random effects model. In contrast, the conventional simulated maximum likelihood routine took 6 hours and 20 minutes just for a single iteration.

In addition to the experiments that show convergence of our algorithm to the correct parameter values, we provide a formal proof 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 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. 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 the unobserved heterogeneities. As discussed earlier, we do so by simultaneously solving for the Dynamic Programming problem and iterating on the MCMC algorithm.

Our estimation method not only makes Bayesian application to DDC models computationally feasible, but even possibly superior to some conventional (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 the 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 in the following two ways. One is by inverting the numerically calculated Information Matrix, which is valid only in a large sample world. The other is 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. In Section 3, we explain our Bayesian DP algorithm and prove convergence. In Section 4, we present estimation results of several experiments applied to the model of entry and exit. Finally, in Section 5, 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.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,

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

where s' is the next period's state variable, U is the current return function. Let ϵ be a vector those a th element ϵ_a is a random shock to current returns to choice a. We assume that ϵ follows a multivariate distribution $F_\epsilon(\epsilon,\theta)$, which is independent over time. β is the discount factor. The expectation is taken with respect to the next period's shock ϵ'. We assume that the next period state variable s' is a deterministic function of current period state variable s, current period action a, and parameter θ[1]. That is,

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

The value function is defined to be as follows.

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

We assume that the dataset for estimation includes variables which corresponds to state vector s and choice a in our model but the choice shock ϵ is not observed. That is, the observed data is $Y_T \equiv \left\{s^d_\tau, a^d_\tau, F^d\right\}_{\tau=1}^{T}$ [2], where

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

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

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 that case, Ψ is a set whose state variable is being an incumbent and the choice variable is staying in. Let $\pi()$ be the prior distribution of θ. Furthermore, let $L(Y_T|\theta)$ be the likelihood of the model, given the parameter θ and the value function $V()$, which is the solution of the Dynamic Programming problem. Then, we have the following posterior distribution function of θ.

$$P(\theta|Y_T) = \pi(\theta)L(Y_T|\theta) \tag{2.}$$

Because ϵ is unobserved to the econometrician, the likelihood is an integral over $\boldsymbol{\epsilon} \equiv \{\epsilon_\tau\}_{\tau=1}^{T}$. That is, if we define $L(Y_T|\boldsymbol{\epsilon},\theta)$ to be the likelihood conditional on $(\boldsymbol{\epsilon},\theta)$ and $F(\boldsymbol{\epsilon}|\theta)$ to be the distribution function of $\boldsymbol{\epsilon}$, then

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

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

$$L(Y_T|\theta) = \int L(Y_T|\boldsymbol{\epsilon},\theta)dF(\boldsymbol{\epsilon}|\theta)$$

The value function enters into the choice probability as follows.

$$P\left[a = a_i^d | s_i^d, V, \theta\right] = \Pr\left[\widehat{\epsilon}_i : a_i^d = \arg\max_{a\in A} \mathcal{V}(s_i^d, a, \widehat{\epsilon}_i, \theta)\right]$$

Choice probability is a component of the likelihood increment of the sample i. Below we describe the various estimation approaches that are possible, including the Bayesian dynamic programming algorithm we propose.

2.1 The Maximum Likelihood Estimation

The conventional estimation procedure of the dynamic programming problem consists of two main steps. First is the solution of the dynamic programming problem and the 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 additional 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, $V^{(0)}(s, \epsilon^{(m)}, \theta)$ for every $s \in S$, $\epsilon^{(m)}$. We also let the expected value function (Emax function) to be $\widehat{E}_{\epsilon'}\left[V^{(0)}(s, \epsilon', \theta)\right] = 0$ 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 future expected value function (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,

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

   This yields the value function,

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

   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.

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

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**

   Computationally, the most demanding part of the likelihood construction is the derivation of the choice probability $P\left[a = a_i^d | s_i^d, V, \theta\right]$. 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 utility. Also, denote,

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

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

$$P\left[a_i^d | s_i^d, V, \theta\right] = P\left[\epsilon_a - \epsilon_{a_i^d} \geq \widetilde{\mathcal{V}}^{(t)}(s, a, \theta) - \widetilde{\mathcal{V}}^{(t)}(s, a_i^d, \theta); a \neq a_i^d | s_i^d, V, \theta\right]$$

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

3. **Maximization routine**

   Now, supppose we have K parameters to estimate. In a typical Maximum Likelihood estimation routine, where one uses Newton hill climbing algorithm, 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 algorithn 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. As can be seen from the above discussion, the main difficulty lies in the fact that the inner loop has to be embedded within the outer loop. The computational burden would be greatly reduced if one can take the inner loop out of the outer loop so that the two loops can be computed simultaneously.

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', $f\left(\theta, \theta'\right)$, which satisfies, among other more technical conditions, the following equality.

$$P(\theta|Y_T) = \int f\left(\theta, \theta'\right) P\left(\theta'|Y_T\right) d\theta'$$

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

Gibbs Sampling is a popular example of the above MCMC strategy that is simple to implement. Gibbs sampling strategy decomposes the transition density $f\left(\theta, \theta'\right)$ into small blocks, where simulation from each block is straightforward. During each MCMC iteration, we also fill in the missing $\epsilon^{(t)}$ following the Data Augmentation strategy (See Tanner and Wong (1987) for more details of Data Augmentation).

The conventional Bayesian estimation method 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 Dynamic Programming 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 $Y_T \equiv \left\{s_\tau^d, a_\tau^d, F^d\right\}_{\tau=1}^{T}$, which does not include the latent shock $\boldsymbol{\epsilon} \equiv \{\epsilon_\tau\}_{\tau=1}^{T}$. In order to 'integrate out' the latent shock, we simulate $\boldsymbol{\epsilon}$. Since the optimal choice is given as a_i^d in the data, we need to simulate $\boldsymbol{\epsilon}$ subject to the constraint that for every sample i, given s_i^d, a_i^d is the optimal choice. That is,

$$a_i^d = \arg\max_{a \in A} \mathcal{V}(s_i^d, a, \widehat{\epsilon}_i, \theta^{(t)})$$

where $\widehat{\epsilon}_i$ is the data augmented shock for sample i.

Gibbs Sampling Step: 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. Let

$$\theta^{(t,-j)} \equiv \left(\theta_1^{(t+1)}, ..., \theta_{j-1}^{(t+1)}, \theta_{j+1}^{(t)}, ..., \theta_J^{(t)}\right).$$

Then,

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

where

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

and $\widehat{\boldsymbol{\epsilon}}$ 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.

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

Although MCMC techniques overcome the computational problem of high dimensionality of parameters, the second problem remains. Since the likelihood is a function of the value function, during the estimation algorithm, the Dynamic Programming 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 at the same time, 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 $f^{(t)}\left(\theta, \theta'\right)$, from which we simulate the sequence of parameters $\left\{\theta^{(\tau)}\right\}_{\tau=1}^{t}$ such that it converges to the correct posterior. We use Gibbs Sampling strategy and decompose the transition density $f^{(t)}\left(\theta, \theta'\right)$ into small blocks, where simulation from each block is straightforward. We also fill in the missing $\epsilon^{(t)}$ 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 Dynamic Programming algorithm only once. As a result of this, in our algorithm, the transition density $f^{(t)}\left(\theta, \theta'\right)$ changes with each iteration since the value function changes between iterations. Thus, the invariant distribution of the transition density $f^{(t)}\left(\theta, \theta'\right)$ in our algorithm varies with each iteration, and the invariant transition density at iteration t depends on the value function approximations derived at iteration t, $V^{(t)}$. The invariant distribution for iteration t is

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

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

$$P^{(t)}(\theta|Y_T) = \int f^{(t)}\left(\theta, \theta'\right) P^{(t)}\left(\theta'|Y_T\right) d\theta'$$

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

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

for any $\theta, \theta' \in \Theta$. Furthermore, we prove that the parameter simulations based on the MCMC using the above sequence of transition densities converges in probability to the parameter simulation generated by the MCMC using the true transition density $f\left(.,.\right)$.

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 $\Omega^{(t)} \equiv \left\{\epsilon^{(\tau)}, \theta^{(\tau)}, V^{(\tau)}\right\}_{\tau=1}^{t-1}$ be the history of shocks and parameters upto the previous iteration $t-1$. Let $\mathcal{V}^{(t)}(s, a, \epsilon^{(t)}, \theta^{(t)}, \Omega^{(t)})$ be the value of choice a and let $V^{(t)}(s, \epsilon^{(t)}, \theta^{(t)}, \Omega^{(t)})$ be the value function derived at iteration t of our solution/estimation algorithm. Then, the value function and the approximation $\hat{E}{}^{(t)}_{\epsilon'}\left[V(s', \epsilon', \theta)|\Omega^{(t)}\right]$ for the expected value function $E_{\epsilon'}\left[V(s', \epsilon', \theta)\right]$ at iteration t are defined recursively as follows.

$$\hat{E}{}^{(t)}_{\epsilon'}\left[V(s', \epsilon', \theta)|\Omega^{(t)}\right] \equiv \sum_{n=1}^{N(t)} V^{(t-n)}(s', \epsilon^{(t-n)}, \theta^{(t-n)} \mid \Omega^{(t-n)}) \frac{K_h(\theta^{(t)} - \theta^{(t-n)})}{\sum_{k=1}^{N(t)} K_h(\theta^{(t)} - \theta^{(t-k)})} \tag{5}$$

and

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

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

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

$$K_h(u) = \frac{1}{h} K(\frac{u}{h})$$

where K is a nonnegative continuous, bounded and symmetric real function which integrates to one. i.e. $\int K(u)du = 1$. Furthermore, we assume that $\int uK(u)du < \infty$. The approximated expected value function is the weighted average of value functions of $N(t)$ past iterations. The sample size of the average, $N(t)$, increases with t. Futhermore, we let $t - N(t) \to \infty$ as $t \to \infty$ as well. The weights are high for the value functions at iterations with parameters close to the current parameter vector θ. This is similar to the idea of Pakes and McGuire (2002), where the expected value function is the average of the past N iterations as well. 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 $s\prime$ as well as parameters that are close to the current parameter θ.

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

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

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

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

2. **Data Augmentation Step**: We simulate $\boldsymbol{\epsilon}$ subject to the constraint that for every sample i, given s_i^d, a_i^d is the optimal choice. That is,

$$a_i^d = \arg\max_{a \in A} \mathcal{V}^{(t)}(s_i^d, a, \widehat{\epsilon}_i, \theta^{(t)})$$

where $\widehat{\epsilon}_i$ is the data augmented shock for sample i. 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

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

and $\widehat{\boldsymbol{\epsilon}}$ is the data augmented shock. Then, given $\theta^{(t)}$, the transition density for the MCMC is 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 Dynamic Programming and Bayesian methods, nonparametric kernel techniques are also used to approximate the value function. Notice that 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.

It can be seen from the above description of the steps, that the Bellman equation step (Step 1) is only done once during a single estimation algorithm. 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.

2.4 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.

1. **Assumption 1**: Parameter space $\Theta \in R^J$ 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.

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

3. **Assumption 3**: We assume that β is known and $\beta < 1$

4. **Assumption 4**: For any $s \in S$, ϵ and $\theta \in \Theta$, $V^{(0)}(s, \epsilon, \theta) < M_I$ for some $M_I > 0$.

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

5. **Assumption 5**: Given V being uniformly bounded, π, L satisfy the following: $\pi(\theta)$ is positive and bounded for any $\theta \in \Theta$ and for any given $\boldsymbol{\epsilon}$, $L(Y_T | \boldsymbol{\epsilon}, \theta, V) > 0$ and bounded for any $\theta \in \Theta$.

6. **Assumption 6**: The support of ϵ is compact.

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

8. **Assumption 8**: For any $\theta \in \Theta$, a_i^d, s_i^d, $i = 1, ..., I$, V,

$$P\left[a = a_i^d | s_i^d, V, \theta\right] = \Pr\left[\widehat{\epsilon}_i : a_i^d = \arg\max_{a \in A} \mathcal{V}(s_i^d, a, \widehat{\epsilon}_i, \theta)\right] > 0$$

9 **Assumption 9**: 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 $t(2) - N(t(2)) = t(1) + 1$. Such $t(2)$ exists from the assumption that $N(t)$ is nondecreasing in t and $t - N(t) \to \infty$. Also, let $\widetilde{N}(2) = N(t(2))$. Similarly, for any $l > 2$, let $t(l+1)$ be such that $t(l+1) - N(t(l+1)) = t(l) + 1$, and let $\widetilde{N}(l+1) = N(t(l+1))$. Assume that there exists a finite constant $A > 0$ such that $\widetilde{N}(l+1) < AN(l)$.

Now, we state the main theorem 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 function $V^{(t)}(s, \epsilon, \theta)$ converges in probability to $V(s, \epsilon, \theta)$ as $t \to \infty$ and $h \to 0$. Also, $\hat{E}_{\epsilon'} \left[V^{(t-1)}(s', \epsilon', \theta^{(t)}) | \Omega^{(t)} \right]$ converges to $E_{\epsilon'} \left[V(s', \epsilon', \theta) \right]$ in probability uniformly.

Proof of Theorem 1 is discussed in the Appendix.

Notice that

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

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

Let

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

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

Then, because $V^{(t)}\to V$ in probability uniformly, and because of the compactness of Θ and support of ϵ, Theorem 1 implies that $f^{(t)}\left(\theta,\theta'\right)$ converges to $f\left(\theta,\theta'\right)$ in probability uniformly.

Theorem 2

Suppose assumptions 1 to 9 are satisfied for $V^{(t)}$, 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 $t\to\infty$ and $h\to 0$. Then, $\theta^{(t)}$ converges to $\widetilde{\theta}^{(t)}$ in probability, where $\widetilde{\theta}^{(t)}$ is a Markov chain with transition density function being f.

Proof of theorem 2 is shown in Appendix 1. Despite the lengthy formal proofs, the basic logic is more straightforward. First, suppose the parameter $\theta^{(t)}$ stays fixed at a value θ^* for all iterations. Then, equation (5) becomes as follows.

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

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, hence in probability. Now, instead of $\theta^{(t)}$ being constant at θ^*, assume that for any $t=1,2,...$, $\theta^{(t)}$ stays within a small open ball around θ^*. Then, after some iterations, the value functions derived from the Dynamic programming algorithm will move closely around the true value function for the parameter θ^* most of the time, because of continuity of value function. Now, let us reconsider the original Bayesian Dynamic Programming algorithm. Because of the assumptions, for any parameter vector $\theta\in\Theta$, the Bayesian MCMC algorithm will produce a sequence of parameters which contains a subsequence $\theta^{(\tau_t)}$ that stays within a small open ball around θ. Because of the compactness, every open cover of

Θ has a subcover, hence after some iterations, the solution of the Dynamic Programming problem will move closely around the true value function uniformly over Θ, most of the time.

Our simultaneous solution and estimation algorithm also can be applied to other settings of dynamic discrete choice models, with some minor modifications. One example is Rust (1997) Random grid approximation. There, given continuous state space vector s, and action a and parameter θ, the transition function from state vector s to next period state vector $s\prime$ is defined to be $f_s(s'|a,s,\theta)$. Then, the Dynamic Programming part of our algorithm is defined as below.

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

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

where s' is the next period state varible, which is assumed to follow the transition function $f_s(s'|a,s,\theta)$. $\hat{E}_{s',\epsilon'}\left[V(s',\epsilon',\theta)|\Omega^{(t)}\right]$ is defined to be the approximation for the expected value function. Furthermore, the value function is defined to be as follows.

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

Let $\theta^{(t-n)}$ be the parameter vector, $\epsilon^{(t-n)}$ be the shock at iteration $t-n$, and let $\Omega^{(t)} \equiv \left\{\epsilon^{(\tau)},\theta^{(\tau)}\right\}_{\tau=1}^{t-1}$. Furthermore, let $K_h(.)$ be the kernel function with bandwidth h. Then, $\hat{E}_{s',\epsilon'}\left[V(s',\epsilon',\theta)|\Omega^{(t)}\right]$ is defined to be as follows.

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

$s^{(\tau)}, \tau = 1,2,...$ are drawn randomly from an i.i.d. distribution. 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 different state vector for each solution/estimation iteration. Hence, even though we only draw one state vector $s^{(\tau)}$at each iteration, the number of random grid points is $N(t)$, which can be made arbitrarily large when we increase the number of iterations. In Rust (1997), if the grid size is N, then the number of computations required for each Dynamic Programming iteration is N. Hence, at iteration τ, the number of Dynamic Programming computations that is required is $N\tau$. In our case as long as $\tau > N(t)$, the total number of Dynamic Programming computation required is τ, which does not depend on the grid size. In other words, the accuracy of the Dynamic Programming computation automatically increases with the iteration.

Rust (1997) assumes that the transition density function $f_s(s'|a,s,\theta)$ 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 those 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 the state variables. 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.

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

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

where s' is the next period state varible, which is assumed to be a deterministic function of s, a, and θ. That is,

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

$\hat{E}_{\epsilon'} \left[V(s', \epsilon', \theta) | \Omega^{(t)} \right]$ is defined to be the approximation for the expected value function. Furthermore, the value function is defined to be as follows.

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

Furthermore, let $K_h(.)$ be the kernel function with bandwidth h. Let Then, it is defined to be as follows.

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

3 Examples

We estimate a simple dynamic discrete choice model of entry and exit, with firms in competitive environment.[3] The firm is either an incumbent (I) or a potential entrant (O). If the incumbent firm chooses to stay, its per period return is,

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

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,

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

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

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.

$$\epsilon_{lt} \sim N(0, \sigma_l)$$

The level of capital K_t evolves as follows.

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

where

$$u_t \sim N(0, \sigma_u),$$

if the incumbent firm decides to stay in, and

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

if the potential entrant decides to enter.

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,

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

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

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

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,

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

is the value of entering during period t and

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

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, $i_{i,t}^d \in \{IN, OUT\}$, that is, whether the firm is an incumbent or a potential entrant.

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

$$\{K_{i,t}^d, \pi_{i,t}^d, i_{i,t}^d\}_{i=1,N_d}^{t=1,T}$$

where,

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

if the firm stays in and 0 otherwise.

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

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

$$\delta_E \text{ ~ } N(\underline{\delta}_i, A_E^{-1})$$

where $\underline{\delta}_x$, $\underline{\delta}_i$ are the prior means and A_x, A_E are the prior precision (inverse of variance) of the exit value and the entry cost, respectively.

For parameters α, b_1, b_2 and b_e, we assume the priors to be uninformative.

We also assume independent Chi squares prior for the precision of the shocks ϵ_1 and u which is the inverse of their variance, i.e. $h_{\epsilon 1} = (\sigma_1^2)^{-1}$, for ϵ_1. That is,

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

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

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

Furthermore, .

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

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 for the next iteration, i.e., $V_\Gamma^{(s+1)}(K, \epsilon^{(s)}, \theta^{(s)}, \Omega^{(s)})$.

1) Suppose we already have calculated the approximation for the expected value function, where the expectation is over the choice shock ϵ, that is, $\widehat{E}_\epsilon^{(s)} V_\Gamma(K'(K, u^{(s)}), \epsilon, \theta^{(s)}|\Omega^{(s)})$. To further integrate the value function over the capital shock u, we use the random grid integration method of Rust (1997). That is, given 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,

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

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.

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

3) Given $\epsilon^{(s)}$ and $\widehat{E}^{(s)} V_\Gamma(K, \epsilon, \theta^{(s)}|\Omega^{(s)})$, solve for 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:

$$\begin{aligned} V_\Gamma^{(s+1)}(K, \epsilon^{(s)}, \theta^{(s)}, \Omega^{(s)}) \quad = \quad & Max\{R_{\Gamma,IN}(K, \epsilon_1^{(s)}, \theta^{(s)}) + \beta \widehat{E}^{(s)}\left[V_I(K'(K, u), \epsilon, \theta^{(s)})|\Omega^{(s)}\right], \\ & R_{\Gamma,OUT}(K, \epsilon_2^{(s)}, \theta^{(s)}) + \beta \widehat{E}^{(s)}\left[V_O(K'(K, u), \epsilon, \theta^{(s)})|\Omega^{(s)}\right]\} \end{aligned}$$

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:

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

and

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

Similarly, for the potential entrant, we define

$$\overline{V}_{O,IN}(K, \theta^{(s)}, \Omega^{(s)}) = -\delta_E^{(s)} + \beta \widehat{E}^{(s)}\left[V_I(K_0, \epsilon, \theta^{(s)})|\Omega^{(s)}\right],$$

and

$$\overline{V}_{O,OUT}(K, \theta^{(s)}, \Omega^{(s)}) = \beta \widehat{E}^s V_O(V_O(0, \epsilon, \theta^{(s)})|\Omega^{(s)}).$$

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

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

The empirical economist does not observe the above statistics directly because he can only obtain data on entry and exit decisions $i_{i,t}^d$ 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, $i_{i,t}^d = IN$, then draw $\eta_{i,t} = \epsilon_{1,i,t} - \epsilon_{2,i,t}$ such that

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

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

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

Once the shock $\widehat{\eta}_{i,t}$ is generated, he can proceed to recover the entry cost and exit value parameters by conducting Bayesian regression of $r_{i,t}^{(s+1)}$ on entry and exit decisions, 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

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

$$\sigma_v^2 = \sigma_1^{(s)2} - \frac{\sigma_1^{(s)4}}{\sigma_1^{(s)2} + \sigma_2^{(s)2}} = \frac{\sigma_1^{(s)2}\sigma_2^{(s)2}}{\sigma_1^{(s)2} + \sigma_2^{(s)2}}$$

and

$$\gamma_1 = \frac{Cov(\epsilon_{1t}, \eta_t)}{\sigma_\eta^2} = \frac{\sigma_1^{(s)2}}{\sigma_1^{(s)2} + \sigma_2^{(s)2}}.$$

Once the profit for firms who exited or stayed out is recovered, we can recover productivity parameters via a simple Bayesian regression.

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

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

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

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 $\mathbf{w}^{(s+1)}$, $\boldsymbol{\pi}^{(s+1)}$ to be the stacked vector of $w_{i,t}^{(s+1)}$ and $\pi_{i,t}^{(s+1)}$.

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

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

where,

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

and

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

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,

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

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

$$k_t = \ln(K_t), \quad \mathbf{k}_{-1} = [k_{11}, k_{12}, ..., k_{1T-1}, ..., k_{N_d1}, k_{N_d2}, ..., k_{N_dT-1}]$$

and

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

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

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

where,

$$\overline{A}_\alpha = (A_\alpha + h_\alpha^{(s)} \mathbf{k}'\mathbf{k})^{-1}$$

and

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

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

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

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

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

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 $(\mathbf{k}, h_b^{(s)})$ as follows.

$$b^{(s+1)} | (\mathbf{k}^{(s+1)}, h_\alpha^{(s)}) \ \tilde{}\ N(\overline{b}, \overline{A}_b),$$

where,

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

and

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

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

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

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

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

Expected Value Function Iteration Step

Next, we update the expected value function for iteration $s+1$, that is, we derive $E_\epsilon^{(s+1)} V_\Gamma(K, \epsilon, \theta^{(s)})$. This is an important step in the alogorithm and is closely related to the algorithm of Pakes and McGuire (2001).

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

where $I()$ is the indicator function, and $K()$ is the kernel function. We adopt the following Gaussian kernel:

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

The expected value function is updated by taking the average over those past $N(s)$ iterations where the parameter vector $\theta^{(j)}$ was close to $\theta^{(s+1)}$. The similarity with PM is that the expected value functions are approximated by averaging over past values of the algorithm, that is, they are never explicitly calculated. Also, the optimization problem is solved only once between iterations. The main difference is that past values are weighted according to the distance between their parameter vectors and the current parameter vector: the shorter is the distance, the higher is the weight.

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 20 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 function $\frac{1}{M}\sum_{m=1}^{M} V_\Gamma(K_m, \epsilon_m^{(j)}, \theta^{(j)})$, where $M = 20$. This obviously increases the accuracy per iteration, and reduces the need to have a large $N(s)$. That is partly why in the below examples, to have $N(s)$ increase up to 2000 turned out to be sufficient for a good estimation performance.

After the above 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 $E^{(s+1)}V(K, \epsilon, \theta^{(s+1)}, \Omega^{(s+1)})$ 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.

4 Simulation and Estimation Exercise.

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 the above DP solution. Then, using the simulated data, we try to estimate the parameter values using the Bayesian DP estimation method. Below, we briefly explain how we solved for the DP problem to generate the data. Notice that for data generation, we only need to solve for 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.

Assume that we already know the expected value function of the s th iteration for all capital grid points.

$$E_\epsilon V(K_m, \epsilon, \theta^*), \quad m = 1, 2, ..., M.$$

Then, following steps are taken to generate the expected value function for $s+1$ th iteration.

Step 1 Given capital stock K, derive

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

for $\Gamma \in \{I, O\}$. Here, K_m $(m = 1, ..., M)$ are grid points and $f(K_m|K, \theta^{(s)})$ is the transition probability from K to K_m

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

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

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

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

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 chocie sequences $\{K_{i,t}^d, \pi_{i,t}^d, i_{i,t}^d\}_{i=1,N_d}^{t=1,T}$ 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 $\beta^* = 0.9$.

4.1 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.

$$\delta_x \sim N(\underline{\delta}_x, A_x^{-1}),\ \underline{\delta}_x = 0.4,\ \underline{A}_x = 1.0,$$

$$\delta_E \sim N(\underline{\delta}_E, A_E^{-1}),\ \underline{\delta}_E = 0.4,\ \underline{A}_E = 1.0$$

$$\alpha \sim N(\underline{\alpha}, A_\alpha^{-1}),\ \underline{\alpha} = 0.2,\ \underline{A}_\alpha = 1.0$$

$$\underline{s_{\epsilon_1}}^2 h_{\epsilon_1} \sim \chi^2(\nu_{\epsilon_1}),\ (\underline{s_{\epsilon_1}}^2)^{-1} = 0.4,\ v_{\epsilon_1} = 400.$$

$$\underline{s_\eta}^2 h_\eta \sim \chi^2(\nu_\eta),\ (\underline{s_\eta}^2)^{-1} = 0.32,\ v_\eta = 400.$$

$$\underline{s_u}^2 h_u \sim \chi^2(\nu_u),\ \underline{s_u}^2 = 1.0,\ v_u = 400.$$

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

We set the initial guess of the expected value function to be 0. We set the initial guesses of the parameters to be the true parameter values given by θ^*. The Gibbs sampling was conducted 10,000 times. The Gibbs sampler for the simulation with sample size 10,000 is shown in figures 1 to 9. In estimation experiments with other sample sizes, the Gibbs sampler converged from around 3,000 iterations as well. 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 some what 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 becomes closer to the truth, even though there are some exceptions.

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. 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 the capital coefficient in the profit equation, α. We can see that there, the value of the first Gibbs sampler jumps from the true value, 0.2 to 0.2367. The upward bias is due to the sample selection bias. However, immediately after a couple of iterations, the Gibbs sampler estimates the true value quite accurately. That is, the algorithm can immediately correct for the sample selection bias even though it has not fully solved the Dynamic programming problem. The Gibbs sampler of the other parameters are reported in Figures 4 to 9. There, we see that all the parameters stay closely around the true value from the start. In a separate experiement, we also have conducted experiments where we set the initial values of the parameters to half the true values and run the Gibbs sampler. As we can see from Table 1, the results turns out to be hardly different from the original ones. In sum, our Bayesian DP algorithm initially estimates the static version of the dynamic model. As we iterate the algorithm, expected value functions become closer to the truth and hence the estimates converge to the true posteriors. These results confirms 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	0.4287 (0.0175)	0.3696 (0.0110)	0.3993 (0.0091)	0.4
δ_E	0.4792 (0.0182)	0.4074 (0.0116)	0.4182 (0.0085)	0.4
α	0.1992 (0.0070)	0.1964 (0.0048)	0.1951 (0.0043)	0.2
σ_1	0.4033 (0.0056)	0.4056 (0.0042)	0.4058 (0.0041)	0.4
σ_2	0.3940 (0.0198)	0.3858 (0.0188)	0.3867 (0.0228)	0.4
b_1	0.0971 (0.0201)	0.1011 (0.0131)	0.1011 (0.0093)	0.1
b_2	0.0882 (0.0370)	0.0978 (0.0239)	0.0966 (0.0173)	0.1
b_e	−0.9932 (0.0141)	−0.9799 (0.0086)	−0.9920 (0.0062)	−1.0
σ_u	0.4102 (0.0045)	0.4031 (0.0030)	0.4022 (0.0021)	0.4
sample size	2,000	5,000	10,000	
CPU time[4]	18 min.10 sec.	41 min.12 sec.	1 hr. 18 min. 59 sec.	

parameter	estimate[5]	true value
δ_x	0.3774 (0.0087)	0.4
δ_E	0.3967 (0.0090)	0.4
α	0.1956 (0.0027)	0.2
σ_1	0.4053 (0.0035)	0.4
σ_2	0.3895 (0.0191)	0.4
b_1	0.1010 (0.0103)	0.1
b_2	0.0961 (0.0062)	0.1
b_e	−0.9923 (0.0595)	−1.0
σ_u	0.4022 (0.0022)	0.4
sample size	10,000	
CPU time[6]	1 hr. 19 min. 5 sec.	

4.2 Experiment 2: Random Effects

We now report estimation results of a model that includes observed and unobserved heterogeneity. We assume that the profit coefficient for each firm i, $\alpha_i$is distributed normally with mean $\mu = 2.0$ and standard error $\sigma_\alpha = 0.04$. Furthermore, we include observed characteristics in our model as well. That is, the transition equation for capital now 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

[4] The estimation exercise was done on a Sun Blade 2000 workstation.

[5] This is the results for the different starting values.

[6] The estimation exercise was done on a Sun Blade 2000 workstation.

draw 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 so far have only used finite number of types, typically less than 10, as an approximation of the random effect. 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 will be shown below.

We set the prior for α_i as follows.

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

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

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

Then, if we denote $\boldsymbol{\alpha'} = (\alpha_1, ..., \alpha_N)$, $\boldsymbol{\pi'} = (\pi_{11}, \pi_{12}, ..., \pi_{1T}, ..., \pi_{N_d 1}, ..., \pi_{N_d T})$ 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 $K_j = [K_{j1}, K_{j2}, ..., K_{jT}]$. Also, e_N is a N by 1 vector of ones.

Then, the prior can be expressed as follows.

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

Let $\theta_{-\alpha}^{(s)}$ be defined as parameters not including α_i. Below, we briefly describe the differences between the earlier estimation routine and that which involves random effects.

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

$$r_{i,t}^{(s+1)} = R_{\Gamma,OUT}(K_{i,t}^d, \epsilon_2, \theta_{-\alpha}^{(s)}, \alpha_i^{(s)}) - R_{\Gamma,IN}(K_{i,t}^d, \epsilon_1, \theta_{-\alpha}^{(s)}, \alpha_i^{(s)}) + \alpha_i^{(s)} K_{i,t}^d \equiv g(K_{i,t}^d, \epsilon_{1,i,t} - \epsilon_{2,i,t}, \theta_{-\alpha}^{(s)}, \alpha_i^{(s)}).$$

- To draw $\eta_{i,t} = \epsilon_{1,i,t} - \epsilon_{2,i,t}$ we follow the below data augmentation steps.

$$\eta_{i,t} \geq \overline{V}_{\Gamma,OUT}(K_{i,t}^d, \theta_{-a}^{(s)}, a_i^{(s)}) - \overline{V}_{\Gamma,IN}(K_{i,t}^d, \theta_{-a}^{(s)}, a_i^{(s)}).$$

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

$$\eta_{i,t} < \overline{V}_{\Gamma,OUT}(K_{i,t}^d, \theta_{-a}^{(s)}, a_i^{(s)}) - \overline{V}_{\Gamma,IN}(K_{i,t}^d, \theta_{-a}^{(s)}, a_i^{(s)}).$$

As we discussed earlier, once the shock $\eta_{i,t}$ is generated, he can proceed to recover the entry cost and exit value parameters by conducting Bayesian regression of $r_{i,t}^{(s+1)}$ on entry and exit decisions, using the following linear relationship.

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

In contrast to the earlier case, 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 standard case 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. The posterior draw of $\boldsymbol{\alpha}$ for iteration s, $\boldsymbol{\alpha}^{(s+1)}$, can be done from the following distribution.

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

with

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

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

One-Step Bellman Equation and Expected Value Function Iteration Step

The only difference between the earlier case is that we solve the one step Bellman equation for each firm i separately. The expected value function. $E_\epsilon^{(s+1)} V(K, \epsilon, \theta_{-\alpha}^{(s+1)}, \alpha_i^{(s+1)})$ is derived as follows.

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

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 those of computing the DP problem again for each firm i, for each simulation draw of α_i as would be done in 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	0.3967 (0.0140)	0.4
δ_E	0.4058 (0.0131)	0.4
$\overline{a}$	0.2086 (0.0053)	0.2
τ	0.0396 (0.0013)	0.04
σ_1	0.4027 (0.0053)	0.4
σ_2	0.3964 (0.0279)	0.4
b_1	0.1006 (0.0137)	0.1
b_2	0.1020 (0.0264)	0.1
b_e	-0.9661 (0.0103)	-1.0
σ_u	0.4059 (0.0034)	0.4
sample size	100×100	
CPU time[7]	10 hrs 47 min 26 sec	

Notice that most of the parameters are close to the true values. The computation time is about 11 hours, which roughly corresponds to those required for a Bayesian estimation of a reasonably complicated static random effects model.

We also conducted some estimation exercise using the conventional simulated ML routine. For each firm, we simulated α_i 100 times (i.e. $M_\alpha = 100$). When we solve for the DP problem, we use Monte-Carlo integration to integrate over the choice shock ϵ. We set the simulation size for ϵ to be 100. A single likelihood calculation took about 35 minues to compute. A single step of the Newton-Raphson method took 11 likelihood calculations. Since we took numerical derivatives, in addition to the likelihood evaluation under the original parameter θ, we calculated the likelihood for the 9 parameter perturbations $\theta + \Delta\theta_i$, $i = 1, ..., 9$. After computing the search direction, we further calculate the likelihood twice to derive the step size. The above computation took us in total 6 hours and 20 minues. By that time, Bayesian DP routine would have completed its $6,744$ iterations. That is, by the time the conventional ML routine finished its first iteration, the Bayesian DP routine would have already converged long ago.

Another estimation strategy for the simulated ML could be to expand the state variables of the DP problem to include both X and a. Then, we have to assign grid points for the 3 dimensional

[7] The estimation exercise was done on a Sun Blade 2000 workstation.

state space points (K, X, a). 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.

4.3 Experiment 3: Infinite Random Grids

As discussed above, instead of fixing the capital grid points throughout the DP solution/estimation algorithm, we we draw different state vector for each solution/estimation iteration. Hence, even though we only draw finite state vector grid points $K_1^{(t)}, ..., K_{M_K}^{(t)}$ (in this example, $M_K = 10$), the number of random grid points can be made arbitrarily large when we increase the number of iterations. That is, the formula for the expected value function for the firm who stays in or enters is as follows.

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

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.

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

We increase the total number of grid points up to 2000.

Table 3 shows the estimation results. We can see that the estimates parameters are close to the true ones. The entire exercise took about 8 hours.

Table 3: Posterior Means and Standard Errors

(standard errors are in parenthesis)

parameter	estimate	true value
δ_x	0.4246 (0.0121)	0.4
δ_E	0.4341 (0.0133)	0.4
a	0.2036 (0.0031)	0.2
σ_1	0.4011 (0.0046)	0.4
σ_2	0.3946 (0.0198)	0.4
b_1	0.1001 (0.0165)	0.1
b_2	0.1033 (0.0097)	0.1
b_e	−0.9844 (0.0097)	−1.0
σ_u	0.4018 (0.0039)	0.4
sample size	10,000	
CPU time[8]	7 hrs39 min 31 sec	

4.4 Experiment 4: Continuous State Space with Deterministic Transition

As discussed above, instead of fixing the capital grid points throughout the DP solution/estimation algorithm, we draw different state vector for each solution/estimation iteration. Hence, even though we only draw finite state vector grid points $K_1^{(t)}, ..., K_{M_K}^{(t)}$ (in this example, $M_K = 10$), the number of random grid points can be made arbitrarily large when we increase the number of iterations.

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

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

That is, the formula for the expected value function for the incument who stays in is as follows.

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

[8]The estimation exercise was done on a Sun Blade 2000 workstation.

where K_{h_K} is the kernel for the capital stock with bandwidth h_K. The expected value function for the entrant is:

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

The formula for the expected value function for either the firm who stays out or the firm who exits is similar as before:

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

We increase the total number of grid points up to 2000.

Table 4 shows the estimation results. We can see that the estimates parameters are close to the truth. The entire exercise took about 5 hours.

Table 4: Posterior Means and Standard Errors

(standard errors are in parenthesis)

parameter	estimate	true value
δ_x	0.3531 (0.0117)	0.4
δ_E	0.3688 (0.0126)	0.4
a	0.0979 (0.0041)	0.1
σ_1	0.4006 (0.0142)	0.4
σ_2	0.4011 (0.0290)	0.4
b_1	0.2180 (0.0222)	0.2
σ_u	0.4005 (0.0142)	0.4
sample size	10,000	
CPU time[9]	5 hrs 6 min 23 sec	

[9] The estimation exercise was done on a Sun Blade 2000 workstation.

5 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 the 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, where he keeps 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 we calculate the Expected value function on is small, the total number of "effective" state space points we evaluate over the entire solution/estimation iteration, which is $N(t)$ in our case, grows with the number of Bayesian DP iterations. The number of the Bayesian DP iterations 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". That is why our nonparmetric approximation of the expected value function works well under the assumption of continuous state space even though the transition function of the state variable is not stochastic. In that case, it is well known that Rust (1997) random grid method faces computational difficulties.

But it is worth mentioning that our algorithm does not come without any cost. 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. But on this issue, so far we remain fairly optimistic. The reason is because 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 function 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.

6 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

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

Davidson, James (1994) "Stochastic Limit Theory, " Oxford 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.

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

6.0.1 Appendix

Appendix 1: Proof of Theorem 1.

What we need to show is that for any $s \in S$, ϵ, $\theta \in \Theta$,

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

But since

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

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

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

Define

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

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.

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

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

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

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

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

$$\equiv A_1 + A_2 + A_3$$

Notice that 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. Lemma 1 below is used for the derivation of the asymptotic lower bound. The lemma shows that the transition density of the parameter process has an important property: regardless of the current location of the parameter or the number of steps, every parameter value is visited with a positive probability.

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

Proof. Recall that

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

By assumptions 1 (Compactness of parameter space), 5 (Strict Positivity and Boundedness of π and L), and 6 (Compactness of support of ϵ), there exist $\eta_1, \eta_2, M_1, M_2 > 0$, such that,

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

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

Therefore, for any θ_j,

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

exists and is positive. Let

$$h(\theta_j) \equiv \inf_{\widehat{\epsilon}, \theta^{(t,-j)}} p^{(t)} \left(\theta_j | \theta^{(t,-j)} \right)$$

Now, define

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

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

$$\varepsilon_0 g(\theta) = \prod_{j=1}^{J} h(\theta_j) \leq \prod_{j=1}^{J} p^{(t)}\left(\theta_j | \theta^{(t,-j)}\right) = f^{(t)}\left(\theta^{(t-1)}, \theta\right).$$

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

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

Proof. Using similar logic as in Lemma 1, one can show that for any θ_j,

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

exists and is bounded. Let

$$\widetilde{h}(\theta_j) \equiv \sup_{\widehat{\epsilon}, \theta^{(-j)}} p^{(t)}\left(\theta_j^{t+1} | \theta^{(t,-j)}\right)$$

Now, let

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

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

Lemma 3 $A_1 \stackrel{P}{\to} 0$ *as* $t \to \infty$

Proof. Recall that,

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

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 as $t \to \infty$ and the denominator is bounded below by a positive number.

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.,

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

Also, using the standard definition of the kernel function, we obtain,

$$\left|X_{N(t)n}\right| \leq \frac{1}{N(t)} \left| \int V(s', \epsilon', \theta) dF_{\epsilon'}(\epsilon', \theta) - V(s', \epsilon^{(t-n)}, \theta) \right| \left[\frac{\sup |K|}{h(N(t))} \right]$$

Now, let

$$C_{N(t)n} = \frac{1}{N(t) h\left(N(t)\right)},$$

where $h(N(t))$ is assumed to satisfy $N(t) h\left(N(t)\right)^2 \to \infty$ as $N(t) \to \infty$. Then,

$$\left| \frac{X_{N(t)n}}{C_{N(t)n}} \right| \leq \left| \int V(s', \epsilon', \theta) dF_{\epsilon'}(\epsilon', \theta) - V(s', \epsilon^{(t-n)}, \theta) \right| \sup |K|$$

Because $V(s', \epsilon^{(t-n)}, \theta)$ is L_2-bounded and $\sup |K|$ is bounded, $\frac{X_{N(t)n}}{C_{N(t)n}}$ is also uniformly L_2-bounded, and thus, uniformly integrable. Furthermore,

$$\lim_{N(t) \to \infty} \sum_{n=1}^{N(t)} C_{N(t)n}^2 = \lim_{N(t) \to \infty} \frac{1}{N(t) h\left(N(t)\right)^2} = 0$$

Therefore, the assumptions for the Corollary 19.10 of Davidson (1994) are satisfied. Hence,

$$\sum_{n=1}^{N(t)} X_{N(t)n} \xrightarrow{L_2} 0$$

Therefore,

$$\sum_{n=1}^{N(t)} X_{N(t)n} = \frac{1}{N(t)} \sum_{n=1}^{N(t)} \left[\int V(s', \epsilon', \theta) dF_{\epsilon'}(\epsilon', \theta) - V(s', \epsilon^{(t-n)}, \theta) \right] K_h(\theta - \theta^{(t-n)}) \xrightarrow{P} 0 \qquad \text{(A1)}$$

as $N(t) \to \infty$. This shows that the numerator in $\frac{A_1}{\beta}$ goes to zero. We next show that the denominator is bounded below by using an argument similar to coupling theory (see Rosenthal, 1995) and the law of large numbers.

Let

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

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

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

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,

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

Furthermore, 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., the standard results on Kernel smoothing holds (see Haerdle (1989)) and

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

Hence, the Law of Large Numbers holds. Therefore, for any $\eta_1 > 0$, $\eta_2 > 0$, there exists $\overline{N} > 0$ such that for any $N(t) > \overline{N}$, $t > N(t)$

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

That is, for any $\eta_1 > 0$, $\eta_2 > 0$, there exists $\overline{N} > 0$ such that for any $N(t) > \overline{N}$, $t > N(t)$,

$$\Pr\left[\frac{1}{N(t)} \sum_{n=1}^{N(t)} Y^{(t-n)} + \eta_1 > \varepsilon_0 g(\theta)\right] > 1 - \eta_2 \tag{A5}$$

Choose $\eta_1 < \frac{1}{2}\epsilon_0 g(\theta)$.Then,

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

Because $\sum_{n=1}^{N} K_h\left(\theta - \theta^{(t-n)}(f^{(t-n)})\right) > \sum_{n=1}^{N} Y^{(t-n)}$, we conclude that for any $\eta_2 > 0$, there exists $\overline{N}$ such that for any $N(t) > \overline{N}$, $t > N(t)$

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

From $A1$ and $A6$, we can see that for any $\eta_1 > 0$, $\eta_2 > 0$, there exists $\overline{N}$ such that for any $N(t) > \overline{N}$, $t > N(t)$,

$$\Pr\left[\frac{\frac{1}{N(t)}\sum_{n=1}^{N(t)}\left[\int V(s',\epsilon',\theta)dF_{\epsilon'}(\epsilon',\theta) - V(s',\epsilon^{(t-n)},\theta)\right]K_h(\theta-\theta^{(t-n)})}{\frac{1}{N(t)}\sum_{n=1}^{N(t)}K_h\left(\theta-\theta^{(t-n)}(f^{(t-n)})\right)} < \frac{\eta_1}{\frac{1}{2}\varepsilon_0 g(\theta)}\right] > 1-\eta_2.$$

Since $\frac{\eta_1}{\frac{1}{2}\varepsilon_0 g(\theta)}$ can be made arbitrarily small by choosing $\eta_1 > 0$ small enough, we have shown that

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

■

Lemma 4

$$A_2 \xrightarrow{P} 0 \ \textit{as } N(t) \to \infty$$

Proof.

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

where $\delta > 0$ is arbitrarily set.

Step 1 of Lemma 4: Show that $H_2 \xrightarrow{P} 0$.

Note that

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

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

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

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,

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

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

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

Now, by change of variables,

$$\begin{aligned}
\int_{\left|\theta - \widetilde{\theta}\right| > \delta} \frac{1}{h}K\left(\frac{\theta - \widetilde{\theta}}{h}\right) g\left(\widetilde{\theta}\right) d\widetilde{\theta} &= \int_{|a| > \frac{\delta}{h}} K\left(a\right) g\left(\theta - ah\right) da \\
&\leq \sup_{\theta \in \Theta} g\left(\theta\right) \int_{|a| > \frac{\delta}{h}} K\left(a\right) da \qquad \text{(A12)}
\end{aligned}$$

Because $\int K\left(a\right) da = 1$ and $\int aK\left(a\right) da$ is bounded by assumption, $\int_{|a| > \frac{\delta}{h}} K\left(a\right) da \to 0$ as $h \to 0$. Therefore, RHS of $A12 \to 0$ as $h \to 0$ and thus,

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

as $h(N(t)) \to 0$.

Now, consider the denominator of $(A9)$. Using (A6) from Lemma 3, for $t > T$,

$$\begin{aligned}
& \Pr\left[2K\frac{\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)})I\left(\left|\theta-\theta^{(t-n)}\right|>\delta\right)}{\sum_{k=1}^{N(t)} K_h(\theta-\theta^{(t-k)})} < \frac{2K\eta}{\frac{1}{2}\varepsilon_0 g(\theta)}\right] \\
> \; & \Pr\{\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)}(f^{(t-n)})) > \frac{1}{2}\varepsilon_0 g(\theta)\right] \\
& \bigcap\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)})I\left(\left|\theta-\theta^{(t-n)}\right|>\delta\right)<\eta\right]\} \\
\geq \; & 1-\Pr\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)}(f^{(t-n)})) \leq \frac{1}{2}\varepsilon_0 g(\theta)\right] \\
& -\Pr\left[\frac{1}{N(t)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-n)})I\left(\left|\theta-\theta^{(t-n)}\right|>\delta\right)\geq\eta\right] \\
> \; & 1-\eta_2-\frac{\varepsilon_1}{\eta}E\left[K_h(\theta-\theta^{(t-n)}(g))I\left(\left|\theta-\theta^{(t-n)}(g)\right|>\delta\right)\right].
\end{aligned}$$

Notice that $g(\theta) > 0$. Hence, $\frac{2KN_A\eta}{\frac{1}{2}\varepsilon_0 g(\theta)}$ can be made arbitrarily small by choosing $\eta > 0$ small enough. Given $\eta > 0$, $\frac{\varepsilon_1}{\eta}E\left[K_h(\theta-\theta^{(t-n)}(g))I\left(\left|\theta-\theta^{(t-n)}(g)\right|>\delta\right)\right]$ can also be made arbitrarily small by choosing h to be small enough.

Thus, we have shown that $H_2 \xrightarrow{P} 0$ as $t \to \infty$, $h \to 0$.

Step 2 of Lemma 4: Show that $H_1 \xrightarrow{P} 0$.

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

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

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

Together, we showed that $A_2 \xrightarrow{P} 0$ as $h \to 0$ ■

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

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

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

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

From Lemma 3 and Lemma 4, we conclude that

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

$$+\beta \left[\sum_{n=1}^{N(t)} \left[V(s', \epsilon^{(t-n)}, \theta^{(t-n)}) - V^{(t-n)}(s', \epsilon^{(t-n)}, \theta^{(t-n)}, \Omega^{(t)}) \right] W_{N(t),h}(\theta, \theta^{(t-n)}, \Omega^{(t)}) \right] \quad \text{(A15)}$$

where

$$A^{(t)}(\theta) = A_1 + A_2 \xrightarrow{P} 0.$$

Because this holds for any $\theta \in \Theta$ and Θ is compact, convergence is uniform. Because $A^{(t)}\left(\theta^{(t)}\right) \leq \sup_{\theta \in \Theta} A^{(t)}(\theta)$,

$$A^{(t)}(\theta^{(t)}) \xrightarrow{P} 0$$

Taking supremum over a and then taking absolute values on both sides of equality A15, we obtain:

$$\begin{aligned} & \left| V(s, \epsilon, \theta) - V^{(t)}\left(s, \epsilon, \theta, \Omega^{(t)}\right) \right| \\ \leq & \left| \sup A^{(t)}(\theta) \right| + \beta \left[\sum_{n=1}^{N(t)} \left| V(s', \epsilon^{(t-n)}, \theta^{(t-n)}) - V^{(t-n)}(s', \epsilon^{(t-n)}, \theta^{(t-n)}, \Omega^{(t)}) \right| W_{N(t),h}(\theta, \theta^{(t-n)}, \Omega^{(t)}) \right] \quad \text{(A15')} \end{aligned}$$

Lemma:

For $t' > t$, let

$$\widetilde{W}\left(t', t\right) \equiv \beta W_{N(t'),h}(\theta^{(t')}, \theta^{(t)}, \Omega^{(t)})$$

Furthermore, let $\tau < t$ and let

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

for $l, m, \underline{N} \geq 1$. Define

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

Then, for any $\underline{N}$,

$$\begin{aligned}
& \left|V(s,\epsilon,\theta) - V^{(t+\underline{N})}\left(s,\epsilon,\theta,\Omega^{(t+\underline{N})}\right)\right| \\
\leq\ & \left|\sup_{a\in A} A^{(t+\underline{N})}\left(\theta^{(t+\underline{N})}\right)\right| \\
& + \sum_{m=0}^{\underline{N}-1} \widehat{W}\left(t+\underline{N}, t+\underline{N}-m, t+\underline{N}-m-1, 1\right) \sup_{a\in A}\left|A^{(t+\underline{N}-m-1)}\left(\theta^{(t+\underline{N}-m-1)}\right)\right| \\
& + \sum_{n=1}^{N(t)} \sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)}) - V^{(t-n)}(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)},\Omega^{(t)})\right| \widehat{W}(t+\underline{N},t,t-n,1). \quad \text{(A16)}
\end{aligned}$$

Furthermore,

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

Proof of Lemma.
For iteration $t+1$, we get

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

Now, we substitute away $\left|V(s',\epsilon^{(t)},\theta^{(t)}) - V^{(t)}(s',\epsilon^{(t)},\theta^{(t)},\Omega^{(t)})\right|$ by using $A15\prime$) and the fact that $N(t) \geq N(t+1)-1$,

$$\begin{aligned}
& \left|V\left(s,\epsilon,\theta^{(t+1)}\right) - V^{(t+1)}\left(s,\epsilon,\theta^{(t+1)},\Omega^{(t+1)}\right)\right| \\
\leq\ & \left|\sup A^{(t+1)}(\theta^{(t+1)})\right| + \sup_{a\in A}\left|A^{(t)}(\theta^{(t)})\right| \widetilde{W}(t+1,t) \\
& + \sum_{n=1}^{N(t)} \sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)}) - V^{(t-n)}(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)},\Omega^{(t-n)})\right| \\
& \{\widetilde{W}(t+1,t)\widetilde{W}(t,t-n) + \widetilde{W}(t+1,t-n)\}
\end{aligned}$$

Notice that

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

Hence,

$$
\begin{aligned}
&\left|V\left(s,\epsilon,\theta^{(t+1)}\right)-V^{(t+1)}\left(s,\epsilon,\theta^{(t+1)},\Omega^{(t+1)}\right)\right| \\
\leq\ &\left|\sup A^{(t+1)}\left(\theta^{(t+1)}\right)\right|+\sup_{a\in A}\left|A^{(t)}\left(\theta^{(t)}\right)\right|\widehat{W}(t+1,t+1,t,1) \\
&+\sum_{n=1}^{N(t)}\sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})-V^{(t-n)}(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)},\Omega^{(t-n)})\right|\widehat{W}(t+1,t,t-n,1)
\end{aligned}
$$

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

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

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

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

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

Thus,

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

Next, suppose that inequality $A16$ holds for $\underline{N}= M$. Then, using $t+1$ instead of t in inequality $A16$ yields,

$$\begin{aligned}
&\left|V\left(s,\epsilon,\theta^{(t+1+M)}\right)-V^{(t+1+M)}\left(s,\epsilon,\theta^{(t+1+M)},\Omega^{(t+1+M)}\right)\right| \\
\leq\ & \left|\sup A^{(t+M+1)}\left(\theta^{(t+M+1)}\right)\right| \\
&+\sum_{m=0}^{M-1}\widehat{W}\left(t+M+1,t+M+1-m,t+M-m,1\right)\sup_{a\in A}\left|A^{(t+M-m)}\left(\theta^{(t+M-m)}\right)\right| \\
&+\sum_{n=1}^{N(t)}\sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)})-V^{(t-n)}(\widehat{s},\epsilon^{(t-n)},\theta^{(t-n)},\Omega^{(t-n)})\right| \\
&\left[\widehat{W}(t+M+1,t+1,t,1)\widetilde{W}(t,t-n)+\widehat{W}(t+M+1,t+1,t-n,1)\right] \qquad \text{(A18)}
\end{aligned}$$

Now, we claim that,

$$\begin{aligned}
&\widehat{W}\left(t+M,t+1,t,1\right)\widetilde{W}(t,t-n)+\widehat{W}(t+M,t+1,t-n,1) \\
=\ &\widehat{W}\left(t+M,t,t-n,1\right) \qquad \text{(A19)}
\end{aligned}$$

Proof of the Claim:

Let $\Psi_{m,1}(t{+}M,t,\tau)\equiv\{J_m=(t_m,t_{m-1},...,t_1,t_0):t_m=t+M>t_{m-1}>....>t_2\geq t+1,t=t_1,t_0=\tau\}$, $\Psi_{m,2}(t{+}M,t,\tau)\equiv\{J_m=(t_m,t_{m-1},...,t_1,t_0):t_m=t+M>t_{m-1}>....>t_2>t_1\geq t+1,t_0=\tau\}$. Then,

$$\Psi_m(t+M,t,\tau)=\Psi_{m,1}(t+M,t,\tau)\cup\Psi_{m,2}(t+M,t,\tau)$$

and

$$\Psi_{m,1}(t+M,t,\tau)\cap\Psi_{m,2}(t+M,t,\tau)=\varnothing$$

Notice that

$$\Psi_{m,1}(t+M,t,\tau)=\Psi_1\left(t,t,\tau\right)\cup\Psi_{m-1}\left(t+M,t+1,t\right)$$

for $m>1$ and

$$\Psi_{1,1}(t+M,t,\tau)=\{J_1=(t_1,t_0):t_0=\tau,t=t_1=t+M\}=\varnothing$$

and

$$\Psi_{m,2}(t+M,t,\tau)=\Psi_m(t+M,t+1,\tau)$$

if $m\leq M$ and because,

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

$$\Psi_{M+1,2}(t+M,t,\tau)=\varnothing$$

Hence,

$$\begin{aligned}
&\widehat{W}(t+M,t,\tau,1)\\
\equiv\ &\sum_{m=1}^{M+1}\left\{\sum_{\Psi_m(\tau,t,t+M)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\\
=\ &\sum_{m=1}^{M+1}\left\{\sum_{\Psi_{m1}(\tau,t,t+M)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}+\sum_{m=1}^{M+1}\left\{\sum_{\Psi_{m2}(\tau,t,t+M)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\\
=\ &\sum_{m=1}^{M+1}\left\{\sum_{\Psi_{m-1}(t,t+1,t+M)}\prod_{k=1}^{m-1}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\widetilde{W}\left(t,\tau\right)+\sum_{m=1}^{M}\left\{\sum_{\Psi_m(\tau,t+1,t+M)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\\
=\ &\sum_{m=1}^{M}\left\{\sum_{\Psi_m(t,t+1,t+M)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}\widetilde{W}\left(t,\tau\right)+\sum_{m=1}^{M}\left\{\sum_{\Psi_m(\tau,t+1,t+M)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}
\end{aligned}$$

By definition,

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

and

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

Therefore,

$$\widehat{W}(t+M,t,\tau,1)=\widehat{W}\left(t+M,t+1,t,1\right)\widetilde{W}(t,\tau)+\widehat{W}(t+M,t+1,\tau,1)$$

and the claim holds. Similar results hold when we substitute $M+1$ for M, i.e.

$$\widehat{W}(t+M+1,t,\tau,l)=\widehat{W}(t+M+1,t+1,t,l-1)\widetilde{W}(t,\tau)+\widehat{W}(t+M+1,t+1,\tau,l)$$

Substituting this into equation $A18$ yields the first part of the lemma. For the second part, note that, for $l=1$, and for any $M>0$,

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

Suppose for some $l>0$,

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

holds for any $M>0$. Then, for $l'=l+1$, by definition of $\widehat{W}$,

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

Therefore, from induction, the Lemma holds.

Now, 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 $t(2)-N(t(2))=t(1)+1$. Such $t(2)$ exists because from assumption, $N(t)$ is nondecreasing in t and $t-N(t)\rightarrow\infty$. Also, let $\widetilde{N}(2)=N(t(2))$. Similarly, for any $l>2$, let $t(l+1)$ be such that $t(l+1)-N(t(l+1))=t(l)+1$, and let $\widetilde{N}(l+1)=N(t(l+1))$. Furthermore, assume that there exists a finite constant $A>0$ such that $\widetilde{N}(l+1)<AN(l)$. Then,

$$\begin{aligned}&\sum_{m=1}^{\widetilde{N}(l+1)} \left|V\left(s,\epsilon^{(t(l)+m)},\theta^{(t(l)+m)}\right)-V^{(t(l)+m)}\left(s,\epsilon^{(t(l)+m)},\theta^{(t(l)+m)},\Omega^{(t(l)+m)}\right)\right|\widetilde{W}(t(l+1),t(l)+m) \\ \leq\ &\sum_{m=1}^{\widetilde{N}(l+1)} \{\left|\sup_{a\in A} A^{(t(l)+m)}\left(\theta^{(t(l)+m)}\right)\right| \\ &+\sum_{i=0}^{m-1}\widehat{W}(t(l)+m,t(l)+m-i,t(l)+m-i-1,1)\sup_{a\in A}\left|A^{(t(l)+m-i-1)}\right|\}\widetilde{W}(t(l+1),t(l)+m) \\ &+\sum_{m=1}^{\widetilde{N}(l+1)}\sum_{n=1}^{\widetilde{N}(l)}\sup_{\widehat{s}\in S}\left|V(\widehat{s},\epsilon^{(t(l)-n)},\theta^{(t(l)-n)})-V^{(t(l)-n)}(\widehat{s},\epsilon^{(t(l)-n)},\theta^{(t(l)-n)},\Omega^{(t(l)-n)})\right| \\ &\widehat{W}(t(l)+m,t(l),t(l)-n,1)\widetilde{W}(t(l+1),t(l)+m) \qquad \text{(A20)}\end{aligned}$$

Now, from the definition of $\widehat{W}(t(l+1),t(l),t(l)-n,l)$, *(the use of l in two ways!)*

$$\widehat{W}(t(l+1),t(l),t(l)-n,l+1)=\sum_{m=l-1}^{\widetilde{N}(l+1)}\widetilde{W}(t(l+1),t(l)+m)\,\widehat{W}(t(l)+m,t(l),t(l)-n,l)$$

I*n the sum, m starts from $l-1$, not* 1.

Hence,

$$\begin{aligned}
\text{RHS of } A20 \quad \le \quad & \sum_{m=1}^{\widetilde{N}(l+1)} \left| \sup_{a \in A} A^{(t(l)+m)}(\theta^{(t(l)+m)}) \right| \widetilde{W}\left(t(l+1), t(l)+m\right) \\
& + \sum_{i=1}^{\widetilde{N}(l+1)} \left\{ \widehat{W}\left(t(l+1), t(l)+i, t(l)+i-1, 2\right) \sup_{a \in A} \left| A^{(t(l)+i-1)}\left(\theta^{(t(l)+i-1)}\right) \right| \right\} \\
& + \sum_{n=1}^{\widetilde{N}(l)} \sup_{\widehat{s} \in S} \left| V(\widehat{s}, \epsilon^{(t(l)-n)}, \theta^{(t(l)-n)}) - V^{(t(l)-n)}(\widehat{s}, \epsilon^{(t(l)-n)}, \theta^{(t(l)-n)}, \Omega^{(t(l)-n)}) \right| \\
& \widehat{W}(t(l+1), t(l), t(l)-n, 2)
\end{aligned} \tag{A21}$$

Now, let

$$\begin{aligned}
A(l) \quad = \quad & \sum_{m=1}^{\widetilde{N}(l+1)} \left| \sup_{a \in A} A^{(t(l)+m)}(\theta^{(t(l)+m)}) \right| \widetilde{W}\left(t(l+1), t(l+1)-m\right) \\
& + \sum_{i=1}^{\widetilde{N}(l+1)} \left\{ \widehat{W}\left(t(l+1), t(l)+i, t(l)+i-1, 2\right) \sup_{a \in A} \left| A^{(t(l)+i-1)}\left(\theta^{(t(l)+i-1)}\right) \right| \right\}
\end{aligned}$$

Recall that

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

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

$$\Pr\left[\sup_{a \in A} \left| A^{(t)} \right| < \eta_1 \right] > 1 - \eta_2$$

Therefore,

$$E\left[\sup_{a \in A} \left| A^{(t)} \right| \right] \le \eta_1 \Pr\left[\sup_{a \in A} \left| A^{(t)} \right| < \eta_1 \right] + \overline{A} \Pr\left[\sup_{a \in A} \left| A^{(t)} \right| \ge \eta_1 \right] \le \eta_1 (1 - \eta_2) + \overline{A} \eta_2 \tag{A22}$$

Next, we show that

$$A(l) \overset{P}{\to} 0 \quad \text{as} \quad l \to \infty$$

Proof: Let

$$A(l) = B_1 + B_2$$

where

$$B_1 \equiv \sum_{m=1}^{\widetilde{N}(l+1)} \left| \sup_{a \in A} A^{(t(l)+m)} \left(\theta^{(t(l)+m)}\right) \right| \widetilde{W}\left(t(l+1), t(l)+m\right)$$

and

$$B_2 \equiv \sum_{i=1}^{\widetilde{N}(l+1)} \widehat{W}\left(t(l+1), t(l)+i, t(l)+i-1, 2\right) \sup_{a \in A} \left| A^{(t(l)+i-1)} \left(\theta^{(t(l)+i-1)}\right) \right|$$

Claim 1:

$$B_1 \xrightarrow{P} 0$$

Proof: For any $t' > t > 0$, denote

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

First, we focus on the numerator part divided by $\widetilde{N}(l+1)$. That is, consider,

$$\frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{K}\left(t(l+1), t(l)+m\right) \left| \sup_{a \in A} A^{(t(l)+m)} \left(\theta^{(t(l)+m)}\right) \right|$$

Notice that,

$$\begin{aligned} & \frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{K}\left(t(l+1), t(l)+m\right) \left| \sup_{a \in A} A^{(t(l)+m)} \left(\theta^{(t(l)+m)}\right) \right| \\ \leq \; & \frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{K}\left(t(l+1), t(l)+m\right) \left| \sup_{a \in A, \theta \in \Theta} A^{(t(l)+m)}(\theta) \right| \end{aligned}$$

Now, for any $0 < m < \widetilde{N}(l+1)$, since $A^{(t(l)+m)}(\theta)$ only depends on $\Omega^{(t(l)+m)}$,

$$\begin{aligned} & E\left\{ \widetilde{K}(t(l+1), t(l)+m) \sup_{a \in A, \theta \in \Theta} \left| A^{(t(l)+m)}(\theta) \right| \right\} \\ = \; & E\left\{ K(\theta^{(t(l+1))}, \theta^{(t(l)+m)}) E_{\Omega^{(t(l)+m)}} \left[\sup_{a \in A, \theta \in \Theta} \left| A^{(t(l)+m)}(\theta) \right| \right] \right\} \end{aligned} \qquad \text{(A23)}$$

From $A22$, $E_{\Omega^{(t(l)+m)}}\left[\sup_{a \in A, \theta \in \Theta} \left|A^{(t(l)+m)}(\theta)\right|\right] \leq \left[\eta_1(1-\eta_2) + \eta_2 \overline{A}\right]$. Hence,

$$RHS \text{ of } A23 \leq E\left\{ K(\theta^{(t(l+1))}, \theta^{(t(l)+m)}) \left[\eta_1(1-\eta_2) + \eta_2 \overline{A}\right] \right\}$$

Therefore,

$$
\begin{aligned}
& \left\{ \frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{K}(t(l+1), t(l)+m) \left| \sup_{a \in A} A^{(t(l)+m)} \left(\theta^{(t(l)+m)} \right) \right| \right\} \\
\leq \; & E \left\{ \frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{K}(t(l+1), t(l)+m) \left[\eta_1 (1-\eta_2) + \eta_2 \overline{A} \right] \right\} \\
\leq \; & \left[\eta_1 (1-\eta_2) + \eta_2 \overline{A} \right] E \left\{ \frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} K_h(\theta^{(t(l+1))} - \theta^{(t(l)+m)}) \right\} \qquad \text{(A24)}
\end{aligned}
$$

Notice that from Lemma 2, for any $0 < m < \widetilde{N}(l+1)$,

$$
\begin{aligned}
& f\left(\theta^{(t(l+1))}, \theta^{(t(l+1)-1)}\right) f\left(\theta^{(t(l+1)-1)}, \theta^{(t(l+1)-2)}\right) \ldots f\left(\theta^{(t(l)+m+1)}, \theta^{(t(l)+m)}\right) f\left(\theta^{(t(l)+m)}, \theta^{(t(l)+m-1)}\right) \ldots f\left(\theta^{(2)} \right. \\
\leq \; & \epsilon_1^2 g(\theta^{(t(l+1))}) f\left(\theta^{(t(l+1)-1)}, \theta^{(t(l+1)-2)}\right) \ldots f\left(\theta^{(t(l)+m+1)}, \theta^{(t(l)+m)}\right) g(\theta^{(t(l)+m)}) \ldots f\left(\theta^{(2)}, \theta^{(1)}\right)
\end{aligned}
$$

Hence,

$$
\begin{aligned}
E\left[K_h(\theta^{(t(l+1))} - \theta^{(t(l)+m)})\right] &= E\left[K_h\left(\theta^{(t(l+1))}(f) - \theta^{(t(l)+m)}(f)\right)\right] \\
&\leq \epsilon_1^2 E_{\theta^{(t(l+1))}, \theta^{(t(l)+m)}} \left[K_h\left(\theta^{(t(l+1))}(g) - \theta^{(t(l)+m)}(g)\right)\right] \\
&\leq \epsilon_1^2 \sup_{\theta \in \Theta} E_{\theta^{(t(l)+m)}} \left[K_h\left(\theta - \theta^{(t(l)+m)}(g)\right)\right]
\end{aligned}
$$

and

$$
\begin{aligned}
RHS \text{ of } A24 &\leq \left[\eta_1 (1-\eta_2) + \eta_2 \overline{A} \right] \frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} E\left\{K_h(\theta^{(t(l+1))} - \theta^{(t(l)+m)})\right\} \\
&\leq \left[\eta_1 (1-\eta_2) + \eta_2 \overline{A} \right] \epsilon_1^2 \frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \sup_{\theta \in \Theta} E\left\{K_h(\theta, \theta^{(t(l)+m)}(g))\right\} \\
&= \left[\eta_1 (1-\eta_2) + \eta_2 \overline{A} \right] \epsilon_1^2 \sup_{\theta' \in \Theta} E\left\{K_h(\theta', \theta(g))\right\} \qquad \text{(A25)}
\end{aligned}
$$

Now, for any $\delta > 0$,

$$
\begin{aligned}
& \left[\sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{W}(t(l+1), t(l)+m) \sup_{a \in A} \left| A^{(t(l)+m)} \right| \leq \delta \right] \\
\supseteq \; & \left[\frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{K}(t(l+1), t(l)+m) \sup_{a \in A} \left| A^{(t(l)+m)} \right| \leq \frac{\delta}{\frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)} \right] \cap \\
& \left[\frac{1}{\widetilde{N}(l+1)} \sum_{k=1}^{N(l+1)} K_h(\theta^{t(l+1)} - \theta^{(t(l+1)-k)}(f^{(t(l+1)-k)})) \geq \frac{1}{2}\varepsilon_0 \inf_\theta g(\theta) \right]
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\left[\sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{W}(t(l+1), t(l)+m) \sup_{a\in A}\left|A^{(t(l)+m)}\right| > \delta\right] \\
\subseteq\; & \left[\frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{K}(t(l+1), t(l)+m) \sup_{a\in A}\left|A^{(t(l)+m)}\right| > \frac{\delta}{\frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)}\right] \cap \\
& \left[\frac{1}{\widetilde{N}(l+1)} \sum_{k=1}^{N(l+1)} K_h(\theta^{t(l+1)} - \theta^{(t(l+1)-k)}(f^{(t(l+1)-k)})) < \frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)\right]
\end{aligned}
$$

and

$$
\begin{aligned}
&\Pr\left[\sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{W}(t(l+1), t(l)+m) \sup_{a\in A}\left|A^{(t(l)+m)}\right| > \delta\right] \\
\leq\; & \Pr\left[\frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{K}(t(l+1), t(l)+m) \sup_{a\in A}\left|A^{(t(l)+m)}\right| > \frac{\delta}{\frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)}\right] \cup \\
& \left[\frac{1}{\widetilde{N}(l+1)} \sum_{k=1}^{N(l+1)} K_h(\theta^{t(l+1)} - \theta^{(t(l+1)-k)}(f^{(t(l+1)-k)})) < \frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)\right] \\
\leq\; & \Pr\left[\frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{K}(t(l+1), t(l)+m) \sup_{a\in A}\left|A^{(t(l)+m)}\right| > \frac{\delta}{\frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)}\right] \\
& + \Pr\left[\frac{1}{\widetilde{N}(l+1)} \sum_{k=1}^{N(l+1)} K_h(\theta^{t(l+1)} - \theta^{(t(l+1)-k)}(f^{(t(l+1)-k)})) < \frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)\right] \qquad \text{(A26)}
\end{aligned}
$$

Now, from Chebychev's Inequality and $A25$,

$$
\begin{aligned}
&\Pr\left[\frac{1}{\widetilde{N}(l+1)} \sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{K}(t(l+1), t(l)+m) \sup_{a\in A}\left|A^{(t(l)+m)}\right| > \frac{\delta}{\frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)}\right] \\
\leq\; & \frac{\left[\eta_1(1-\eta_2) + \eta_2 \overline{A}\right] \epsilon_1^2 \sup_{\theta'\in\Theta} E\left\{K_h(\theta', \theta(g))\right\}}{\delta / \left[\frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)\right]} \qquad \text{(A27)}
\end{aligned}
$$

Furthermore,

$$
\sum_{k=1}^{\widetilde{N}(l+1)} K_h(\theta^{t(l+1)} - \theta^{(t(l+1)-k)}) \leq \inf_{\theta\in\Theta} \sum_{k=1}^{\widetilde{N}(l+1)} K_h(\theta - \theta^{(t(l+1)-k)}(f^{(t(l+1)-k)}))
$$

and from $A6$, we know that for any $\eta_3 > 0$, there exists L such that for any $l > L$,

$$
\Pr\left[\frac{1}{\widetilde{N}(l+1)} \sum_{k=1}^{\widetilde{N}(l+1)} K_h(\theta^{t(l+1)} - \theta^{(t(l+1)-k)}) < \frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)\right] < \eta_3 \qquad \text{(A28)}
$$

Now, given arbitrarily small $\eta > 0$, choose η_1, η_2, η_3 such that

$$\frac{\left[\eta_1(1-\eta_2)+\eta_2\overline{A}\right]\epsilon_1^2 \sup_{\theta'\in\Theta} E\left\{K_h(\theta',\theta(g))\right\}}{\delta/\left[\frac{1}{2}\varepsilon_0 \inf_\theta g(\theta)\right]} + \eta_3 < \eta \tag{A29}$$

Let L_1 such that $L_1 > L$, $t(L_1) > T$. Then, from $A26$ to $A29$,for any $l > L_1$,

$$\Pr\left[\sum_{m=1}^{\widetilde{N}(l+1)} \widetilde{W}(t(l+1),t(l)+m)\sup_{a\in A}\left|A^{(t(l)+m)}\right| > \delta\right] < \eta$$

Hence, Claim 1 holds.

Claim 2:

$$B_2 \equiv \sum_{i=1}^{\widetilde{N}(l+1)} \widehat{W}\left(t(l+1),t(l)+i,t(l)+i-1,2\right)\sup_{a\in A}\left|A^{(t(l)+i-1)}\right| \xrightarrow{P} 0$$

Define $W^*\left(t(l+1)-m,t(l),t,k\right)$ recursively to be as follows.

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

Similarly,

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

Now, for given l,

$$\widehat{W}(t(l+1),t(l),\tau,l_1) \equiv \sum_{m=l_1}^{\widetilde{N}(l+1)}\left\{\sum_{\Psi_m(t(l+1),t,\tau)}\prod_{k=1}^{m}\widetilde{W}\left(t_k,t_{k-1}\right)\right\}$$

$$= \sum_{k=l_1}^{\widetilde{N}(l+1)} W^*(t(l+1),t(l),\tau,k) \tag{A30}$$

Hence,

$$
\begin{aligned}
& \sum_{i=1}^{\widetilde{N}(l+1)} \left\{ \widehat{W}\left(t(l+1), t(l)+i, t(l)+i-1, l_1\right) \sup_{a\in A} \left| A^{(t(l)+i-1)} \right| \right\} \\
= & \sum_{i=1}^{\widetilde{N}(l+1)} \left\{ \sum_{k=l_1}^{\widetilde{N}(l+1)} W^*(t(l+1), t(l)+i, t(l)+i-1, k) \sup_{a\in A} \left| A^{(t(l)+i-1)} \right| \right\} \\
= & \sum_{k=l_1}^{\widetilde{N}(l+1)} \left\{ \sum_{i=1}^{\widetilde{N}(l+1)} W^*(t(l+1), t(l)+i, t(l)+i-1, k) \sup_{a\in A} \left| A^{(t(l)+i-1)} \right| \right\}
\end{aligned}
$$

For any $j > 0$, and for any $j \geq k \geq 1$,

$$
\frac{1}{\widetilde{N}(l+1)^k} \sum_{j_1,\ldots,j_{k+1}} I(t(l)+i \leq j_1 < \ldots < j_{k+1} = t(l)+i+j) = \frac{1}{\widetilde{N}(l+1)^k} \left(\frac{j!}{k!(j-k)!} \right) \tag{A31}
$$

$$
\leq \frac{\left[j/\widetilde{N}(l+1) \right]^k}{k!} \leq \frac{1}{k!} \tag{2}
$$

By induction, we can derive the following two equations.

$$
\begin{aligned}
& E\left\{ \sum_{i=1}^{\widetilde{N}(l+1)} K^*(t(l+1), t(l)+i, t(l)+i-1, k) \right\} \\
\leq & A^k \epsilon_1^{k+1} \sup_{\theta\in\Theta} E\left[\beta K_h\left(\theta - \theta(g)\right)\right]^k \frac{1}{k!}
\end{aligned} \tag{A32}
$$

Also,

$$
\begin{aligned}
& E\left\{ \sum_{i=1}^{\widetilde{N}(l+1)} K^*(t(l+1), t(l)+i, t(l)+i-1, k) | \theta^{(t(l+1))} \right\} \\
\leq & A^k \epsilon_1^{k+1} \sup_{\theta\in\Theta} E\left[\beta K_h\left(\theta - \theta(g)\right)\right]^k \frac{1}{k!}
\end{aligned} \tag{A33}
$$

As in Claim 1, we can show that, for $k = 1$,

$$
\begin{aligned}
& E\left\{ \sum_{i=1}^{\widetilde{N}(l+1)} K^*(t(l+1), t(l)+i, t(l)+i-1, 1) \right\} \\
\leq & E\left\{ \frac{\widetilde{N}(l+1)}{\widetilde{N}(l)} \frac{1}{\widetilde{N}(l+1)} \sum_{i=1}^{\widetilde{N}(l+1)} \widetilde{K}(t(l+1), t(l)+i-1) \right\} \\
\leq & A\epsilon_1^2 E\left[\beta K_h\left(\theta'(g) - \theta(g)\right)\right] \leq A\epsilon_1^2 \sup_{\theta'\in\Theta} E_\theta\left[\beta K_h\left(\theta' - \theta(g)\right)\right]
\end{aligned}
$$

Also,

$$
\begin{aligned}
& E\left\{\sum_{i=1}^{\widetilde{N}(l+1)} K^*(t(l+1), t(l)+i, t(l)+i-1, 1)|\theta^{(t(l+1))}\right\} \\
= \;& E\left\{\frac{\widetilde{N}(l+1)}{\widetilde{N}(l)} \frac{1}{\widetilde{N}(l+1)} \sum_{i=1}^{\widetilde{N}(l+1)} \widetilde{K}(t(l+1), t(l)+i-1)|\theta^{(t(l+1))}\right\} \\
\leq \;& A\epsilon_1^2 E\left[\beta K_h\left(\theta^{(t(l+1))} - \theta(g)\right)|\theta^{(t(l+1))}\right] \leq A\epsilon_1^2 \sup_{\theta' \in \Theta} E_\theta\left[\beta K_h\left(\theta' - \theta(g)\right)\right]
\end{aligned}
$$

Hence, $A32$ and $A33$ hold for $k=1$. Now, suppose that $A32$ and $A33$ hold for $k=M$. Then, for $k=M+1$,

$$
\begin{aligned}
& E\left\{\sum_{i=1}^{\widetilde{N}(l+1)} K^*(t(l+1), t(l)+i, t(l)+i-1, M+1)\right\} \\
\leq \;& E\left\{A\frac{1}{\widetilde{N}(l+1)} \sum_{i=1}^{\widetilde{N}(l+1)} \sum_{j=M}^{\widetilde{N}(l+1)-i} \widetilde{K}(t(l+1), t(l)+i+j) K^*\left(t(l)+i+j, t(l)+i, t(l)+i-1, M\right)\right\} \\
\leq \;& A\frac{1}{\widetilde{N}(l+1)} \sum_{i=1}^{\widetilde{N}(l+1)} \sum_{j=M}^{\widetilde{N}(l+1)-i} E\left\{\widetilde{K}(t(l+1), t(l)+i+j) E\left[K^*\left(t(l)+i+j, t(l)+i, t(l)+i-1, M\right)|\theta^{(t(l)+i+j)}\right]\right\} \\
\leq \;& \frac{A}{\widetilde{N}(l+1)}\epsilon_1 E\left[\beta K_h\left(\theta''(g) - \theta'(g)\right) A^M \epsilon_1^{M+1} \sup_{\theta' \in \Theta} E\left[\beta K_h\left(\theta' - \theta(g)\right)\right]^M\right] \\
& \frac{1}{\widetilde{N}(l+1)^{M+1}} \sum_{j_1, \ldots, j_{M+2}} I(t(l)+i \leq j_1 < j_2 < \ldots < j_{M+2} = t(l+1)) \\
\leq \;& \frac{A^{M+1}}{\widetilde{N}(l+1)} \epsilon_1^{M+2} \sup_{\theta' \in \Theta} E\left[\beta K_h\left(\theta' - \theta(g)\right)\right]^{M+1} \frac{1}{(M+1)!}.
\end{aligned}
$$

Therefore, A32 holds for $k=M+1$. The proof for $A33$ for $k=M+1$ is similar to that for $A32$.

Also for any $\eta_3 > 0$, there exists L such that for any $l > L$, $t = t(l)$ and for $t = t(l) + \widetilde{N}(l)/2$,

$$
\Pr\left[\frac{1}{\widetilde{N}(l)/2} \sum_{n=1}^{\widetilde{N}(l)/2} K_h(\theta - \theta^{(t-k)}(f^{(t-k)})) \geq \frac{1}{2}\varepsilon_0 g(\theta)\right] > 1 - \eta_3
$$

Now, notice that for any $t(l) \leq t \leq t(l+1)$, either $\left[t(l) - \widetilde{N}(l)/2, t(l)\right] \subset [t - N(t), t]$ or $\left[t(l)+1, t(l) + \widetilde{N}(l)/2\right] \subset [t - N(t), t]$ or both. Hence, for any t such that $t(l) \leq t \leq t(l+1)$

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

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

$$\Pr\left[\frac{1}{\widetilde{N}(l+1)}\sum_{n=1}^{N(t)} K_h(\theta-\theta^{(t-k)}(f^{(t-k)})) \geq \frac{1}{4A}\varepsilon_0 g\left(\theta\right)\right] > 1-\eta_3 \tag{A35}$$

Notice that

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

where

$$D\left(t\right) \equiv \sum_{i=1}^{N(t_j)}\widetilde{K}\left(t,t-i\right)$$

Now, we get

$$\begin{aligned}
&\Pr\left[\sum_{i=1}^{\widetilde{N}(l+1)} W^*(t(l+1),t(l)+i,t(l)+i-1,k)\sup_{a\in A}\left|A^{(t(l)+i-1)}\right| > \delta^k\right] \\
\leq\ &\Pr\left[\sum_{m=1}^{\widetilde{N}(l+1)} K^*(t(l+1),t(l)+i,t(l)+i-1,k)\sup_{a\in A}\left|A^{(t(l)+m)}\right| > \left[\frac{\delta}{4A}\varepsilon_0\inf_{\theta} g\left(\theta\right)\right]^k\right] \\
&+\Pr\left[\inf_{t(l)\leq t\leq t(l+1)}\left[\frac{1}{\widetilde{N}(l+1)}\sum_{i=1}^{N(t)}\widetilde{K}(t,t-i)\right] < \frac{1}{4A}\varepsilon_0\inf_{\theta} g\left(\theta\right)\right]
\end{aligned}$$

From Chebychev Inequality, we get

$$\begin{aligned}
&\Pr\left[\sum_{m=1}^{\widetilde{N}(l+1)} K^*(t(l+1),t(l)+i,t(l)+i-1,k)\sup_{a\in A}\left|A^{(t(l)+m)}\right| > \left[\frac{\delta}{4A}\varepsilon_0\inf_{\theta} g\left(\theta\right)\right]^k\right] \\
\leq\ &\frac{\left[\eta_1\left(1-\eta_2\right)+\eta_2\overline{A}\right]A^k\epsilon_1^k\sup_{\theta\in\Theta}E\left[\beta K_h\left(\theta-\theta(g)\right)\right]^k\frac{1}{k!}}{\left[\delta\frac{1}{4A}\varepsilon_0\inf_{\theta} g\left(\theta\right)\right]^k}
\end{aligned}$$

Now, from $A35$, we get

$$\Pr\left[\inf_{t(l)\leq t\leq t(l+1)}\left[\frac{1}{\widetilde{N}(l+1)}\sum_{i=1}^{N(t)}\widetilde{K}(t,t-i)\right]<\frac{1}{4A}\varepsilon_0\inf_{\theta}g(\theta)\right]\leq\eta_3 \tag{3}$$

Together, we establish that

$$\begin{aligned}
&\Pr\left[\sum_{k=l_1}^{\widetilde{N}(l+1)}\sum_{i=1}^{\widetilde{N}(l+1)}W^*(t(l+1),t(l)+i,t(l)+i-1,k)\sup_{a\in A}\left|A^{(t(l)+i-1)}\right|>\frac{\delta-\delta^{\widetilde{N}(l+1)}}{1-\delta}\right]\\
\leq\ &\Pr\left[\bigcup_{k=l_1}^{\widetilde{N}(l+1)}\left\{\sum_{i=1}^{\widetilde{N}(l+1)}W^*(t(l+1),t(l)+i,t(l)+i-1,k)\sup_{a\in A}\left|A^{(t(l)+i-1)}\right|\geq\delta^k\right\}\right]\\
\leq\ &\sum_{k=l_1}^{\widetilde{N}(l+1)}\Pr\left[\sum_{i=1}^{\widetilde{N}(l+1)}W^*(t(l+1),t(l)+i,t(l)+i-1,k)\sup_{a\in A}\left|A^{(t(l)+i-1)}\right|\geq\delta^k\right]\\
\leq\ &\sum_{k=l_1}^{\widetilde{N}(l+1)}\Pr\left[\sum_{m=1}^{\widetilde{N}(l+1)}K^*(t(l+1),t(l)+i,t(l)+i-1,k)\sup_{a\in A}\left|A^{(t(l)+m)}\right|\geq\left[\frac{\delta}{4A}\varepsilon_0\inf_{\theta}g(\theta)\right]^k\right]\\
&+\Pr\left[\inf_{t(l)\leq t\leq t(l+1)}\left[\frac{1}{\widetilde{N}(l+1)}\sum_{i=1}^{N(t)}\widetilde{K}(t,t-i)\right]<\frac{1}{4A}\varepsilon_0\inf_{\theta}g(\theta)\right]\\
\leq\ &\left[\eta_1(1-\eta_2)+\eta_2\overline{A}\right]e^{\lambda}\sum_{k=l_1}^{\widetilde{N}(l+1)}\left[e^{-\lambda}\frac{\lambda^k}{k!}\right]+\eta_3
\end{aligned}$$

where

$$\lambda=\frac{4A^2\epsilon_1\sup_{\theta\in\Theta}E\left[K_h(\theta-\theta(g))\right]}{\delta\varepsilon_0\inf_{\theta}g(\theta)}>0 \tag{4}$$

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

$$\sum_{i=l_1}^{\widetilde{N}(l+1)}e^{-\lambda}\frac{\lambda^k}{k!}\leq 1$$

Together, we have shown that

$$\begin{aligned}
&\Pr\left[\sum_{k=l_1}^{\widetilde{N}(l+1)}\sum_{i=1}^{\widetilde{N}(l+1)}W^*(t(l+1),t(l)+i,t(l)+i-1,k)\sup_{a\in A}\left|A^{(t(l)+i-1)}\right|>\frac{\delta-\delta^{\widetilde{N}(l+1)}}{1-\delta}\right]\\
\leq\ &\left[\eta_1(1-\eta_2)+\eta_2\overline{A}\right]\exp(\lambda)+\eta_3
\end{aligned}$$

RHS of the equation can be made arbitrarily small by choosing l, large enough. Hence $B_2\xrightarrow{P}0$.

Together, we have shown that

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

Now, iterating it once again from $A21$, we obtain,

$$\begin{aligned}
& \sum_{m=1}^{\widetilde{N}(l)} \left| V\left(s, \epsilon^{(t(l-1)+m)}, \theta^{(t(l-1)+m)}\right) - V^{(t(l-1)+m)}\left(s, \epsilon^{(t(l-1)+m)}, \theta^{(t(l-1)+m)}, \Omega^{(t(l-1)+m)}\right)\right| \\
& \widehat{W}\left(t(l+1), t(l), t(l-1)+m, 2\right) \\
\leq & \sum_{m=1}^{\widetilde{N}(l)} \{ \left|\sup_{a\in A} A^{(t(l-1)+m)}\right| \\
& + \sum_{i=0}^{m-1} \widehat{W}\left(t(l-1)+m, t(l-1)+m-i, t(l-1)+m-i-1, 1\right) \sup_{a\in A} \left|A^{(t(l-1)+m-i-1)}\right| \} \\
& \widehat{W}\left(t(l+1), t(l), t(l-1)+m, 2\right) \\
& + \sum_{m=1}^{\widetilde{N}(l)} \sum_{n=1}^{\widetilde{N}(l-1)} \sup_{\widehat{s}\in S} \left| V(\widehat{s}, \epsilon^{(t(l-1)-n)}, \theta^{(t(l-1)-n)}) - V^{(t(l-1)-n)}(\widehat{s}, \epsilon^{(t(l-1)-n)}, \theta^{(t(l-1)-n)}, \Omega^{(t(l-1)-n)})\right| \\
& \widehat{W}(t(l-1)+m, t(l-1), t(l-1)-n, 1) \widehat{W}\left(t(l+1), t(l), t(l-1)+m, 2\right) \qquad \text{(A36)}
\end{aligned}$$

Notice that,

$$\begin{aligned}
& \sum_{m=1}^{\widetilde{N}(l)} \widehat{W}\left(t(l+1), t(l), t(l-1)+m, 2\right) \widehat{W}(t(l-1)+m, t(l-1), t(l-1)-n, 1) \\
= \; & \widehat{W}\left(t(l+1), t(l-1), t(l-1)-n, 3\right)
\end{aligned}$$

Hence,

$$\begin{aligned}
& \text{RHS of } A36 \\
= & \sum_{m=1}^{\widetilde{N}(l)} \widehat{W}\left(t(l+1), t(l), t(l-1)+m, 2\right) \left|\sup_{a\in A} A^{(t(l-1)+m)}\right| \\
& + \sum_{i=0}^{\widetilde{N}(l)} \widehat{W}\left(t(l+1), t(l-1)+i, t(l-1)+i-1, 3\right) \sup_{a\in A} \left|A^{(t(l-1)+i-1)}\right| \\
& \sum_{n=1}^{\widetilde{N}(l-1)} \sup_{\widehat{s}\in S} \left| V(\widehat{s}, \epsilon^{(t(l-1)-n)}, \theta^{(t(l-1)-n)}) - V^{(t(l-1)-n)}(\widehat{s}, \epsilon^{(t(l-1)-n)}, \theta^{(t(l-1)-n)}, \Omega^{(t(l-1)-n)})\right| \\
& \widehat{W}\left(t(l+1), t(l-1), t(l-1)-n, 3\right)
\end{aligned}$$

Denote

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

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

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

$$\overline{W}\left(l+1,k\right) \equiv \left[\widehat{W}(t(l+1),t(l+2-k),t(l+2-k)-m,k)\right]_{m=1}^{\widetilde{N}(l+2-k)}$$

Then, $\overline{W}\left(l+1,k\right)'\iota \leq \beta^k$ and from $A20$,

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

The sum for A should start from i = 2. V should have l+2-k as an argument. The first term on the RHS, $\sum_{i=0}^{k} A\left(l-k+i\right)$ converges to 0 in probability as $l \to \infty$ given $k > 0$, and since $\Delta V\left(l+2-k\right)$ is bounded and $\overline{W}\left(l+1,k\right)'\iota \leq \beta^k$, the second term can be made arbitrarily small by chosing large enough k. Therefore, $\Delta V\left(l\right)' W\left(l,1\right)$ converges to zero in probability as $l \to \infty$.

Now, from (4), we know that for $t \geq t(l)$

$$\begin{aligned}&\left|\mathcal{V}\left(s,a,\epsilon^{(t)},\theta\right)-\mathcal{V}^{(t)}\left(s,a,\epsilon^{(t)},\theta,\Omega^{(t)}\right)\right|\\ \leq\ &\left|A^{(t)}\right|+\Delta V(l)'W(l,1)\end{aligned} \tag{A37}$$

Since RHS converges to 0 in probability as $l \to \infty$,

$$\left|\mathcal{V}\left(s,a,\epsilon^{(t)},\theta\right)-\mathcal{V}^{(t)}\left(s,a,\epsilon^{(t)},\theta,\Omega^{(t)}\right)\right| \xrightarrow{P} 0$$

as $t \to \infty$.

Proof of Theorem 2

Suppose there is a Markov chain with the transition function $f^{(t)}\left(.,.\right)$ which converges to $f\left(.,.\right)$ in probability uniformly. Also, suppose that that there is a density $g(.)$ and a constant $\varepsilon > 0$ such that for any $\theta \in \Theta$,

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

Also, define $\nu^{(t)}$ as follows.

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

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, assume the following coupling scheme. Let $X^{(t)}$ be a random variable, and $Y^{(t)}$ is the Markov process tht follows the transition probability $f(x,.)$. Suppose $X^{(t)} \neq Y^{(t)}$.

With probability $\varepsilon > 0$, let

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

with probability $1-\varepsilon$,

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

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

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

with probaiblity $1 - v_k$,

$$\begin{aligned} X^{(t+1)} &\sim \frac{1}{1-v^{(t)}}\left[f^{(t)}\left(X^{(t)},.\right) - \varepsilon g(.)\right] \\ Y^{(t+1)} &\sim \frac{1}{1-v^{(t)}}\left[f^{(t)}\left(Y^{(t)},.\right) - \varepsilon g(.)\right] \end{aligned}$$

As $f^{(t)}(x,.) \xrightarrow{P} f(x,.)$ uniformly over the compact parameter set Θ, $v^{(t)}$ converges to v in probability. Let $w^{(t)} = 1 - v^{(t)}$. Then, $w^{(t)} \xrightarrow{P} 0$. Let state 1 be the case when $X^{(t)} = Y^{(t)}$, and let state 2 be the case when $X^{(t)} \neq Y^{(t)}$. Then, the process $\left(X^{(t)}, Y^{(t)}\right)$ follows the Markov process with the below transition matrix.

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

Denote the unconditional probability of state 1 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,

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

Prove that $\pi^{(t)} \xrightarrow{P} 1$.

Define W_{tm} to be

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

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,

$$\begin{aligned} \Pr\left\{|W_{km} - 0| < \frac{\overline{\delta}_1}{5}\right\} &\geq \Pr\left\{\bigcap_{j=t-m}^{t} \left|w^{(j)} - 0\right| < \delta_1\right\} \\ &= 1 - \Pr\left\{\bigcup_{j=t-m}^{t} \left|w^{(j)} - 0\right| \geq \delta_1\right\} \\ &\geq 1 - \sum_{j=t-m}^{t} \Pr\left\{\left|w^{(j)} - 0\right| \geq \delta_1\right\} \geq 1 - \overline{\delta}_2 \end{aligned} \quad (12)$$

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

$$\begin{aligned} \Pr\left[\left|\pi^{(t+1)} - 1\right| < \overline{\delta}_1\right] &= \Pr\left[\left|\pi^{(t-m)}(1-\varepsilon)^m - \varepsilon^{m+1} + W_{km}\right| < \overline{\delta}_1\right] \\ &\geq \Pr\left[\left|\pi^{(t-m)}(1-\varepsilon)^m - \varepsilon^{m+1}\right| < \frac{2\overline{\delta}_1}{5}, |W_{km}| < \frac{\overline{\delta}_1}{5}\right] \\ &= \Pr\left[|W_{km}| < \frac{\overline{\delta}_1}{5}\right] \end{aligned} \quad (13)$$

Last equality holds because

$$0 \leq \pi^{(t-m)} \leq 1$$

$$\left|\pi^{(t-m)}(1-\varepsilon)^m - \varepsilon^{m+1}\right| \leq \left|(1-\varepsilon)^m - \varepsilon^{m+1}\right| \leq \left|(1-\varepsilon)^m\right| + \left|\varepsilon^{m+1}\right| < \frac{2\overline{\delta}_1}{5}$$

From (12) and (13), 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)

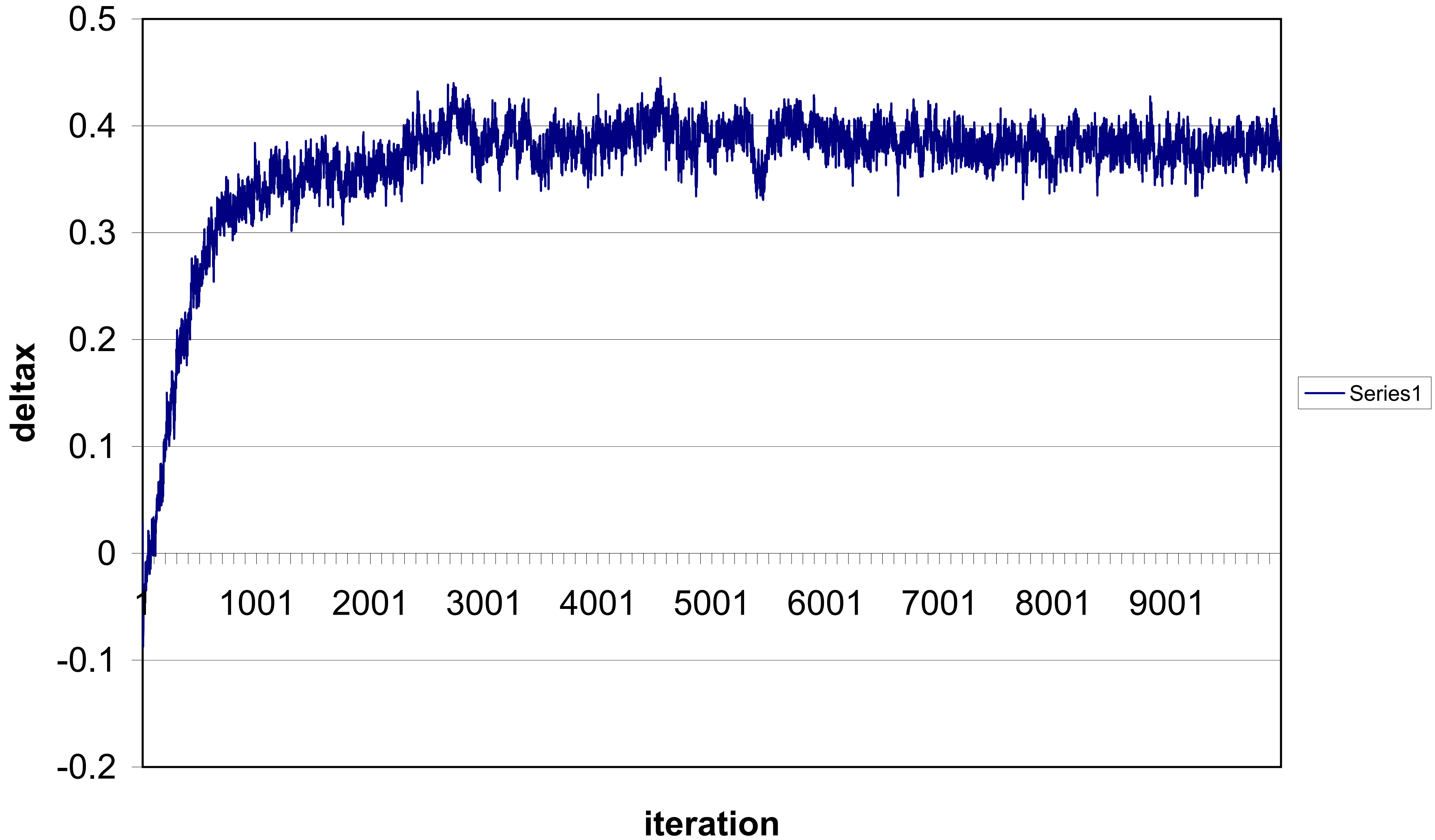

Figure 2: Gibbs Sampler Output of Entry Cost (True Value:0.4)

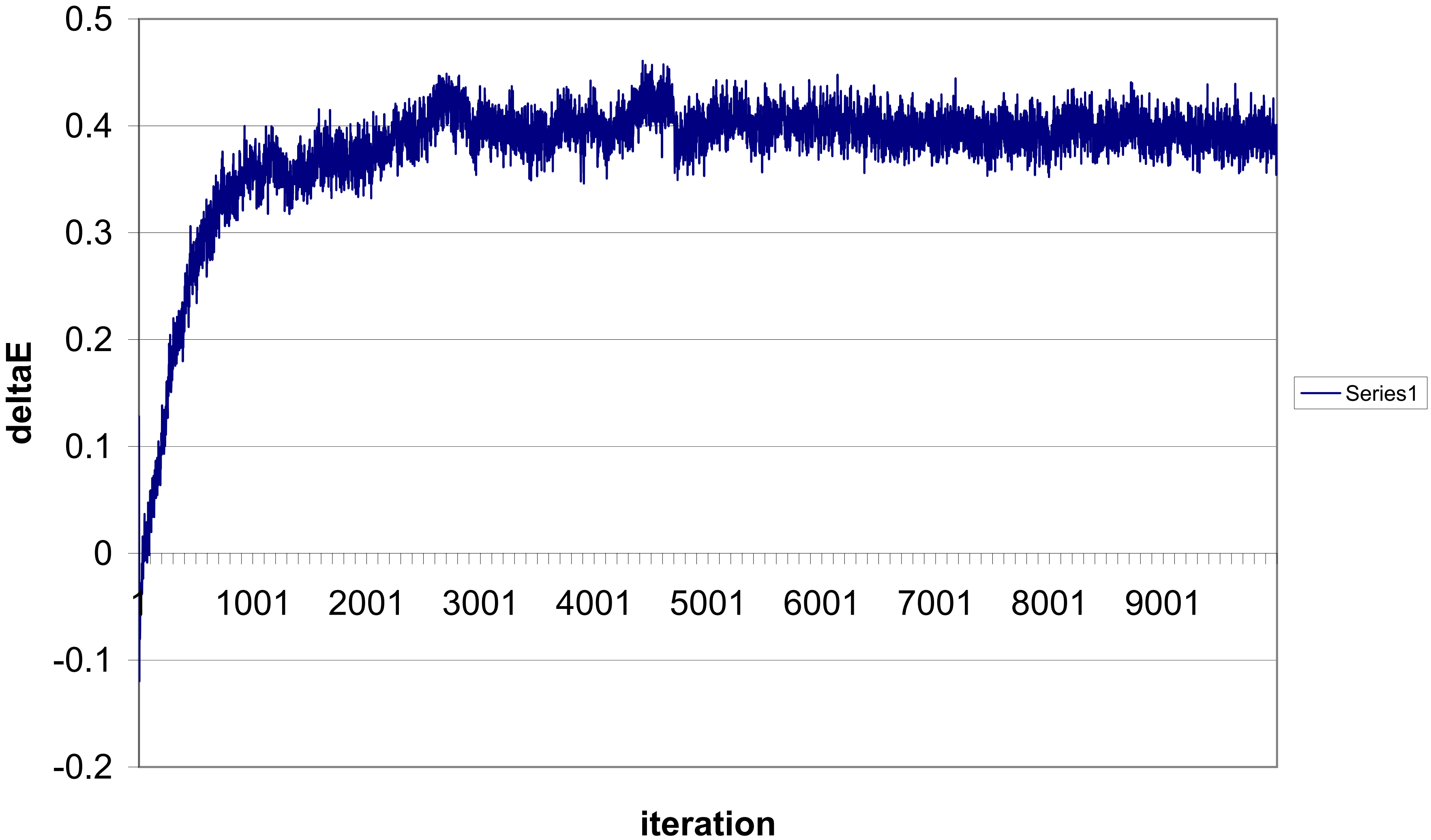

Figure 3: Gibbs Sampler Output of Capital Coefficient
(True Value: 0.2)

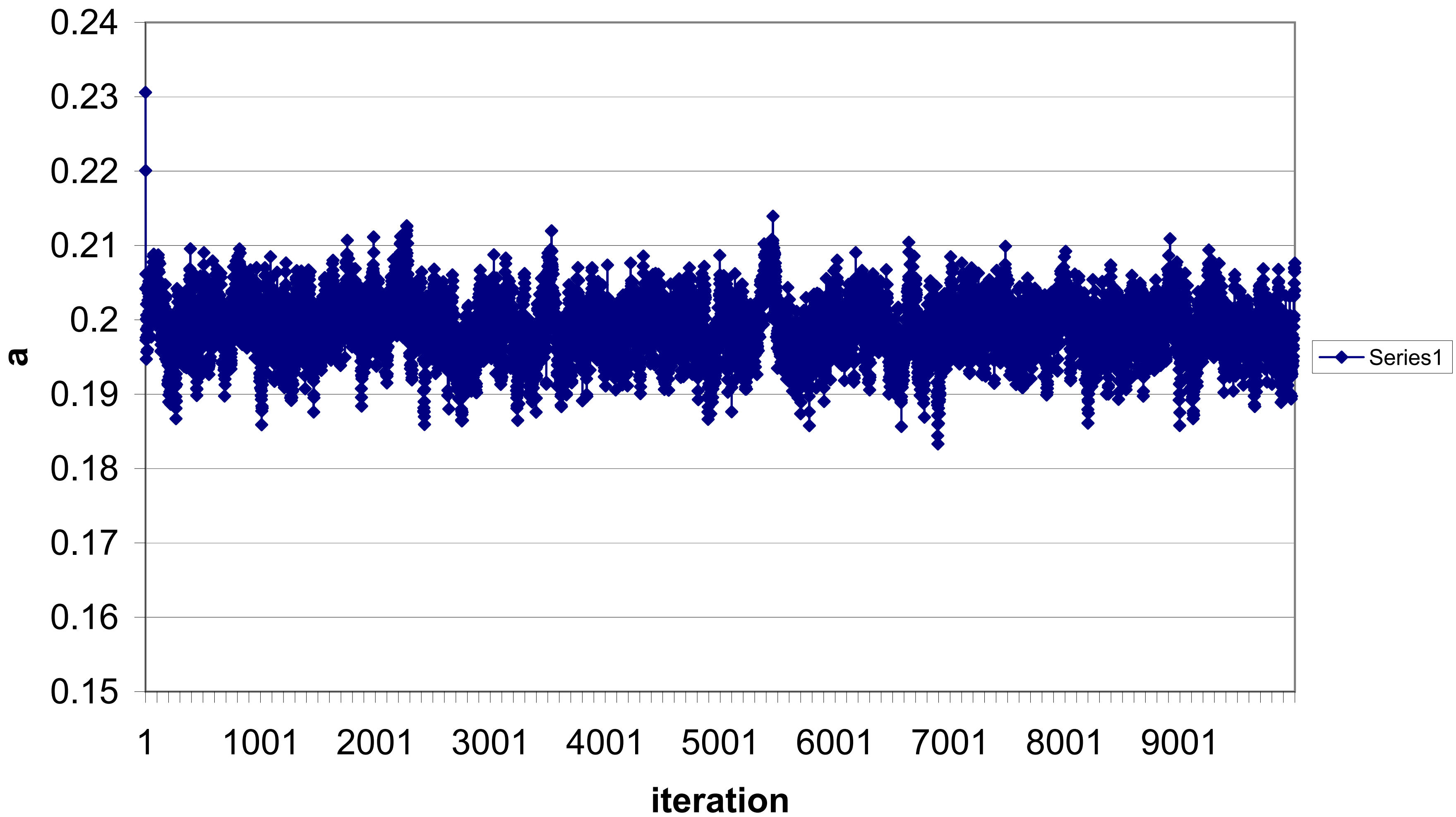

Figure 4: Gibbs Sampler Output of the Profit Shock Standard Error (True Value: 0.4)

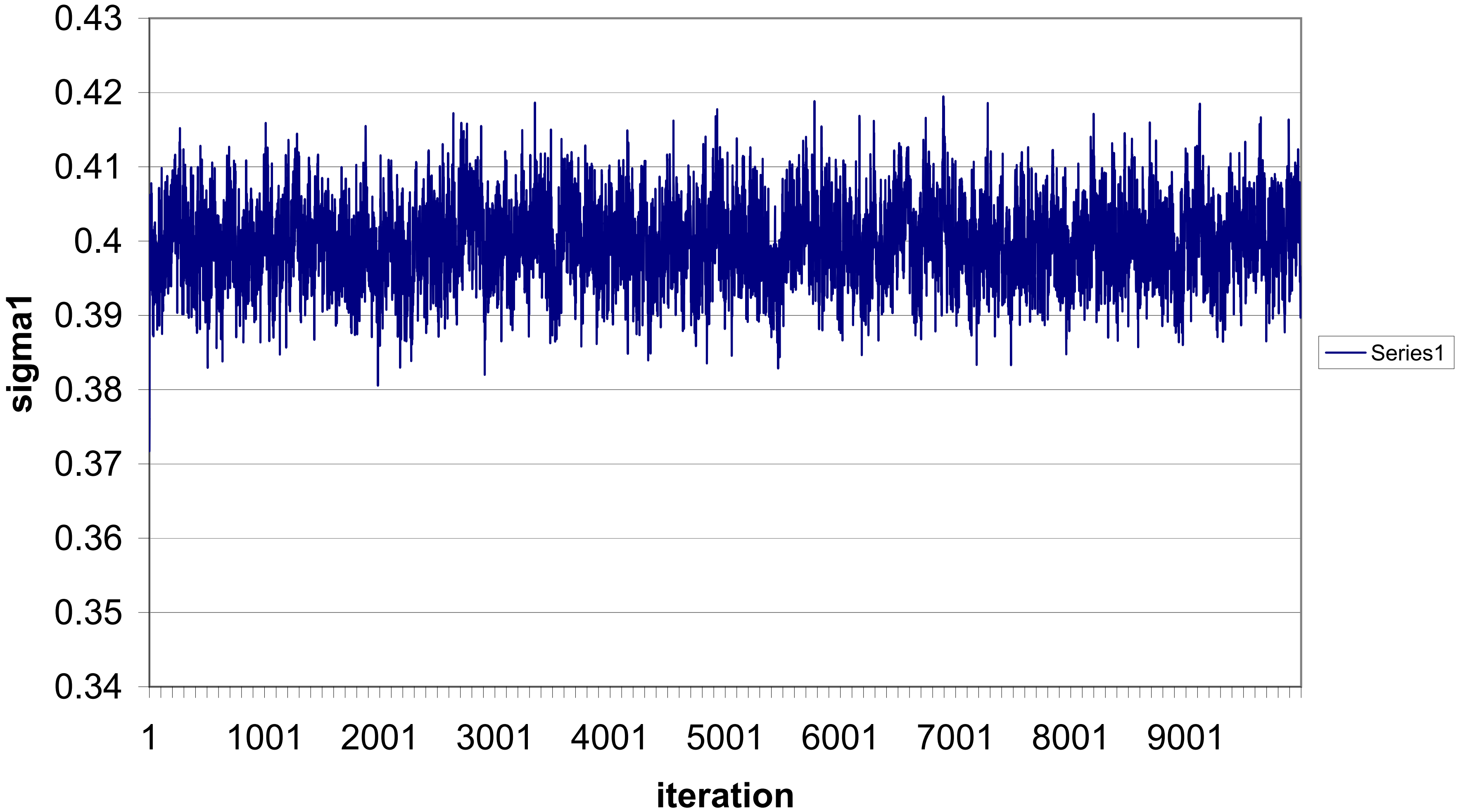

Figure 5: Gibbs Sampler Output of the Entry and Exit Shock Standard Error (True Value: 0.4)

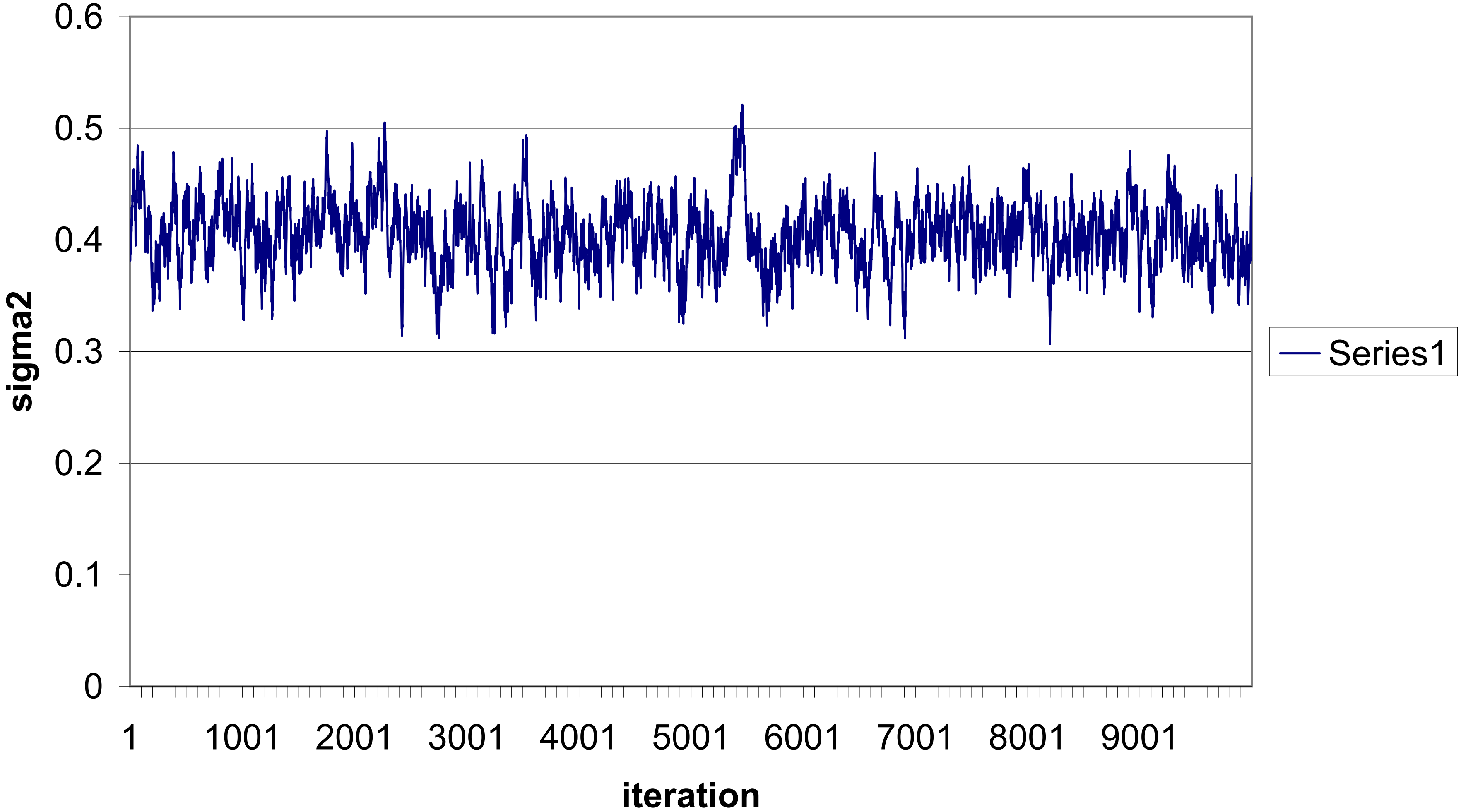

Figure 6: Gibbs Sampler Output of the Capital Stock Transition Parameter b1 (True Value: 0.1)

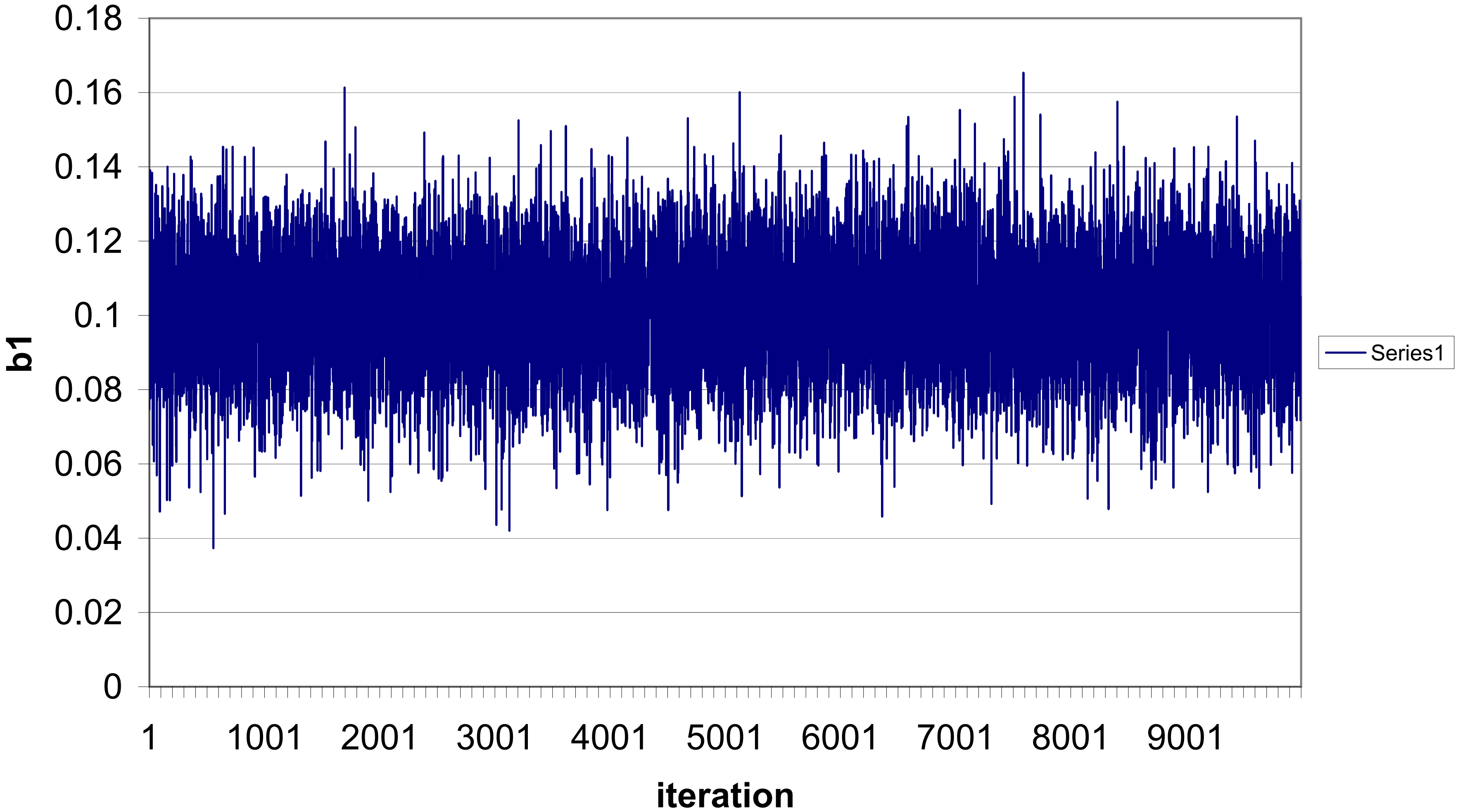

Figure 7: Gibbs Sampling Output of Capital Transition Parameter b2 (True Value: 0.1)

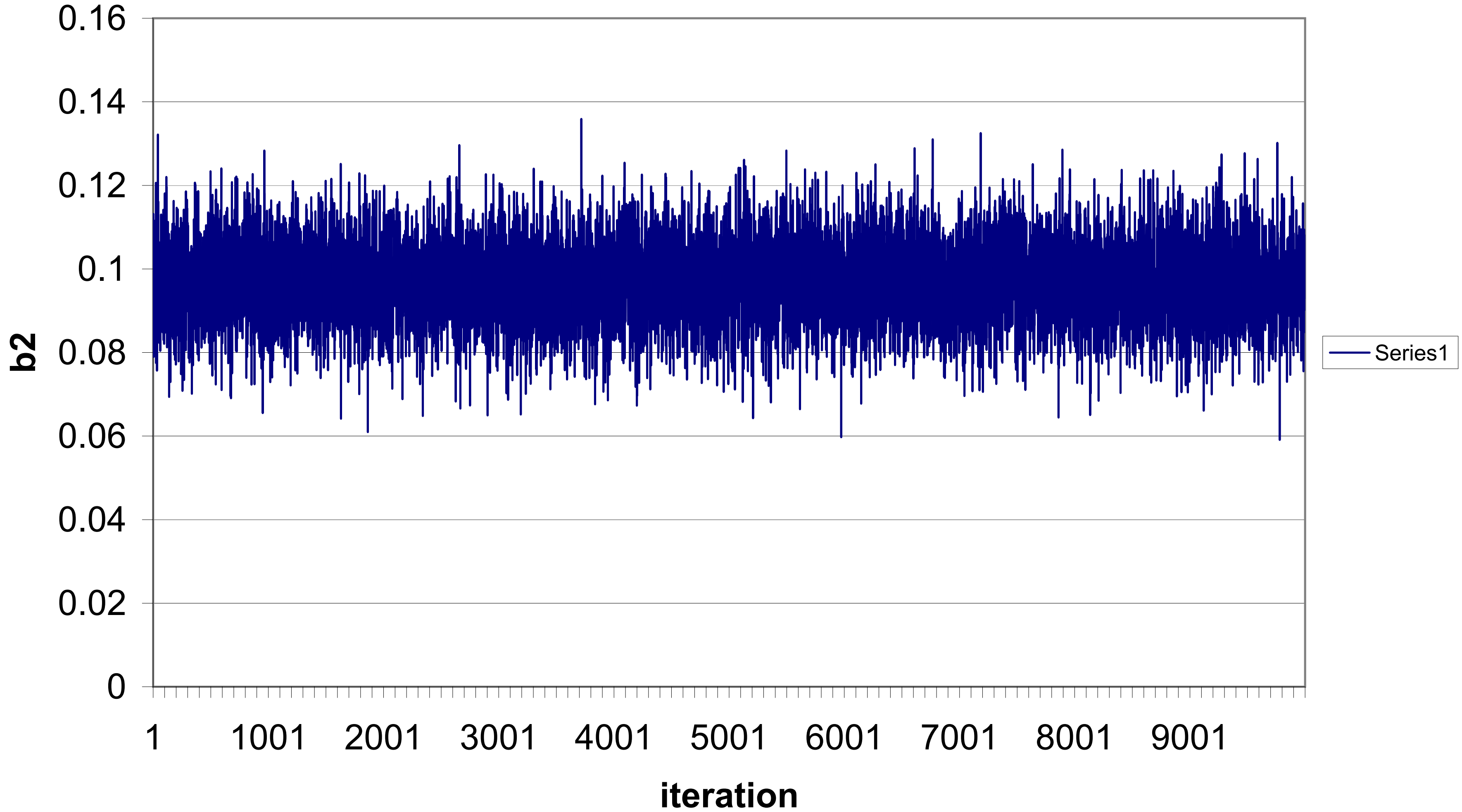

Figure 8: Gibbs Sampler Output of Capital Stock Transition Shock Standard Error sigmau (True Value: 0.4)