Minimum Distance Estimation for a Class of Markov Decision Processes*

Sorawoot Srisuma†,‡

London School of Economics and Political Science

August 2, 2010

Abstract

We develop a two-step estimator for a class of Markov decision processes with continuous control that is intuitive and simple to implement. Making use of the monotonicity assumption we estimate the expected continuation value functions nonparametrically in the first stage. In the second stage our estimator minimizes a minimum distance criterion that measures the divergence between the nonparametric conditional distribution function and a model implied simulated semiparametric counterpart. We show that our minimum distance estimator is asymptotically normal and converges at the parametric rate under some regularity conditions. We estimate the expected value function by kernel smoothing and derive its pointwise distribution theory. We illustrate how our estimation methodology forms a basis for the estimation of dynamic models with different class of control variable(s) as well as a class of Markovian games.

KEYWORDS: Markov Decision Models, Kernel Smoothing, Semiparamatric Estimation with Nonsmooth Objective Functions

---

*Previously circulated under the title "*Semiparametric Estimation of Markov Decision Processes: Continuous Control.*"

†I am most grateful to my PhD advisor, Oliver Linton, for his encouragement and guidance. I wish to thank Emmanuel Guerre, Jim Heckman, Javier Hidalgo, Stefan Hoderlein, Tatiana Komarova, Martin Pesendorfer, Marcia Schafgans, Philipp Schmidt-Dengler, Myung Seo, and seminar participants at many universities for valuable comments and suggestions. I have also benefited from various discussions with Joachim Groeger. This research is partially supported by the ESRC.

‡Department of Economics, London School of Economics, Houghton Street, London, WC2A 2AE, United Kingdom. E-mail address: `s.t.srisuma@lse.ac.uk`

1 Introduction

In this paper we propose a two-step estimation methodology for a class of Markov decision models without solving the numerical dynamic programming problem. Our work builds on the literature for estimating Markov decision problems based on the conditional independence assumption of Rust (1987) and the subsequent two-step approaches following Hotz and Miller (1993). We first focus on a single agent problem with continuously distributed control variable. We then show the same estimation principles can be used to estimate dynamic models with different types of controls, as well as a popular class of Markovian games. Most methodology papers in the literature consider a single class of control, many with purely discrete choice. One notable exception to this is Bajari, Benkard and Levin (2007), hereafter BBL, who illustrate that their forward simulation method can be used to estimate dynamic discrete choice models as well as those with continuous components.

A common obstacle in estimating a general dynamic programming model is the presence of the value function. Although some purely continuous control problems can use Euler equations to bypass this issue, e.g. see Hansen and Singleton (1982), more generally Euler equations may not be easy, if at all possible, to derive, see Pakes (1994). The latter point is particularly relevant for dynamic models with strategic interactions, which is of significant interest in the empirical I.O. literature. In dealing with the value functions, inspired by Hotz and Miller (1993), our general two-step approach is similar in spirit to the work of Pesendorfer and Schmidt-Dengler (2008) who use the policy value equation in the first stage to identify and estimate their dynamic discrete action games. This is in contrast to the forward simulation methodology of BBL, which has been motivated by Hotz, Miller, Sanders and Smith (1994).[1]

We define our estimator to minimize the distance between the conditional distribution functions implied by the model and the data. Under purely discrete choice setting, this is a version of Hotz and Miller conditional choice probabilities estimator. To estimate a continuous control problem, we make use of the monotonicity assumption on the policy function to recover the unobserved state variables nonparametrically. This assumption has been utilized in Olley and Pakes (1996), and more recently by BBL and Hong and Shum (2009).

Although there is an intrinsic value in providing an estimator for purely continuous control prob-

[1] Under the infinite time horizon framework, which has become increasingly popular in applied work, the computational advantage that motivates the simulation approach of Hotz et al. (1994) over Hotz and Miller (1993) is removed.

lems that do not rely on Euler's equation, since there are few methodologies available at present.[2,3] The main goal of this paper is to provide a unified approach to estimate a variety of dynamic optimization models. Our work is complementary to BBL's in providing general methods to estimate structural models without solving out the equilibrium. However, a distinctive feature in our approach is that we use the information of the model to estimate our parameter of interest. We provide a set of primitive conditions to ensure consistent estimation under an interpretable identifying assumption. In constrast, most known applications of BBL's methodologies consider a class of inequalities that are ad hoc random perturbations of the policy observed from the observed data. Since there is no theory for selecting an appropriate class of inequalities even if one assumes a high level identification assumption analogous to ours, the objective functions contructed from these local perturbations may fail to provide a consistent estimator for the parameter of interest, see Srisuma (2010).

We take a semiparametric approach. Since the transition law of the observables is one of the model primitives, which we often have no prior information on, it is desirable to be as flexible as possible when modeling this. In constrast to discrete decision problems, where the traditional nonparametric technique leads to the frequency estimator, models with continuous choice require smoothing. A well known drawback from using nonparametric methods in practice is the curse of dimensionality. This is less of an issue in this literature as applied researchers often assume that the observable state variables only take on finitely many values and the continuous component of the control variable is usually one dimensional; so the effective rate of convergence of the nonparametric estimator is determined by a one dimensional object.

The paper proceeds as follows. The next section begins by describing the Markov decision model of interest for a single agent problem and provides a simple example that motivates our methodology, it then outlines the estimation strategy and discusses the computational aspect. Section 3 provides the conditions to obtain the desired distribution theory and discuss inference based on semiparametric bootstrap. Section 4 reports a Monte Carlo study of our estimator and illustrates the affects of ignoring the model dynamics. Section 5 extends our methodology to estimate Markovian processes with discrete/continuous controls, ordered discrete response as well as a class of Markovian games and considers the estimation problem when the observable state space is uncountably infinite. Section 6 concludes. The proofs of our Theorems can be found in the Appendix and we collect the Tables at the end of the paper.

[2]Berry and Pakes (2000) propose an estimation method to estimate a dynamic oligopoly model with purely continuous control. However, their framework is rather different to ours, in particular they do have any unobserved state variables in their model.

[3]Another example can be found in the dynamic auction literature, see Jofre-Bonet and Pesendorfer (2003).

2 Markov Decision Processes

2.1 Basic Framework

Time is indexed by t, each economic agent, i, is forward looking in solving an infinite horizon intertemporal problem. The random variables in the model are the control and state variables, denoted by a_{it} and s_{it} respectively. The support of control variable is a convex set $A \subset \mathbb{R}$. and the state space S is a subset of $\mathbb{R}^{L+1}$. In each period, the economic agent observes s_{it} and chooses an action a_{it} in order to maximize her discounted expected utility. The per period utility function is time separable and is represented by $u\left(a_{it}, s_{it}\right)$. The agent's action today affects the uncertain future states according to a Markovian transition law $F\left(s_{it+1}|s_{it}, a_{it}\right)$. Next period's utility is subjected to discounting at a rate $\beta \in (0,1)$, which is assumed to be known. Formally the agent is represented by a triple of primitives (u, β, F), who is assumed to behave according to an optimal decision rule $\left\{\alpha\left(s_{it}\right)\right\}_{t=\tau}^{\infty}$, in solving the following sequential problem:

$$V\left(s_{it}\right) = \max_{\left\{a\left(s_{i\tau}\right)\right\}_{\tau=t}^{\infty}} E\left[\left.\sum_{\tau=t}^{\infty} \beta^{\tau-t} u\left(a\left(s_{i\tau}\right), s_{i\tau}\right)\right| s_{it}\right], \text{s.t. } a\left(s_{it}\right) \in A \text{ for all } \tau \geq t. \quad (1)$$

Under some regularity conditions, there exists a time invariant Markovian optimal decision rule $\alpha\left(\cdot\right)$ so that

$$\alpha\left(s_{it}\right) = \arg\max_{a \in A} \left\{u\left(a, s_{it}\right) + \beta E\left[V\left(s_{it+1}\right)|s_{it}, a_{it} = a\right]\right\}. \quad (2)$$

Furthermore, the value function, V, is the unique stationary solution to the Bellman's equation

$$V\left(s_{it}\right) = \max_{a \in A} \left\{u\left(a, s_{it}\right) + \beta E\left[V\left(s_{it+1}\right)|s_{it}, a_{it} = a\right]\right\}. \quad (3)$$

More details on the mathematical theory and of specific Markov decision models that are commonly used in economics can be found in Pakes (1994) and Rust (1994). In order to avoid a degenerate model, we assume that the state variables $s_{it} = \left(x_{it}, \varepsilon_{it}\right) \in \mathbb{R}^L \times \mathbb{R}$ can be separated into two parts, which are observable and unobservable respectively to the econometrician; see Rust (1994) for various interpretations of this unobserved heterogeneity. We next provide an economic example that naturally fits in our dynamic decision making framework.

Dynamic Price Setting Example:

Consider a dynamic price setting problem for a firm. Each period t, firm i faces the demand for some goods summarized by $D\left(a_{it}, x_{it}, \varepsilon_{it}\right)$ where: a_{it} denotes the price; x_{it} is some measure of the consumer's satisfaction that affects the level of the demand for the immediate period that is publically observed; ε_{it} is the firm's private demand shock. The firm sets a price and earns the following immediate profit

$$u\left(a_{it}, x_{it}, \varepsilon_{it}\right) = D\left(a_{it}, x_{it}, \varepsilon_{it}\right)\left(a_{it} - c\right),$$

where c denotes a constant marginal cost. The price setting decision affects the sentiment of the demand of the consumers for the next period, x_{it+1}, that can be modelled by some Markov process. So the firm chooses price a_{it} to maximize its discounted expected profit

$$a_{it} = \arg\max_{a \in A} \{u(a, x_{it}, \varepsilon_{it}) + \beta E[V(s_{it+1}) | x_{it}, \varepsilon_{it}, a_{it} = a]\}.$$

In Section 5, we focus on a specific example of this dynamic price setting problem and use a Monte Carlo experiment to illustrate the finite sample behavior of our estimator as well as the effects of ignoring the underlying dynamics in the model.

Now we introduce the main modelling assumptions that are standard assumptions in this literature; unless stated otherwise they are assumed to hold throughout the paper.

ASSUMPTION M1: *(Conditional Independence) The transitional distribution has the following factorization:* $F(x_{it+1}, \varepsilon_{it+1} | x_{it}, \varepsilon_{it}, a_{it}) = Q(\varepsilon_{it+1}) p_{X'|X,A}(x_{it+1} | x_{it}, a_{it})$*, where* Q *is a known cdf of* ε_t *that is absolutely continuous with respect to the Lebesgue measure on* $\mathcal{E}$ *with density* q.

The conditional independence assumption of Rust (1987) imposes restrictions on the dependent structure of our stochastic process. In particular, it implies that $\{\varepsilon_{it}\}$ is an i.i.d. process independent of $\{x_{it}\}$ and all variables in the past, and ε_{it} only correlated to x_{it+1} through the choice variable a_{it}. It is conceptually straightforward to relax the former condition and have ε_{it} to be conditionally independent of the past given x_{it}. The functional form for the cdf of ε_t can be relaxed to upto any finite dimensional parameterization.

ASSUMPTION M2: *The support of* $s_{it} = (x_{it}, \varepsilon_{it})$ *is* $X \times \mathcal{E}$*, where* $X = \{1, \ldots, J\}$*, for some* $J < \infty$*, denotes the observable state space and* $\mathcal{E} \subseteq \mathbb{R}$.

Finiteness of X is a simplifying assumption. It is generally feasible to allow for continuous component in x_{it}. A rigorous treatment of this extension can be found in Srisuma and Linton (2009), although they only consider a discrete control problem, their techniques can directly be applied to other dynamic models as well.

ASSUMPTION M3: *(Monotone Choice) The per period payoff function* $u : A \times X \times \mathcal{E} \to \mathbb{R}$ *has increasing differences in* (a, ε) *for all* x.

The monotonocity assumption is crucial in our methodology since ε necessarily enters u non-additively. However, this condition can be empirically motivated, see Olley and Pakes (1996), BBL and Hong and Shum (2009). In particular, the implication of M3 together with M1 is that policy function is increasing on $\mathcal{E}$; this follows since M1 implies the policy function can be defined as (cf. (2))

$$\alpha(x, \varepsilon) = \arg\max_{a \in A} \{u(a, x, \varepsilon) + \beta E[V(s_{it+1}) | x_{it} = x, a_{it} = a]\}. \quad (4)$$

The function to be maximized on the RHS is supermodular in (a, ε) for all x, the claim follows from Topkis' theorem (see Topkis (1998)).

2.2 Estimation Methodology

In practice, researchers are often interested in some structural parameters $\theta \in \Theta \subset \mathbb{R}^M$ that parameterize the payoff function u_θ. Suppose we have a set of balanced panel data $\{a_{it}, x_{it}\}_{i=1,t=1}^{N,T}$ from N i.i.d. agents that have been generated by a controlled Markovian process with primitives (u_{θ_0}, β, F). Throughout this paper, for notational simplicity, we suppress the dependence on θ_0 for all functions generated from (u_{θ_0}, β, F).

Methodology Outline

Our starting point is the (a.e.) one-to-one relation between the optimal policy function that generates the data and its corresponding conditional distribution function

$$F_{A|X}(a|x) = \Pr\left[\alpha(x_{it}, \varepsilon_{it}) \leq a | x_{it} = x\right].$$

We consider their reduced form counterparts, indexed by θ, which are generated from a family of policy value functions $\{V_\theta\}_{\theta\in\Theta}$, to be defined below.[4] We define the model implied policy and distribution functions as follows

$$\begin{aligned} \alpha_\theta(x, \varepsilon) &= \arg\max_{a\in A}\left\{u_\theta(a, x, \varepsilon) + \beta E\left[V_\theta(s_{it+1})|x_{it} = x, a_{it} = a\right]\right\}, & (5) \\ F_{A|X}(a|x;\theta) &= \Pr\left[\alpha_\theta(x_{it}, \varepsilon_{it}) \leq a|x_{it} = x\right]. & (6) \end{aligned}$$

We shall show that $F_{A|X}(\cdot;\theta) = F_{A|X}$ when $\theta = \theta_0$, and base our estimation criterion on the distance between the empirical versions of $F_{A|X}(\cdot;\theta)$ and $F_{A|X}$. However, the choice specific expected future value $E\left[V_\theta(s_{it+1})|x_{it}, a_{it}\right]$, which we denote by g_θ, is a solution to some fixed point problem that does not have a closed form. The first part of our two-stage procedures consists two steps: (i) estimation of g_θ; (ii) use the estimated g_θ to construct the objective functions on the RHS of (5) that allows us to simulate the model implied distribution function $F_{A|X}(\cdot;\theta)$. The second is the optimization stage, where a class of minimum distance estimators can then be derived from minimizing the distribution functions observed from the data and simulated from the first stage.

First Stage Nonparametrics

STEP 1: POLICY VALUE FUNCTIONS

[4]See Rust (1994) and Magnac and Thesmar (2002) for the definitions of *reduced forms* in closely related contexts.

For any $\theta \in \Theta$ we define a *policy value function*, V_θ, by

$$V_\theta(s) = E\left[\sum_{t=0}^{\infty} \beta^t u_\theta(\alpha(s_{it}), s_{it}) \middle| s_{i0} = s\right] \quad \text{for any} \quad s \in S.$$

By definition of the policy function α that generates the data, V_θ coincides with the solution to the sequential problem defined in (1) when $\theta = \theta_0$. From the Neumann series expansion, V_θ can also be written as a (stationary) solultion to the following policy value equation (cf. equation (3))

$$V_\theta(s_{it}) = u_\theta(\alpha(s_{it}), s_{it}) + \beta E[V_\theta(s_{it+1})|s_{it}]. \tag{7}$$

We can interpret V_θ as the value function for an economic agent whose underlying preference is θ but is using the policy function that is optimal with respect to θ_0. For notational simplicity, we henceforth replace $\alpha(s_{it})$ with a_{it}.

The goal of the first step is to estimate g_θ, where under M1 it can be written as the following conditional expectation

$$g_\theta(a, x) = E[E[V_\theta(s_{it+1})|x_{it+1}]|x_{it} = x, a_{it} = a] \text{ for any } (a, x) \in A \times X.$$

Marginalizing out the unobserved states in policy equation (7), under M1, we obtain an analogous characterization of the conditional policy value function from

$$E[V_\theta(s_{it})|x_{it}] = E[u_\theta(a_{it}, x_{it}, \varepsilon_{it})|x_{it}] + \beta E[E[V_\theta(s_{it+1})|x_{it+1}]|x_{it}].$$

Since $|X| = J$, the equation above can be conveniently summarized by a matrix equation as

$$m_\theta = r_\theta + \mathcal{L} m_\theta. \tag{8}$$

where r_θ denotes a $J-$dimensional vector of $(E[u_\theta(a_{it}, x_{it}, \varepsilon_{it})|x_{it} = x])_{x=1}^J$, and $\mathcal{L}$ is a $J \times J$ discounted stochastic matrix whose $(x', x)-$th entry represents $\beta \Pr[x_{it+1} = x'|x_{it} = x]$. The matrix $(I - \mathcal{L})$ is invertible by the dominant diagonal theorem, therefore the $J-$dimensional vector of conditional value functions, m_θ, is the unique solution to equation (8). Therefore the choice specific expected value can be written in a linear functional notation

$$g_\theta = \mathcal{H} m_\theta, \tag{9}$$

where $\mathcal{H}$ is a conditional expectation operator so that, $\mathcal{H}\phi(x, a) = \sum_{x' \in X} \phi(x') p_{X'|X,A}(x'|x, a)$ for any function ϕ of x_{it}.

To implement, one proceeds by estimating the matrix equation in (8). This requires estimators for r_θ and $\mathcal{L}$. Although we do not observe $\{\varepsilon_{it}\}$, we can use the monotonicity assumption to generate their nonparametric estimates from the following relation

$$\widehat{\varepsilon}_{it} = Q^{-1}\left(\widehat{F}_{A|X}(a_{it}|x_{it})\right), \tag{10}$$

where Q is the known distribution function of ε_{it};[5] and $\widehat{F}_{A|X}(a|x) = \frac{1}{NT}\sum_{i=1,t=1}^{N,T} w_{itN}(x)\mathbf{1}[a_{it} \leq a]$ is an estimator for $F_{A|X}$, where $w_{itN}(x) = \frac{\mathbf{1}[x_{it}=x]}{\widehat{p}_X(x)}$ with $\widehat{p}_X(x) = \frac{1}{NT}\sum_{i=1,t=1}^{N,T}\mathbf{1}[x_{it} = x]$. Given $\{a_{it}, x_{it}, \widehat{\varepsilon}_{it}\}_{i=1,t=1}^{N,T}$, we can estimate r_θ by

$$\widetilde{r}_\theta(x) = \frac{1}{NT}\sum_{i=1,t=1}^{N,T} w_{itN}(x)\, u_\theta(a_{it}, x_{it}, \widehat{\varepsilon}_{it}). \tag{11}$$

Assuming further that $\widehat{p}_X(x) > 0$ for all x, we again use the frequency estimators to estimate the elements in $\mathcal{L}$. The dominant diagonal theorem implies $\left(I - \widehat{\mathcal{L}}\right)^{-1}$ exists, we can then uniquely estimate the conditional value function by the relation

$$\widehat{m}_\theta = \left(I - \widehat{\mathcal{L}}\right)^{-1}\widetilde{r}_\theta. \tag{12}$$

In order to estimate g_θ, we now only need to estimate the conditional expectation operator $\mathcal{H}$. We employ the Nadaraya-Watson type estimator:

$$\widehat{g}_\theta = \widehat{\mathcal{H}}\widehat{m}_\theta, \tag{13}$$

where for any $a \in int(A)$

$$\begin{aligned}
\widehat{g}_\theta(a,x) &= \sum_{x' \in X}\widehat{m}_\theta(x')\frac{\widehat{p}_{X',X,A}(x',x,a)}{\widehat{p}_{X,A}(x,a)}, \qquad (14)\\
\widehat{p}_{X',X,A}(x',x,a) &= \frac{1}{NT}\sum_{i=1,t=1}^{N,T}\mathbf{1}[x_{it+1} = x', x_{it} = x]\, K_h(a_{it} - a),\\
\widehat{p}_{X,A}(x,a) &= \frac{1}{NT}\sum_{i=1,t=1}^{N,T}\mathbf{1}[x_{it} = x]\, K_h(a_{it} - a),
\end{aligned}$$

where $\widehat{p}_{X',X,A}$ denotes our estimator for $p_{X',X,A}$, the mixed-continuous joint density of $(x_{it+1}, x_{it}, a_{it})$; $\widehat{p}_{X,A}$ and $p_{X,A}$ are defined similarly; $K_h(\cdot) = \frac{1}{h}K\left(\frac{\cdot}{h}\right)$ denotes a user-chosen kernel and h is the bandwidth. Typically we may need to trim off the estimates near the boundaries or the tails of the distribution, this discussion is deferred until Section 3.

Step 2: Model Implied Distribution Functions

For any $\theta \in \Theta$, we consider the model implied objective function π_θ where

$$\pi_\theta(a,x,\varepsilon) = u_\theta(a,x,\varepsilon) + \beta g_\theta(a,x).$$

[5] BBL also uses the one to one correspondence between a_{it} and ε_{it} in their forward simulation method, where they draw $\{\varepsilon_b\}$ and generate the corresponding optimal choice from $\left\{\widehat{F}_{A|X}^{-1}(Q_\varepsilon(\varepsilon_b)|x_b)\right\}$, for any state x_b.

Since π_θ is the objective function on the RHS of (5) that defines α_θ, naturally we define the model implied policy by plugging in $\widehat{g}_\theta$, namely

$$\widehat{\alpha}_\theta(x,\varepsilon) = \arg\max_{a\in A}\left\{\widehat{\pi}_\theta(a,x,\varepsilon)\right\},$$

where $\widehat{\pi}_\theta(a,x,\varepsilon) = u_\theta(a,x,\varepsilon) + \beta\widehat{g}_\theta(a,x)$. The corresponding conditional distribution function can then be simulated by to approximate $F_{A|X}(\cdot;\theta)$. In particular, we define

$$\widetilde{F}_{A|X}(a|x;\theta) = \frac{1}{R}\sum_{r=1}^{R}\mathbf{1}\left[\widehat{\alpha}_\theta(x,\varepsilon_r)\le a\right],$$

where $\{\varepsilon_r\}_{r=1}^R$ is a random sample drawn from the known distribution of ε_{it}.

Second Stage Optimization

By construction of the policy values and their related functions, we must have $V_\theta = V$, $\alpha_\theta = \alpha$ and $F_{A|X}(;\theta) = F_{A|X}$ when $\theta = \theta_0$. We can then write down the following set of continuum of moment restrictions (see Carrasco and Florens (2000))

$$E\left[\mathbf{1}\left[a_{it}\le a\right] - F_{A|X}(a|x_{it};\theta)\big|\, x_{it}\right] = 0, \text{ for } a\in A \text{ when } \theta=\theta_0. \tag{15}$$

Since no general theory for semiparametric moment estimation with a continuum of moments is available at present. We instead use the following equivalent interpretation to these moment conditions as a basis for our estimation problem

$$F_{A|X}(a|x) = F_{A|X}(a|x;\theta) \quad a.e. \quad \text{when} \quad \theta=\theta_0.$$

We focus on a class of minimum distance estimators.[6] Wolfowitz (1953) introduce the minimum distance method that since has developed into a general estimation technique that has well known robustness and efficiency properties, see Koul (2002) for a review. In this paper, we define a class of estimators that minimize the following Cramér von-Mises type objective function that defines some L^2-distance between the CDF implied the model and that of the data

$$M_N(\theta) = \sum_{x\in X}\int_A\left[\widetilde{F}_{A|X}(a|x;\theta) - \widehat{F}_{A|X}(a|x)\right]^2\mu_x(da). \tag{16}$$

Here $\{\mu_x\}$ is a sequence of user chosen sigma-finite (for now, assume non-random) measures on A. Clearly the property of $\widehat{\theta}$ will depend on the choice of measures. In Section 3, we provide a discussion on how to select the measures to ensure consistent estimation under some regularity conditions. However, we leave the issue of choosing $\{\mu_x\}$ efficiently for future work.

[6] Another alternative to the moment based estimator is to maximize the conditional maximum likelihood function, however the maximum likelihood estimator (MLE) is much more computationally demanding. Although one can proceed with our minimum distance approach and perform Newton Ralphson type iterations to ensure we get the same first order asymptotic distribution as the conditional MLE.

2.3 Practical Aspects

First note that all elements we require to solve and transform the linear equations in (8) and (9) have explicit functional forms, so they are easy to program. In addition, similar to Hotz et al. (1994) and BBL, we can also take advantage of the linear structure the policy value equation. In particular, if the parameterization of θ in u is linear so that $u_\theta = \theta^\intercal u_0$, then $\widetilde{r}_\theta(j)$ can be writen as $\theta^\intercal \widetilde{r}_0$ where $\widetilde{r}_0(x) = \frac{1}{NT}\sum_{i=1,t=1}^{N,T} w_{itN}(x) u_0(a_{it}, x_{it}, \widehat{\varepsilon}_{it})$ for each x. In a vector form, $\widetilde{r}_\theta = W\widehat{u}_\theta = W_{\widehat{u}}\theta$ for some matrix $W_{\widehat{u}}$. To estimate the conditional value function $\widehat{m}_\theta = \left(I - \widehat{\mathcal{L}}\right)^{-1} W_{\widehat{u}}\theta$, since we estimate $\mathcal{L}$ nonparametrically, therefore we only have to compute the matrix $\left(I - \widehat{\mathcal{L}}\right)^{-1} W_{\widehat{u}}$ once as it does not depend on θ. We comment that, since the dimension of A and $\mathcal{E}$ are both one, this allows one to rapidly search for the maximizer of $\widehat{\pi}_\theta(\cdot, x, \varepsilon)$ for the entire set of $\{\varepsilon_r\}_{r=1}^R$ by grid search for different values of θ.

It is also straightforward to carry out our methodology in a fully parametric framework. One can choose to parameterize the transition law $p_{X'|X,A}(\cdot;\theta_{tr})$, for some θ_{tr}. The choice specific continuation value function is still defined through equation (9) where the conditional expectation operator becomes $\mathcal{H}_{\theta_{tr}}$. For a fixed θ_{tr}, we can estimate $\widehat{g}_\theta$ using the relation (13) by simply replacing $\frac{\widehat{p}_{X',X,A}}{\widehat{p}_{X,A}}$ in equation (14) by $p_{X'|X,A}(\cdot;\theta_{tr})$. Although the conditional expectation operator $\mathcal{H}_{\theta_{tr}}$ depends on θ_{tr}, it does not affect how we estimate $\widehat{m}_\theta$. Note also that all the subsequent stages of the methodology only assume we have $\widehat{g}_\theta$ and not how they are obtained, therefore the remaining steps in our procedure remains unchanged.

3 Distribution Theory

Our goal is to estimate the structural parameter, θ_0, that generates $\{a_{it}, x_{it}\}$. Our minimum distance estimator, which falls in the class of a profiled semiparametric M-estimator with non-smooth objective function, is any sequence of $\widehat{\theta}$ that satisfies

$$M_N\left(\widehat{\theta}\right) \leq \inf_{\theta\in\Theta} M_N(\theta) + o_p\left(N^{-1}\right), \tag{17}$$

where $\widehat{g}_\theta$ and M_N are defined in (14) and (16) in the previous section.

We now provide sufficient conditions to derive the asymptotic theory for our estimators.

Assumption M4:

(i) $\{a_{it}, x_{it}\}_{i=1,t=1}^{N,T}$ is i.i.d. across i, for each i, $\{a_{it}, x_{it}\}_{t=1}^{T}$ is a strictly stationary realizations of the controlled Markov process for a fixed periods of T with exogenous initial values; (ii) A and $\mathcal{E}$ are compact and convex subsets of $\mathbb{R}$; (iii) Let Θ be a compact subset of $\mathbb{R}^M$, then the following

condition holds for all x

$$\alpha_\theta(x,\cdot)=\alpha(x,\cdot) \quad Q-a.e. \text{ if and only if } \theta=\theta_0,$$

where $\theta_0\in int(\Theta)$*; (iv) For each* x,μ_x *is a finite measure on* A *that dominates* Q *and has zero measure on the boundary of* A*;*

(i) Exogenous initial value is often assumed in this literature. We anticipate a data set with short panel, however, we can also allow for large T asymptotics with no modification for the procedure, see Srisuma and Linton (2009); (ii) We assume the support of A (by monotonicity of the policy function, hence $\mathcal{E}$) to be compact for simplicity, otherwise one just add some standard trimming conditions discussed in Robinson (1988); (iii) This is the main identification assumption. It can be shown directly that the conditions we impose on the policy functions above is equivalent to imposing that the moment restrictions in (15) hold if and only if $\theta=\theta_0$. This condition has the interpretation that the only model implied policy function induced by θ that is consistent with the data is only when $\theta=\theta_0$; (iv) We only consider the object functions that make use of all the information from the moment restrictions, which ensures they will converge to functions that have a unique minimum at θ_0 in the limit. One possible choice for $\{\mu_x\}$ that satisfies this condition is simply the Lebesgue measure. Alternatively, we can also use random measures, as long as they converge weakly to $\{\mu_x\}$ that satisfy the finiteness and dominant conditions.[7] The most natural choice of random measures that one may use is the empirical measure, which puts equal mass on each observed data points $\{a_{it},x_{it}\}$ and zero measure outside it.

Next we impose some conditions on the kernel function and simulation size:

Assumption M5:

(i) K *is an even and continuously differentiable 4-th order kernel function on* $[-1,1]$*; (ii) The bandwidth sequence* h_N *satisfies* $h_N=d_NN^{-\varsigma}$ *for* $1/8<\varsigma<1/6$*, with* d_N *is a sequence of real numbers that is bounded away from zero and infinity; (iii) Trimming factor* $\gamma_N=o(1)$ *and* $h_N=o(\gamma_N)$*; (iv) The simulation size* R *satisfies* $N/R=o(1)$*;*

(i) We need to use a higher order kernel despite the fact that $\dim(A)=1$ since the effective nonparametric rates we are dealing with corresponds to that of a nonparametric estimator for the derivative of a density function; this can be seen from defining the policy function as an implicit function satisfying the following first order condition $\frac{\partial}{\partial a}\pi_\theta(\alpha_\theta(x,\varepsilon),x,\varepsilon)=0$; (ii) We provide the range of bandwidths, which will ensure root$-N$ consistent estimation of our semiparametric estimator; (iii)

[7] In the Appendix we simply consider sequence of nonrandom measures $\{\mu_j\}$. However, the proofs can be lengthened leading to the same asymptotic results for random measures $\{\mu_{Nj}\}$, where $\mu_{Nj}\Rightarrow\mu_j$ for each j; this can be shown to follow from repeated applications of continuous mapping theorem following the results from Ranga Rao (1962).

Since we assume bounded support on A, we need to trim off our estimates near the boundary. In practice, this can be done by simply approximating the policy function on a compact subset of A. (iv) requires the simulation size to diverge faster than N to ensure we can ignore the approximation error of $F_{A|X}(\cdot|\cdot;\theta)$, which arises from simulation, in our asymptotic results.

We can also write down the general expressions for higher order kernels of order r, see Robinson (1988). Since $\dim(A) = 1$, we explicitly work with a $4-th$ order kernel for transparency that avoids imposing unnecessary arbitrary degree of smoothness assumptions. In particular, we impose the following standard smoothness conditions on various density functions and the payoff function:

Assumption M6:

(i) For all $x \in X$, the density $p_{X,A}(x,\cdot)$ is 5-times continuously differentiable on A and $\inf_{a\in A} p_{X,A}(x,a) > 0$; (ii) For all $x', x \in X$ the density $p_{X',X,A}(x',x,\cdot)$ is 5-times continuously differentiable on A; (iii) The known distribution function of ε_{it}, Q, is Lipschitz continuous and twice continuously differentiable; (iv) For all $x \in X$, $u_\theta(a,x,\varepsilon)$ is twice continuously differentiable in a and θ, once continuously differentiable in ε, these continuous derivatives exist for all a, ε and θ. In addition we assume $\frac{\partial^2}{\partial a \partial \varepsilon} u_\theta(a,x,\varepsilon) > 0$ and $\frac{\partial^4}{\partial a^2 \partial \varepsilon^2} u_\theta(a,x,\varepsilon)$ exists and is continuous for all a, ε and θ;

We note that, given our smoothness assumptions, $\frac{\partial^2}{\partial a \partial \varepsilon} u_\theta(a,x,\varepsilon) > 0$ is the analytical condition for M3.

We need to introduce some notations before stating the last set of assumptions for our main theorems. By M1 and M3, our model implied policy functions are monotone in ε, so we can define a corresponding inverse of the policy function by $\rho_\theta(a,x)$. In addition, by implicit function theorem in Banach space, α_θ and ρ_θ (and $F_{A|X}(;\theta)$) depend on the function $\partial_a g_\theta$, where ∂_a is short for $\frac{\partial}{\partial a}$, we define the relevant normed space and make these dependence explicit in the Appendix. We denote the (partial-) Fréchet differential operators by D_θ and D_g, where the indices θ and g denote the finite and infinite dimensional argument used in differentiating.[8]

Assumption M7:

(i) The inverse of the policy function is twice Fréchet differentiable and $\|D_g \rho_\theta\| < \infty$; (ii) For some x in X, the following $M \times M$ matrix

$$\int_A [q(\rho_{\theta_0}(\cdot,x))]^2 \left[D_\theta \rho_{\theta_0}(\cdot,x) D_\theta \rho_{\theta_0}(\cdot,x)'\right] \mu_x(da)$$

[8]For any map $T : X \to Y$ and some Banach spaces X and Y, we say that T is Fréchet differentiable at x, that belongs to some open neighborhood of X, if and only if there exists a linear bounded map $D_T : X \to Y$ such that $T(x+f) - T(x) = D_T(x) f + o(\|f\|)$ with $\|f\| \to 0$ for all f in some neighborhood of x; we denote the Fréchet differential at x in a particular direction f by $D_T(x)[f]$.

is positive definite; (iii) For all a and x, the Fréchet differential of $\rho_{\theta_0}(a,x)$ w.r.t. $\partial_a g$ in the direction $[\partial_a \widehat{g}_{\theta_0}(\cdot,x) - \partial_a g_{\theta_0}(\cdot,x)]$ is asymptotically linear, in particular for any $a \in int(A)$

$$D_g \rho_{\theta_0}(a,x)[\partial_a \widehat{g}_{\theta_0}(\cdot,x) - \partial_a g_{\theta_0}(\cdot,x)] = \frac{1}{NT}\sum_{i=1,t=1}^{N,T} \psi_0(a_{it}, x_{it}; a, x) + o_p\left(N^{-1/2}\right),$$

with $E[\psi_0(a_{it}, x_{it}; a, x)] = 0$ and $E\left[\psi_0^2(a_{it}, x_{it}; a, x)\right] < \infty$ for all i,t; in addition, w.p.a. 1, the expression above holds uniformly on any strict compact subset A_δ of A and $\psi_0(a_{it}, x_{it}; \cdot) \in \Psi_\delta$ where Ψ_δ is some class of functions on A_δ that is a Donsker class.

(i) We impose the primitive conditions on the inverse of the policy function, as opposed to the policy function, for notation convenience. This is without any loss of generality by using implicit, inverse and Taylor's theorems in Banach space; (ii) This is a standard rank condition on the second derivative on the Hessian; (iii) We implicitly assume that the effects from using the nonparametric estimator can be captured by a leading term that satisfies the Donsker theorem. The pointwise linearization is a reasonable assumption that we expect to be satisfied in most applications.[9] For the Donsker property, first note that $\{\Psi_\delta\}$ only depends on δ through the support of A, we need the space of functions $\{\Psi_\delta\}$ to not be too rich (in our case, sufficiently smooth). The Donsker property that generalizes the central limit theorem to random elements is satisfied by a large class of functions, see van der Vaart and Wellner (1996).

THEOREM 1: *Under M1-M7: For $\widehat{g}_\theta$ that satisfies* (13), *if $\widehat{\theta}$ satisfies (17) then $\widehat{\theta} \xrightarrow{p} \theta_0$, and*

$$\sqrt{N}\left(\widehat{\theta} - \theta_0\right) \Rightarrow N\left(0, H_0^{-1}\Omega H_0^{-1}\right),$$

where

$$\begin{aligned}
\Omega &= \lim_{N\to\infty} \mathrm{var}\left(-2\sum_{x\in X}\int \begin{bmatrix} \left[D_\theta F_{A|X}(a|x;\theta_0)\right] \\ \times \\ \sqrt{N}\begin{bmatrix} \left(\widehat{F}_{A|X}(a|x) - F_{A|X}(a|x)\right) \\ -\left(D_g F_{A|X}(a|x;\theta_0)[\partial_a \widehat{g}_{\theta_0}(\cdot,x) - \partial_a g_{\theta_0}(\cdot,x)]\right) \end{bmatrix} \end{bmatrix} \mu_x(da)\right), \\
H_0 &= 2\sum_{x\in X}\int_A \left[D_\theta F_{A|X}(a|x;\theta_0)\, D_\theta F_{A|X}(a|x;\theta_0)\right]' \mu_x(da).
\end{aligned}$$

[9] Suppressing x and θ_0, M1- M6 ensure that $\partial_a g$ is a square integrable function and, by Taylor's theorem in Banach space, $D_g\rho(a)[\partial_a \widehat{g} - \partial_a g]$ is a smooth linear functional of $\partial_a \widehat{g} - \partial_a g$ for each $a \in int(A)$. Then, by Riesz representation theorem, there is a unique square integrable function $\gamma(\cdot;a)$, defined on A, so that $D_g\rho(a)[\partial_a \widehat{g} - \partial_a g] = \int \gamma(a';a)[\partial_a \widehat{g}(a') - \partial_a g(a')]dF_\alpha(a')$ for some measure F_α. Since $\partial_a \widehat{g}$ is a kernel estimator that is well approximated by a closed form expression, it follows from standard change of variables and integration by parts that the integral will lead to an average of mean zero terms and a smaller order term. See Chen, Linton and van Keilegom (2003) for some explicit examples that make use of this argument.

Next theorem provides the pointwise distribution theory of $\widehat{g}_{\widehat{\theta}}$. Denote $\int u^j K(u)\,du$ and $\int K^j(u)\,du$ by $\mu_j(K)$ and $\kappa_j(K)$ respectively:

THEOREM 2: *Under M1-M7: For any $a \in int(A)$ and $x \in X$, if $\widehat{g}_\theta$ satisfies (13) and $\widehat{\theta}$ satisfies (17) then*

$$\sqrt{Nh}\left(\widehat{g}_{\widehat{\theta}}(a,x) - g_{\theta_0}(a,x) - B_N(a,x;m_{\theta_0})\right) \Rightarrow \mathcal{N}\left(0, \frac{\kappa_2(K)}{Tp_{X,A}(x,a)}\mathrm{var}\left(m_{\theta_0}(x_{it+1})\,|x_{it}=x, a_{it}=a\right)\right),$$

where

$$B_N(a,x;m) = \frac{1}{4!}h^4\mu_4(K)\sum_{x'\in X} m(x')\left(\frac{\frac{\partial^4}{\partial a^4}p_{X',X,A}(x',x,a)}{p_{X,A}(x,a)} + \frac{p_{X',X,A}(x',x,a)\frac{\partial^4}{\partial a^4}p_{X,A}(x,a)}{p^2_{X,A}(x,a)}\right),$$

furthermore, $\widehat{g}_{\widehat{\theta}}(a,x)$ and $\widehat{g}_{\widehat{\theta}}(a',x')$ are asymptotically independent when $a' \neq a$ or $x' \neq x$.

The pointwise asymptotic property of $\widehat{g}_{\widehat{\theta}}(a,x)$ in Theorem 1 is identical to that of a Nadaraya-Watson estimator of the regression $E[m_{\theta_0}(x_{it+1})\,|x_{it}=x, a_{it}=a]$ when m_{θ_0} is known. In other words, the nonparametric estimation of m_θ, as well as the generation of the nonparametric residuals (10), does not affect the first order asymptotic of $\widehat{g}_\theta$. The reason behind this is due to the fact that $\left(\widetilde{r}_\theta, \widehat{m}_\theta, \widehat{\mathcal{L}}\right)$ converges uniformly (over $\Theta \times X$) in probability to $(r_\theta, m_\theta, \mathcal{L})$ at the rate close to $N^{-1/2}$, which is much faster than the nonparametric rate (cf. Theorems 4 and 5 of Srisuma and Linton (2009)).

The asymptotic variance of the finite dimensional estimator in semiparametric models can have a complicated form that generally is a functional of the infinite dimensional parameters and their derivatives. Not only it is difficult to estimate such object, the estimate often works poorly in finite sample. The semiparametric bootstrap seems to be a natural resampling method since we know the DGP for the controlled processes up to an estimation error. In the parametric setting, Kasahara and Shimotsu (2008a) develops a bootstrap procedure for discrete Markov decision models, where they use parametric bootstrap framework of Andrews (2005).

4 Numerical Example

In this section we illustrate some finite sample properties of our proposed estimator in a small scale Monte Carlo experiment. We consider a dynamic price setting problem for a representative firm described in Section 2 with the following specification.

DESIGN:

For some unknown $\theta_0 = (\theta_{01}, \theta_{02})$, each firm faces the following demand

$$D(a_{it}, x_{it}, \varepsilon_{it}) = \overline{D} - \theta_{01} a_{it} + \theta_{02}(x_{it} + \varepsilon_{it}).$$

The observable state x_{it} takes value either 1 or -1, where 1 signifies an increase in demand towards the firm's product and vice versa; the firm's private shock in the demand ε_{it} is not observed to the econometrician. $\overline{D}$ can be interpreted as the upper bound of the supply and θ_0 are the parameters representing the market elasticities. Unlike x_{it}, the evolution of the private shocks are completely random and transitory. The distribution of the consumer satisfaction measure depends on the previous period's price set by the firm, which is summarized by $\Pr[x_{t+1} = -1 | x_{it}, a_{it}] = \frac{a_{it} - \underline{a}}{\overline{a} - \underline{a}}$, where $\underline{a}$ and $\overline{a}$ denote the minimum and maximum possible prices respectively. It is a simple exercise to show that the policy function can be characterized in terms of the conditional value function $E[V(s_{it+1}) | x_{it}]$, in particular, the firm's optimal pricing strategy has the following explicit form

$$\alpha(x, \varepsilon) = \left(\overline{D} + \theta_{02}(x + \varepsilon) + c\theta_{01} - \beta \frac{\lambda_1 - \lambda_2}{\overline{a} - \underline{a}}\right) / 2\theta_{01}, \tag{18}$$

where $\lambda_1 = E[V(s_{it}) | x_{it} = 1]$ and $\lambda_2 = E[V(s_{it}) | x_t = -1]$. It follows that $D(a, x, \varepsilon)(a - c)$ is supermodular in (a, ε_i) for any x if θ_{01} and θ_{02} are positive, which ensures the policy function above will then be strictly increasing in ε. If we ignore that the firm is forward looking, the optimal static profit can be obtained from the following pricing policy

$$\alpha_s(x, \varepsilon) = \left(\overline{D} + \theta_{02}(x + \varepsilon) + c\theta_{01}\right) / 2\theta_{01}. \tag{19}$$

Intuitively, we expect firms which do not take into the account of the consumer's adverse response to high prices will overprice relative to their forward looking counterparts. This is confirmed by the expressions in the displays above since we expect $\lambda_1 - \lambda_2$ (and θ_{01}) to be positive, i.e. the latter implies $\alpha_s(x, \varepsilon) > \alpha(x, \varepsilon)$ for any pair of (x, ε).

In our design, we assign the following values to the parameters

$$\overline{D} = 3, \theta_{01} = 1, \theta_{02} = 1/2, c = 1,$$

and let $\varepsilon_{it} \sim Uni[-1, 1]$. Then it can be shown that $\overline{a} - \underline{a} = 1$ and the stochastic matrix $\mathcal{L}$ is symmetric, in particular

$$\begin{aligned} \mathcal{L} &= \beta \begin{pmatrix} \Pr[x_{it+1} = 1 | x_{it} = 1] & \Pr[x_{it+1} = -1 | x_{it} = 1] \\ \Pr[x_{it+1} = 1 | x_{it} = -1] & \Pr[x_{it+1} = -1 | x_{it} = -1] \end{pmatrix} \\ &= \beta \begin{pmatrix} 0.25 & 0.75 \\ 0.75 & 0.25 \end{pmatrix}. \end{aligned}$$

A numerical method that mirrors our estimation of the policy value equation in Section 2 can be used to calculate the exact value of $\lambda_1 - \lambda_2 \,(= 1/1.45)$. These information allow us to simulate the controlled Markov processes that are consistent with optimal pricing behavior in (18). We generate 1000 replications of such controlled Markov processes with for various sizes of $N \in \{20, 100, 200\}$ random samples of decision series over 5 time periods; this leads to five sets of experiments with the total sample size, NT, of $100, 500$ and 1000.

Implementation:

We are interested in obtain estimates for the demand parameters θ_0 and assume the knowledge of $\left(\overline{D}, c\right)$. In estimating the nonparametric estimator of g_θ, we construct a truncated 4-th order kernel based on the density of a standard normal random variable, see Rao (1983). For each replication, we experiment with the 3 different bandwidths $\left\{h_\varsigma = 1.06s\,(NT)^{-\varsigma} : \varsigma = \frac{1}{6}, \frac{1}{7}, \frac{1}{8}\right\}$; the rates of decay for h_1 and h_3 lie on the boundary off the admissible range stated in M5. We trim out the calculations involving $\widehat{g}_\theta\,(a, x)$ when a that lies within a bandwidth neighborhood of the boundary. For the simulation of $F_{A|X}\,(\cdot;\theta)$, we take $R = N \log(N)$ random draws from Q. We approximate the model implied policy function by grid-search. We use two sets of measures $\left\{\mu_x^{UM}\right\}$ and $\left\{\mu_x^{EM}\right\}$ to construct the objective functions defined in (16) that lead to $\widehat{\theta}^{UM}$ and $\widehat{\theta}^{EM}$ respectively. The uniform measures $\left\{\mu_x^{UM}\right\}$ simply put equal weights on all a and x, whilst $\left\{\mu_x^{EM}\right\}$ are the empirical measures. For a comparison, we also compute the structural estimators for the static model, which is much simpler to estimate since the policy function in such framework has a closed form as described in (19).

Results:

We report the bias, median of the bias, standard deviation, interquartile range (scaled by 1.349) and the mean square error for the estimators in Table 1 and Table 2 that can be found at the end of the paper; the tables give the results for the estimators of θ_{01} and θ_{02} respectively. The rows are arranged according to the total sample size and bandwidths. We have the following general observations for both of our estimators regardless across all bandwidths and measures: (i) the median of the bias is similar to the mean; (ii) the estimators converge to the true values as N increases as their respective mean and standard deviations (and MSE) are converging to zero; (iii) the standard deviation figures are similar to the corresponding scaled interquartile range.[10] There does not seem to be any significant difference in the performance of the estimators that are defined using different measures. We also report analogous summary statistics when the model is wrongly assumed to be static, they can also be found in Table 1 and 2 in the rows labelled *static*. It is clear that the estimators under static environment do not converge to θ_0, instead they display a significant positive bias. This finding is very intuitive, since our minimum distance estimators reflect the model that

[10] (iii) is a characteristic of a normal random variable.

best fit the observed data, the upwards bias of the elasticity parameters estimates is highly plausible. To see this, first recall from (19) that firms who do not take into the account of the future dynamics will overprice relative to the forward looking firms. The firms that only maximize their static profits will therefore, on average, need to expect the market elasticities to be more sensitive in order to generate more conservative pricing schemes consistent with the behaviors of their forward looking counterparts.

5 Other Dynamic Models

The estimation methodology discussed in Section 2 can be adapted to estimate different classes of dynamic problems. Notice that, once we can provide a consistent estimator for the conditional value functions, the model implied distribution function can be readily obtained and used to construct a variety of objective functions to estimate the finite dimensional parameter of interest. Therefore we only discuss how to obtain consistent estimators for the conditional value function defined as the solution to some policy value equation, analogous to the solution to equation (8), under different settings.

Discrete and Continuous Controls:

An important contribution of BBL is to be able to estimate models that have discrete as well as continuous control variables without solving the model equilibrium. This class of problems is useful for many empirical models, for instance where the economic agents endogenously choose whether to participate in the market before deciding on the price or investment decisions, see the dynamic oligopoly studies of Ryan (2009) and Santos (2009).

Our modeling framework here is similar to Section 4 of Arcidiacono and Miller (2008). For each economic agent, the model now consists of the control variables $(a_{it}, d_{it}) \in A \times D$, where $A \subset \mathbb{R}$ and $D = \{1, \ldots, K\}$, and the state variables $s_{it} = \left(x_{it}, \varepsilon_{it}, v_{it}^K\right) \in X \times \mathcal{E} \times \mathcal{V}^K$, where $X = \{1, \ldots, J\}, \mathcal{E} \subset \mathbb{R}$ and $\mathcal{V}^K \subset \mathbb{R}^K$ so $v_{it}^K = (v_{it}(1), \ldots, v_{it}(K))$. In each period the economic agent makes a sequential decision, first on the discrete choice then the continuous one. The decisions made within and across period generally are allowed to affect the consequential state variables. The decision problem leads to the following pair of policy functions

$$\begin{aligned} \delta\left(s_{it}^D\right) &= \arg\max_{1 \le d \le K} \left\{ E\left[u\left(\alpha\left(s_{it}^C\right), d_{it}, x_{it}, \varepsilon_{it}, v_{it}^K\right) | x_{it}, d_{it} = d\right] + \beta E\left[V\left(s_{it+1}\right) | x_{it}, d_{it} = d\right] \right\}, \\ \alpha\left(s_{it}^C\right) &= \arg\max_{a \in A} \left\{ u\left(a, d_{it}, x_{it}, \varepsilon_{it}, v_{it}^K\right) + \beta E\left[V\left(s_{it+1}\right) | x_{it}, a_{it} = a, d_{it}\right] \right\}, \end{aligned}$$

where $s_{it}^D = \left(x_{it}, v_{it}^K\right)$ and $s_{it}^C = \left(x_{it}, \varepsilon_{it}, v_{it}^K, d_{it}\right)$.

For any payoff function u that is parameterized by some $\theta \in \Theta$, we can define a policy value function that is a stationary solution to the following linear equation (cf. equation (7))

$$V_\theta(s) = u_\theta\left(\alpha\left(s_{it}^C\right), \delta\left(s_{it}^D\right), x_{it}, \varepsilon_{it}, v_{it}^K\right) + \beta E\left[V_\theta(s_{it+1}) | s_{it} = s\right] \text{ for any } s \in S.$$

Simplify the notation by replacing $\left(\alpha\left(s_{it}^C\right), \delta\left(s_{it}^D\right)\right)$ with (a_{it}, d_{it}), under a conditional independence assumption analogous to M1, the conditional value function is then the solution to this matrix eqation (cf. equation (8))

$$E\left[V_\theta(s_{it}) | x_{it}\right] = E\left[u_\theta\left(a_{it}, d_{it}, x_{it}, \varepsilon_{it}, v_{it}^K\right) | x_{it}\right] + \beta E\left[E\left[V_\theta(s_{it+1}) | x_{it+1}\right] | x_{it}\right].$$

The matrix equation (and its solution) is identified under some familiar assumptions. In particular, let $u_\theta\left(a, d, x, \varepsilon, v^K\right) = u_\theta^C(a, d, x, \varepsilon) + v(d)$ and u_θ^C has increasing differences in (a, ε) for all d, x and θ, then

$$\begin{aligned} E\left[u_\theta\left(a_{it}, d_{it}, x_{it}, \varepsilon_{it}, v_{it}^K\right) | x_{it}\right] &= E\left[u_\theta^C(a_{it}, d_{it}, x_{it}, \varepsilon_{it}) | x_{it}\right] + E\left[v_{it}(d_{it}) | x_{it}\right], \\ &= \sum_d \Pr\left[d_{it} = d | x_{it}\right] E\left[u_\theta^C(a_{it}, d_{it}, x_{it}, \varepsilon_{it}) | x_{it}, d_{it} = d\right] \\ &\quad + \sum_d \Pr\left[d_{it} = d | x_{it}\right] E\left[v_{it}(d_{it}) | x_{it}, d_{it} = d\right]. \end{aligned}$$

For instance, the conditional choice probabilities can be estimated nonparametrically using the frequency estimator. Although we do not observe $\{\varepsilon_{it}\}$, under analogous conditions to M1 and M3, they can be generated by the relation $\widehat{\varepsilon}_{it} = Q^{-1}\left(\widehat{F}_{A|X,D}(a_{it}|x_{it}, d_{it})\right)$, where $\widehat{F}_{A|X,D}(a|x, d)$ is the nonparametric estimator for $\Pr\left[a_{it} \leq a | x_{it} = x, d_{it} = d\right]$. The selectivity term, $E\left[v_{it}(d_{it}) | x_{it}, d_{it} = d\right]$, that arises from the discrete choice problem can be estimated using Hotz and Miller's inversion theorem as in a purely discrete choice problem. The transition laws and conditional distributions of the observables are nonparametrically identified under some regularity conditions. The estimates of $E\left[V_\theta(s_{it}) | x_{it}\right]$ can then be use to estimate the choice specific conditional value function, which can once again be written as a conditional expectation of the observables, namely $E\left[V_\theta(s_{it+1}) | x_{it}, a_{it}, d_{it}\right] = E\left[E\left[V_\theta(s_{it+1}) | x_{it+1}\right] | x_{it}, a_{it}, d_{it}\right]$.

Ordered Discrete Response:

A dynamic problem with ordered choice is precisely the discrete counterpart of the continuous control framework. Practical application includes investment models where firms purchase or rent goods in discrete units, e.g. see Gowrisankaran et al. (2010).

In this case, the support of a_{it}, A, is an ordered set $\left\{a^1, \ldots, a^K\right\}$. The policy function that solves the dynamic discrete ordered choice problem satisfies

$$\alpha(s_{it}) = \arg\max_{1 \leq k \leq K} \left\{u\left(a^k, s_{it}\right) + \beta E\left[V(s_{it+1}) | s_{it}, a_{it} = a^k\right]\right\}.$$

The conditional policy value equation can be generated for each θ, analogous to the matrix equation (8). We are only concerned with consistent estimation of $E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}\right]$, where

$$E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}\right] = \sum_{k=1}^{K} \Pr\left[a_{it} = a^k | x_{it}\right] E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}, a_{it} = a^k\right].$$

The conditional choice probabilities are nonparametrically identified under some regularity conditions. The per period payoff function contains the unobserved state variable ε_{it}. Analogous to the case with continuous control case, the policy function is weakly monotone in ε_{it} under conditional independence and increasing differences assumptions. We can utilize the quantile invariance property between a_{it} and ε_{it} to identify the choice specific expected payoff. In particular, for $k > 1$ let $\mathcal{I}_k = \left[Q^{-1}\left(F_{A|X}\left(a^{k-1}|x_{it}\right)\right), Q^{-1}\left(F_{A|X}\left(a^k|x_{it}\right)\right)\right]$, we have

$$\begin{aligned} E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}, a_{it} = a^k\right] &= E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}, \varepsilon_{it} \in \mathcal{I}_k\right] \\ &= \frac{\int_{\mathcal{I}_k} u_\theta\left(a^k, x_{it}, \varepsilon\right) Q\left(d\varepsilon\right)}{F_{A|X}\left(a^k|x_{it}\right) - F_{A|X}\left(a^{k-1}|x_{it}\right)}, \end{aligned}$$

when $k = 1$ let $\mathcal{I}_1 = \left[Q^{-1}\left(0\right), Q^{-1}\left(F_{A|X}\left(a^1|x_{it}\right)\right)\right]$, and

$$E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}, a_{it} = a^1\right] = \frac{\int_{\mathcal{I}_1} u_\theta\left(a^1, x_{it}, \varepsilon\right) Q\left(d\varepsilon\right)}{F_{A|X}\left(a^1|x_{it}\right)}.$$

We can consistently estimate $\{\mathcal{I}_k\}$ from the nonparametric estimators $\left\{F_{A|X}\left(a^k|x_{it}\right)\right\}$ and the assumed distribution function Q. Therefore we can estimate $\int_{\mathcal{I}_k} u_\theta\left(a^k, x_{it}, \varepsilon\right) Q\left(d\varepsilon\right)$ for any k. The conditional policy value function can be estimated and solved. The feasible model implied policy and conditional distribution functions can then be used to construct a minimum distance criterion as before.

Mixed Discrete-Continuous Control:

The control variable in many investment and pricing problems naturally have both continuous and discrete components. As in the empirical example in Hong and Shum (2009), firms may choose to not invest with positive probability, or for a pricing problem, price may be regulated to lie without certain bounds (which is binding). For simplicity, suppose that $A = \{0\} \cup (\underline{a}, \overline{a}]$, where $0 \leq \underline{a} < \overline{a} \leq \infty$.

The policy function for this class of controlled problem satisfies

$$\alpha\left(s_{it}\right) = \alpha^C\left(s_{it}\right) \mathbf{1}\left[\begin{array}{c} u\left(0, s_{it}\right) + \beta E\left[V\left(s_{it+1}\right)|s_{it}, a_{it} = 0\right] \\ < u\left(\alpha^C\left(s_{it}\right), s_{it}\right) + \beta E\left[V\left(s_{it+1}\right)|s_{it}, a_{it} = \alpha^C\left(s_{it}\right)\right] \end{array}\right],$$

where $\alpha^C\left(s_{it}\right) = \arg\max_{a \in [\underline{a}, \overline{a}]} \left\{u\left(a, s_{it}\right) + \beta E\left[V\left(s_{it+1}\right)|s_{it}, a_{it} = a\right]\right\}$. As seen in the previous cases, the main feature that distinguishes the conditional policy equations (see equation (8)) for different type of control variables is the conditional mean of the payoff function $E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}\right]$. In this setting, we can write

$$\begin{aligned} E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}\right] &= \Pr\left[a_{it}=0|x_{it}\right] E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}, a_{it}=0\right] \\ &\quad + \Pr\left[a_{it}>0|x_{it}\right] E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}, a_{it}>0\right]. \end{aligned}$$

Again, under the conditional independence and increasing differences assumptions we rely on the monotonicity condition of the policy function. For the discrete part, similar to the ordered choice case, we can estimate $E\left[u_\theta\left(a_{it}, s_{it}\right)|x_{it}, a_{it}=0\right]$ by $\int_{\mathcal{I}} u_\theta\left(0, x_{it}, \varepsilon\right) Q\left(d\varepsilon\right) / \Pr\left[a_{it}=0|x_{it}\right]$, where $\mathcal{I} = \left[Q^{-1}\left(0\right), Q^{-1}\left(\Pr\left[a_{it}=0|x_{it}\right]\right)\right]$. For the continuous contribution, we can again generate ε_{it} that corresponds to positive values of a_{it} by same relation in Equation (10). Given the monotonicity framework, it is straightforward to allow for a control variable that has multiple mass points.

Markovian Games:

The development of empirical methods to analyze of dynamic games has been growing in the empirical industrial organization literature, we refer to Ackerberg, Benkard, Berry and Pakes (2005) and Aguirregabiria and Mira (2008) for recent surveys. To avoid repetition, we illustrate that our methodology can be extended to estimate a class of Markovian games when the action variable is purely continuous. We consider the independent private value framework similar to Aguirregabiria and Mira (2007), BBL, Pesendorfer and Schmidt-Dengler (2008).

For each period t, there are N players, indexed by the ordered set $\{i\}$. Assuming that the game has a unique equilibrium so that each player i's optimal strategy satisfies

$$\alpha_i\left(s_{it}\right) = \arg\max_{a_i \in A_i} \left\{ \begin{array}{c} E\left[u\left(a_{it}, \boldsymbol{\alpha}_{-i}\left(s_{-it}\right), s_{it}\right)|s_{it}, a_{it}=a_i\right] \\ +\beta_i E\left[V_i\left(s_{it+1}; \boldsymbol{\alpha}_{-i}\right)|s_{it}, a_{it}=a_i\right] \end{array} \right\},$$

where $\boldsymbol{\alpha}_{-i}$ and s_{-it} denotes the strategy profile and information of all players other than player i respectively, and

$$V_i\left(s_{it}; \boldsymbol{\alpha}_{-i}\right) = \max_{a_i \in A_i} \left\{ \begin{array}{c} E\left[u_i\left(a_{it}, \boldsymbol{\alpha}_{-i}\left(s_{-it}\right), s_{it}\right)|s_{it}, a_{it}=a_i\right] \\ +\beta_i E\left[V_i\left(s_{it+1}; \boldsymbol{\alpha}_{-i}\right)|s_{it}, a_{it}=a_i\right] \end{array} \right\}.$$

Let $s_{it} = \left(x_{it}, \varepsilon_{it}\right)$ denote the public and private information for each player i. We can define a family of policy value functions $\{V_{\theta,i}\}$ that is generated from using the players equilibrium strategy profile $\{\alpha_i\}$. So that, for each θ and i, the conditional policy value function is the solution to the following matrix equation

$$E\left[V_{\theta,i}\left(s_{it}\right)|x_{it}\right] = E\left[u_{\theta,i}\left(a_{it}, \mathbf{a}_{-it}, s_{it}\right)|x_{it}\right] + \beta_i E\left[E\left[V_{\theta,i}\left(s_{it+1}\right)|x_{it+1}\right]|x_{it}\right].$$

The linear equation above is the continuous action counterpart of Equation (6) in Pesendorfer and Schmidt-Dengler (2008). Once again, under the monotonicity assumption, we can generate $\widehat{\varepsilon}_{it}$ from $Q_i^{-1}\left(\widehat{F}_{A_i|X_i}\left(a_{it}|x_{it}\right)\right)$, so the matrix equation has to be estimated and solved N times, once for each player i. The model implied strategy/policy functions and conditional distribution of the actions can then be estimated for each players, which can then be used to construct the minimum distance estimator for dynamic games.

6 Conclusion

In this paper we develop a new two-step estimator for a class of Markov decision processes with continuous control that forms a basis to estimate a larger class of structural dynamic models. Our criterion function has a simple interpretation and is also simple to construct; we minimize a minimum distance criterion that measures the divergence between two estimators of the conditional distribution function of the observables. In particular, we compare the conditional distribution functions, one implied purely by the data with another constructed from the structural model. We provide some conditions to ensure our estimator is consistent for the structure parameter of interest. As an alternative estimator to BBL, which is designed to estimate the same class of models without having to solve for the dynamic programming problem, for a parametric model we can simply use the empirical measure to construct our objective functions hence there is no additional decisions to be made by the practitioners (e.g. choosing classes of inequalities). We also explicitly work with the framework where we do not need to impose any distributional assumptions on the transition law of the observables. This additional flexibility is important since the transition law is a model primitive. We provide the distribution theory of both the finite dimensional parameters as well as the conditional value functions. We illustrate the performance of our estimator in a Monte Carlo experiment on a dynamic pricing problem, where our criterion function provides a simple intuition for the direction of the biased of the estimator which ignore the model dynamics. We also demonstrate how the general approach we take to estimate dynamic models with continuous control, analogous to the discrete choice counterparts proposed by Hotz and Miller (1993), can easily be adapted to estimate other class of interesting and practically relevant dynamic models.

There are also other important aspects of dynamic models we do not discuss in this paper. We end with a brief note of two issues that are particularly relevant to our framework. The first is regarding unobserved heterogeneity. The absence of unobserved heterogeneity has long been the main criticism against two-step approaches developed along the line of HM. Recently, finite mixtures have been used to add unobserved components in related two-step estimation methodologies, for example see Aguirregabiria and Mira (2007) and Arcidiacono and Miller (2008), Kasahara and Shimotsu

(2008a,b). Finite mixture models can also be used with the estimator developed in this paper. Secondly, our paper focuses on estimation and assumes the model can be identified from some moment conditions. There are ongoing research on the nonparametric and semiparametric identification of dynamic decision models of single and multiple agents, for some recent examples, we refer interested readers to Aguirregabiria (2008), Bajari et al. (2009), Heckman and Navarro (2007), Hu and Shum (2009) and Pesendorfer and Schmidt-Dengler (2008).

7 Appendix

We first make an explicit assumption on the choice of norms. In a general semiparametric problem, the parameter space $\Theta \times \mathcal{G}$ is the Cartesian product of $\Theta \subset \mathbb{R}^M$ and $\mathcal{G}$, a Banach space of $J-$vector-valued functions defined with the supremum norm on A. Since the policy values are indexed by θ, a generic element in g can be written explicitly as $g(\cdot,\theta) = (g_x(\cdot,\theta))_{x=1}^J$, so that $\|g\|_{\mathcal{G}} = \max_{1\leq x\leq J}\sup_{\theta\in\Theta}\|g_x(\cdot,\theta)\|_\infty$, where $\|\cdot\|_\infty$ denotes the usual sup-norm. Since $\mathcal{G}$ is a Cartesian product of J spaces of functions $\times_{x\in X}\mathcal{G}_x$, the norm $\|\cdot\|_{\mathcal{G}_x}$ are defined for a particular element of the $J-$vector in $\mathcal{G}$ accordingly. The norm on $\Theta\times\mathcal{G}$ is defined additively with respect to the finite and infinite dimensional components, namely $\|(\theta,g)\|_{\Theta\times\mathcal{G}} = \|\theta\| + \|g\|_{\mathcal{G}}$, where the norm on the finite dimensional vector space is the usual matrix norm; the norms for other Cartesian products of any subset of $\mathbb{R}^D$ (e.g. $A\times\Theta$) and $\mathcal{G}$ are defined similarly. In relation to our problem, the space of functions of interest contains the vector of expected policy value functions g_θ, defined in Equation (9), we henceforth denote this by $g_0(\cdot,\theta)$ for any θ. We are also interested in the space of functions of derivatives of the expected value functions, we denote this by $\mathcal{G}^{(1)}$, where it is equipped with an analogous sup-norm as $\mathcal{G}$; the norm $\|\cdot\|_{\Theta\times\mathcal{G}^{(1)}}$ are be defined on $\Theta\times\mathcal{G}^{(1)}$ accordingly.

We have used several abbreviations to simplify the notations in the main text. Throughout the appendix, we denote the model implied objective policy functions $\alpha_\theta(x,\varepsilon)$, defined in Equation (5), and its inverse by $\alpha_x(\varepsilon,\theta,\partial_a g_x(\cdot,\theta))$ and $\rho_x(a,\theta,\partial_a g_x(\cdot,\theta))$, for any $a,x,\varepsilon,\theta,g_x$, respectively. The explicit dependence on the derivative of the function g follows from implicit function theorem in Banach space as we assume that the policy function is characterized as the solution to the first order condition of the objective function on the RHS of Equation (5). We denote the corresponding conditional distribution functions $F_{A|X}(a|x;\theta)$, defined in Equation (6), by $F_{A|X=x}(a,\theta,\partial_a g_x(\cdot,\theta))$ for any a,x,θ,g_x. We also use the short-hand notation $F_{A|X=x}(\theta,\partial_a g_x(\cdot,\theta))$ and $F_{A|X=x}$ for the functions $F_{A|X=x}(\cdot,\theta,\partial_a g_x(\cdot,\theta))$ and $F_{A|X=x}(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0))$ respectively. Lastly, we denote the feasible objective function $M_N(\theta)$, defined in Equation (16), by $\widehat{M}_N(\theta,\widehat{g}(\cdot,\theta))$ for any θ.

We let VW abbreviates van der Vaart and Wellner (1996). We also use the multi-index notation for the higher order derivatives with respect to a and θ, in particular $\partial_\theta^{|\eta|} = \partial^{|\eta|}/\partial\theta_1^{\eta_1}\ldots\partial\theta_M^{\eta_M}$ when $|\eta| = \sum_{l=1}^M \eta_l$ for any given natural number $|\eta|$ and any combination $\{\eta_l\}$. In what follows we let: $\xi > 0$ be a number that is arbitrarily close to 0; C_0 denotes a positive constant that may take different values in various places.

7.1 Theorem G and Some Lemmas

We first present a general theorem that will be useful to obtain the asymptotic normality of our estimator. We denote the (partial-) Fréchet differential operators by $D_\theta, D_g, D_{\theta\theta}, D_{\theta g}$ and D_{gg}, where the indices denote the argument(s) used in differentiating and double indexing denotes second derivative. For any map $T : X \to Y$ and some Banach spaces X and Y, we say that T is Fréchet differentiable at x, that belongs to some open neighborhood of X, if and only if there exists a linear bounded map $D_T : X \to Y$ such that $T(x+f) - T(x) = D_T(x) f + o(\|f\|)$ with $\|f\| \to 0$ for all f in some neighborhood of x; we denote the Fréchet differential at x in a particular direction f by $D_T(x)[f]$. For Theorem G below, let θ_0 and g_0 denote the true finite and infinite dimensional parameters that lie in Θ and $\mathcal{G}$ respectively. Since we only need to focus on the local behavior around (θ_0, g_0), for any $\delta > 0$ we define $\Theta_\delta = \{\theta \in \Theta : \|\theta - \theta_0\| < \delta\}$ and $\mathcal{G}_\delta = \left\{g \in \mathcal{G} : \|g - g_0\|_{\mathcal{G}} < \delta\right\}$, here δ can also be replaced by some positive sequence $\delta_N = o(1)$. The pseudo-norm on $\mathcal{G}_\delta$ can be suitably modified to reflect the smaller parameter space Θ_δ, and the choice of δ for Θ_δ and $\mathcal{G}_\delta$ can be distinct, but for notational simplicity we ignore this. Let $M(\theta, g(\cdot,\theta))$ denote the population objective function that is minimized at $\theta = \theta_0$, and $M_N(\theta, g(\cdot,\theta))$ denote the sample counterpart so that $M(\theta, g(\cdot,\theta)) = \lim_{N\to\infty} E[M_N(\theta, g(\cdot,\theta))]$. Further, we denote $D_\theta M(\theta, g(\cdot,\theta))$ by $S(\theta, g(\cdot,\theta))$ and $D_{\theta\theta} M(\theta, g(\cdot,\theta))$ by $H(\theta, g(\cdot,\theta))$.

Theorem G: *Suppose that* $\widehat{\theta} \xrightarrow{p} \theta_0$, *and for some positive sequence* $\delta_N = o(1)$,

G1 $M_N\left(\widehat{\theta}, \widehat{g}\left(\cdot, \widehat{\theta}\right)\right) \leq \inf_{\theta\in\Theta} M_N(\theta, \widehat{g}(\cdot,\theta)) + o_p(N^{-1})$

G2 For all θ, $\widehat{g}(\cdot,\theta) \in \mathcal{G}_{\delta_N}$ *w.p.a. 1 and* $\sup_{\theta\in\Theta} \|\widehat{g}(\cdot,\theta) - g_0(\cdot,\theta)\|_\infty = o_p\left(N^{-1/4}\right)$

G3 For some $\delta > 0$, $M(\theta, g)$ *is twice continuously differentiable in* θ *at* θ_0 *for all* $g \in \mathcal{G}_\delta$. $H(\theta, g)$ *is continuous in* g *at* g_0 *for* $\theta \in \Theta_\delta$. *Further,* $S(\theta_0, g_0(\cdot,\theta_0)) = 0$ *and* $H_0 = H(\theta_0, g_0(\cdot,\theta_0))$ *is positive definite.*

G4 For some $\delta > 0$, $S(\theta, g(\cdot,\theta))$ *is (partial-) Fréchet differentiable with respect to* g, *for any* $\theta \in \Theta_\delta$ *and for all* $g \in \mathcal{G}_\delta$. *Further* $\|S(\theta_0, g(\cdot,\theta_0)) - D_g S(\theta_0, g_0(\cdot,\theta_0))[g(\cdot,\theta_0) - g_0(\cdot,\theta_0)]\| \leq B_N \times \sup_{\theta\in\Theta} \|g(\cdot,\theta) - g_0(\cdot,\theta)\|_\infty^2$ *for some* $B_N = O_p(1)$.

G5 (Stochastic Differentiability)

$$\sup_{\|\theta-\theta_0\|<\delta_N} \left| \frac{\mathcal{D}_N(\theta, \widehat{g}(\cdot,\theta))}{1 + \sqrt{N}\|\theta - \theta_0\|} \right| = o_p(1),$$

where there exist some sequence C_N, *so that*

$$\begin{aligned} & \mathcal{D}_N(\theta, \widehat{g}(\cdot,\theta)) \\ = \;& \sqrt{N}\frac{\left[M_N(\theta, \widehat{g}(\cdot,\theta)) - M_N(\theta_0, \widehat{g}(\cdot,\theta_0)) - (M(\theta, \widehat{g}(\cdot,\theta)) - M(\theta_0, \widehat{g}(\cdot,\theta_0))) - (\theta - \theta_0)' C_N\right]}{\|\theta - \theta_0\|}. \end{aligned} \tag{20}$$

G6 For some finite positive definite matrices Ω_0 *and* Ω, *we have the following weak convergence* $\sqrt{N} C_N \Rightarrow N(0, \Omega_0)$ *and* $\sqrt{N} D_N = \sqrt{N}(C_N + D_g S(\theta_0, g_0(\cdot,\theta_0))[\widehat{g} - g_0]) \Rightarrow \mathcal{N}(0, \Omega)$.

Then

$$\sqrt{N}\left(\widehat{\theta} - \theta_0\right) \Rightarrow N\left(0, H_0^{-1}\Omega H_0^{-1}\right).$$

Lemma 1. *Under M1 - M4 and M6:* $\left\|\widehat{\mathcal{L}}-\mathcal{L}\right\|=O_p\left(N^{-1/2}\right)$.

Lemma 2. *Under M1 - M6: For any* $x\in X$, $\widetilde{r}_\theta(x)=r_\theta(x)+\widetilde{r}_\theta^R(x)$ *such that* $\max_{1\le x\le J}\sup_{\theta\in\Theta}\left|\widetilde{r}_\theta^R(x)\right|=o_p\left(N^{-\lambda}\right)$ *for any* $\lambda<1/2$.

Lemma 3. *Under M1 - M6: For any* $x\in X$, $\widehat{m}_\theta(x)=m_\theta(x)+\widehat{m}_\theta^R(x)$ *such that* $\max_{1\le x\le J}\sup_{\theta\in\Theta}\left|\widehat{m}_\theta^R(x)\right|=o_p\left(N^{-\lambda}\right)$ *for any* $\lambda<1/2$.

Lemma 4. *Under M1 - M6: For any* $\theta\in\Theta$, $x\in X$, *and* $a\in A$, $\widehat{g}_x(a,\theta)=g_x(a,\theta)+\widehat{g}_x^B(a,\theta)+\widehat{g}_x^S(a,\theta)+\widehat{g}_x^R(a,\theta)$ *such that*

$$\begin{aligned}\max_{1\le x\le J}\sup_{\theta,a\in\Theta\times A_N}\left|\widehat{g}_x^B(a,\theta)\right| &= O_p\left(h^4\right),\\ \max_{1\le x\le J}\sup_{\theta,a\in\Theta\times A_N}\left|\widehat{g}_x^S(a,\theta)\right| &= o_p\left(\frac{N^\xi}{\sqrt{Nh}}\right),\\ \max_{1\le x\le J}\sup_{\theta,a\in\Theta\times A_N}\left|\widehat{g}_x^R(a,\theta)\right| &= o_p\left(h^4+\frac{N^\xi}{\sqrt{Nh}}\right).\end{aligned}$$

Lemma 5. *Under M1 - M6: For all* $x\in X, \max_{0\le l\le 2,1\le x\le J}\sup_{\theta,a\in\Theta\times A_N}\left|\partial_a^{|l|}\widehat{g}_x(a,\theta)-\partial_a^{|l|}g_{0,x}(a,\theta)\right|=o_p(1)$.

Lemma 6. *Under M1 - M6:* $\max_{0\le l,p\le 2,1\le x\le J}\sup_{\theta,a\in\Theta\times A_N}\left|\partial_a^{|l|}\partial_\theta^{|p|}\widehat{g}_x(a,\theta)-\partial_a^{|l|}\partial_\theta^{|p|}g_{0,x}(a,\theta)\right|=o_p(1)$.

Lemma 7. *Under M1 - M7 for all* $x\in X$, $\mathcal{F}_x=\{\mathbf{1}[\cdot\le\rho_x(a,\theta,\partial_a g_x)]:a\in A,\theta\in\Theta,g_x\in\mathcal{G}\}$ *is a Donsker class.*

Lemma 8. *Under M1 - M7: For any* $x\in X$ *and some positive sequence* $\delta_N=o(1)$ *as* $N\to\infty$

$$\lim_{N\to\infty}\sup_{\substack{(a,\theta,\partial_a g_x)\in A\times\Theta\times\mathcal{G}_x^{(1)},\\ \|(a'-a,\theta'-\theta,\partial_a g_x'-\partial_a g_x)\|<\delta_N}}\left|\begin{array}{l}\frac{1}{N}\sum_{i=1}^N\left(\mathbf{1}\left[\varepsilon_i\le\rho_x\left(a',\theta',\partial_a g_x'\right)\right]-Q\left(\rho_x\left(a',\theta',\partial_a g_x'\right)\right)\right)\\ -\frac{1}{N}\sum_{i=1}^N\left(\mathbf{1}\left[\varepsilon_i\le\rho_x\left(a,\theta,\partial_a g_x\right)\right]-Q\left(\rho_x\left(a,\theta,\partial_a g_x\right)\right)\right)\end{array}\right|=0.$$

Lemma 9 *Under M4: For any* $x\in X$

$$\sqrt{N}\left(\widehat{F}_{A|X=x}-F_{A|X=x}\right)\rightsquigarrow\mathbb{F}_x,$$

where $\mathbb{F}_x$ *is a tight Gaussian process that belongs to* $l^\infty(A)$.

Lemma 10: *Under M1 - M7 For any* $x\in X$

$$\sqrt{N}(F_{A|X=x}(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0))-F_{A|X=x}(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)))\rightsquigarrow\mathbb{G}_x,$$

where $\mathbb{G}_x$ *is a tight Gaussian process that belongs to* $l^\infty(A)$.

7.2 Proofs of Theorems 1 and 2

For the proof of Theorems 1, it is convenient to introduce the following notations: (i) $E_x(\theta,\partial_a g_x(\cdot,\theta))$ where $E_x(\theta,\partial_a g_x(\cdot,\theta))=F_{A|X=x}(\theta,\partial_a g_x(\cdot,\theta))-F_{A|X=x}$; (ii) $E_{N,x}(\theta,\partial_a g_x(\cdot,\theta))=\widetilde{F}_{A|X=x}(\theta,\partial_a g_x(\cdot,\theta))-$

$\widehat{F}_{A|X=x}$, where $\widetilde{F}_{A|X=x}(\theta, \partial_a g_x(\cdot,\theta))$ and $\widehat{F}_{A|X=x}$ are functions defined on A that are the short for $\widetilde{F}_{A|X}(\cdot|x;\theta,\partial_a g_x(\cdot,\theta))$ and $\widehat{F}_{A|X}(\cdot|x)$ respectively. Therefore we can write the infeasible counterpart of the objective function $\widehat{M_N}(\theta, g(\cdot,\theta))$ as

$$M_N(\theta, g(\cdot,\theta)) = \sum_{x\in X}\int_A E^2_{N,x}(\theta,\partial_a g_x(\cdot,\theta))\,d\mu_x,$$

whose limit is

$$M(\theta, g(\cdot,\theta)) = \sum_{x\in X}\int_A E^2_{x}(\theta,\partial_a g_x(\cdot,\theta))\,d\mu_x.$$

In addition, for $x\in X$, let $\nu_{R,x}$ denote the empirical process indexed by $(\theta,\partial_a g_x)\in\Theta\times\mathcal{G}^{(1)}_x$ to be a random element that takes value over A, i.e. $\nu_{R,x}(\theta,\partial_a g_x) = \frac{1}{\sqrt{R}}\sum_{r=1}^R \mathbf{1}\left[\varepsilon_r \le \rho_x(\cdot,\theta,\partial_a g_x)\right] - Q(\rho_x(\cdot,\theta,\partial_a g_x))$.

Proof of Theorem 1. We first show that $\widehat{\theta}$ is a consistent estimator for θ_0. To do this, we show that $M(\theta, g_0(\cdot,\theta))$ has a well separated minimum at θ_0. By assumption M4 we have $M(\theta, g_0(\cdot,\theta)) \ge M(\theta_0, g_0(\cdot,\theta_0))$ for all θ in the compact set Θ with equality only holds for $\theta=\theta_0$. We also have $F_{A|X}(\cdot|x;\theta,\partial_a g_x(\cdot,\theta)) = Q(\rho_x(\cdot,\theta,\partial_a g_0(\cdot,\theta)))$ that is continuous in θ given, this ensures a well-separated minimum. By standard arguments, consistency will now follow if we can show

$$\sup_{\theta\in\Theta}\left|M_N(\theta,\widehat{g}(\cdot,\theta)) - M(\theta,g_0(\cdot,\theta))\right| = o_p(1). \tag{21}$$

By the triangle inequality, we have

$$\begin{aligned}
\left|M_N(\theta,\widehat{g}(\cdot,\theta)) - M(\theta,g_0(\cdot,\theta))\right| \;&\le\; 4\sum_{x\in X}\int\left|\widetilde{F}_{A|X=x}(\theta,\partial_a\widehat{g}_x(\cdot,\theta)) - F_{A|X=x}(\theta,\partial_a\widehat{g}_x(\cdot,\theta))\right|d\mu_x \\
&\quad +4\sum_{x\in X}\int\left|F_{A|X=x}(\theta,\partial_a\widehat{g}_x(\cdot,\theta)) - F_{A|X=x}(\theta,\partial_a g_{0,x}(\cdot,\theta))\right|d\mu_x \\
&\quad +4\sum_{x\in X}\int\left|\widehat{F}_{A|X=x} - F_{A|X=x}\right|d\mu_x \\
&= A_1 + A_2 + A_3.
\end{aligned}$$

For A_1, for each x and any $\eta>0$ we have

$$\begin{aligned}
&\Pr\left[\sup_{\theta\in\Theta}\left|\widetilde{F}_{A|X}(a|x;\theta,\partial_a\widehat{g}(\cdot,\theta)) - F_{A|X}(a|x;\theta,\partial_a\widehat{g}(\cdot,\theta))\right| > \eta\right] \\
\le\;& \Pr\left[\sup_{\theta,a\in\Theta\times A_N}\left|\frac{1}{R}\sum_{r=1}^R \mathbf{1}\left[\varepsilon_r\le\rho_x(a,\theta,\partial_a\widehat{g}_x)\right] - Q(\rho_x(a,\theta,\partial_a\widehat{g}_x))\right| > \eta\right] \\
\le\;& \Pr\left[\sup_{\theta,a,\partial_a g_x\in\Theta\times A_N\times\mathcal{G}^{(1)}_x}\left|\frac{1}{R}\sum_{r=1}^R \mathbf{1}\left[\varepsilon_r\le\rho_x(a,\theta,\partial_a g_x)\right] - Q(\rho_x(a,\theta,\partial_a g_x))\right| > \eta\right] \\
&+\Pr\left[\partial_a\widehat{g}_x(\cdot,\theta)\notin\mathcal{G}^{(1)}_x\right].
\end{aligned}$$

From Lemma 7, $\mathcal{F}_x$ is $Q-$Glivenko-Cantelli by Slutsky's theorem, therefore the first term of the last inequality above converges to zero as $R \to \infty$. By Lemma 6, $\Pr\left[\partial_a \widehat{g}_x(\cdot,\theta) \notin \mathcal{G}_x^{(1)}\right] = o(1)$ hence by finiteness of μ_x it follows that $|A_1| = o_p(1)$ uniformly over Θ. For A_2, for each x we have

$$\begin{aligned}\left|F_{A|X=x}\left(\theta, \partial_a \widehat{g}_x(\cdot,\theta)\right) - F_{A|X=x}\left(\theta, \partial_a g_x(\cdot,\theta)\right)\right| &= \left|Q\left(\rho_x\left(a,\theta,\partial_a \widehat{g}_x(\cdot,\theta)\right)\right) - Q\left(\rho_x\left(a,\theta,\partial_a g_{0,x}(\cdot,\theta)\right)\right)\right| \\ &\leq C_0\left|\rho_x\left(a,\theta,\partial_a \widehat{g}_x(\cdot,\theta)\right) - \left(\rho_x\left(a,\theta,\partial_a g_{0,x}(\cdot,\theta)\right)\right)\right|,\end{aligned}$$

where the inequality follows from the mean value theorem (MVT) and the fact that the derivative of Q is uniformly bounded. Given the smoothness assumption on (ρ_x) in assumption (xiii), by MVT in Banach space $\sup_{a\in A_N}\left|\rho_x\left(a,\theta,\partial_a \widehat{g}_x(\cdot,\theta)\right) - \left(\rho_x\left(a,\theta,\partial_a g_{0,x}(\cdot,\theta)\right)\right)\right| \leq \sup_{\theta,a,\partial_a g_x \in \Theta\times A\times\mathcal{G}_x^{(1)}} \left\|D_{\partial_a g}\rho_x\left(a,\theta,\partial_a g_x\right)\right\|$ $\times \sup_{\theta,a\in\Theta\times A_N}\left|\partial_a \widehat{g}_x(a,\theta) - \partial_a g_{0,x}(a,\theta)\right|$. Since μ_x has zero measure on the boundary of A, by Lemma 5, $\int\left|F_{A|X=x}\left(\theta,\partial_a \widehat{g}_x(\cdot,\theta)\right) - F_{A|X=x}\left(\theta,\partial_a g_{0,x}(\cdot,\theta)\right)\right| d\mu_x \leq C_0 \sup_{\theta,a\in\Theta\times A_N}\left|\partial_a \widehat{g}_x(a,\theta) - \partial_a g_{0,x}(a,\theta)\right| + 2\mu_x(A\backslash A_N) = o_p(1)$. So we also have $|A_2| = o_p(1)$ uniformly over Θ. Lastly for A_3, for each x we write

$$\begin{aligned}\widehat{F}_{A|X}(a|x) - F_{A|X}(a|x) &= \frac{1}{p_X(x)}\left[\widehat{F}_{A,X}(a,x) - F_{A,X}(a,x)\right] \\ &\quad -\frac{\widehat{F}_{A|X}(a|x)}{p_X(x)}\left[\widehat{p}_X(x) - p_X(x)\right],\end{aligned}$$

where $\widehat{F}_{A,X}(a,x) = \frac{1}{NT}\sum_{i=1,t=1}^{N,T}\mathbf{1}\left[a_{it}\leq a, x_{it}=x\right]$, then w.p.a. 1

$$\begin{aligned}&\max_{1\leq x\leq J}\sup_{a\in A}\left|\widehat{F}_{A|X}(a|x) - F_{A|X}(a|x)\right| \\ \leq\ & \frac{1}{\min_{1\leq x\leq J} p_X(x)}\max_{1\leq x\leq J}\sup_{a\in A}\left|\widehat{F}_{A,X}(a,x) - F_{A,X}(a,x)\right| \\ &+\left|\frac{\max_{1\leq x\leq J}\left[\widehat{p}_X(x) - p_X(x)\right]}{\min_{1\leq x\leq J} p_X(x)}\right|.\end{aligned}$$

By Lemma 9, the class of functions $\{\mathbf{1}\left[\cdot\leq a, x_{it}=x\right] - F_{A,X}(\cdot,x) : a\in A\}$ is also a Glivenko-Cantelli class, so: the first term on the RHS of the inequality above converges in probability to zero; the second term also converges in probability to zero since $\widehat{p}_X(x) - p_X(x) = o_p(1)$ for each $x\in X$. Since A_3 is independent of θ, the uniform convergence in (21) holds and consistency follows.

For the asyptotic normality, we set out to show that our assumptions imply we satisfy all the conditions of Theorem G. We have just shown the consistency of our estimator. G1 is the definition of the estimator. For G2, it suffices to show $\partial_a \widehat{g}_x(\cdot,\theta)\in\mathcal{G}_{\delta_N,x}$ w.p.a. 1 and $\sup_{\theta\in\Theta}\left\|\partial_a \widehat{g}_x(\cdot,\theta) - \partial_a g_{0,x}(\cdot,\theta)\right\|_\infty$ $= o_p\left(N^{-1/4}\right)$ for all $x\in X$. The former is implied by Lemma 6, from the proof of Lemma 5, the latter holds if $h^4 + \frac{N^\xi}{\sqrt{Nh^3}} = o\left(N^{-1/4}\right)$, this is certainly the case when h is in the suggested range. G3 and G4 simply requires translating the smoothness we impose in M6 and M7 to satisfy these conditions. Now we show G5, in particular we need to show that

$$
\begin{aligned}
& M_N(\theta, \widehat{g}(\cdot, \theta)) - M_N(\theta_0, \widehat{g}(\cdot, \theta_0)) - (M(\theta, \widehat{g}(\cdot, \theta)) - M(\theta_0, \widehat{g}(\cdot, \theta_0))) - (\theta - \theta_0)' C_N \quad (22) \\
= \; & o_p\left(\|\theta - \theta_0\|^2 + \frac{\|\theta - \theta_0\|}{\sqrt{N}} + \frac{1}{N}\right),
\end{aligned}
$$

holds uniformly for $\|\theta - \theta_0\| < \delta_N$ for some sequence C_N. For any pair $(\theta, \partial_a \widehat{g}_x(\cdot, \theta))$ we can write

$$
\begin{aligned}
E_x^2(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - E_x^2(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0)) = & \left(F_{A|X=x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - F_{A|X=x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))\right) \times \\
& \left(F_{A|X=x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) + F_{A|X=x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0)) - 2F_{A|X=x}\right),
\end{aligned}
$$

and analogously

$$
\begin{aligned}
E_{N,x}^2(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - E_{N,x}^2(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0)) = & \left(\widetilde{F}_{A|X=x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - \widetilde{F}_{A|X=x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))\right) \times \\
& \left(\widetilde{F}_{A|X=x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) + \widetilde{F}_{A|X=x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0)) - 2\widehat{F}_{A|X=x}\right).
\end{aligned}
$$

Combing these we have

$$
\begin{aligned}
& M_N(\theta, \widehat{g}(\cdot, \theta)) - M_N(\theta_0, \widehat{g}(\cdot, \theta_0)) \\
= & \sum_{x \in X} \int \left[\begin{array}{c} \left[\begin{array}{c} R^{-1/2}(\nu_{R,x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - \nu_{R,x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))) \\ + \left(F_{A|X=x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - F_{A|X=x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))\right) \end{array} \right] \\ \times \left[\begin{array}{c} R^{-1/2}(\nu_{R,x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) + \nu_{R,x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))) \\ + \left(F_{A|X=x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) + F_{A|X=x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0)) - 2\widehat{F}_{A|X=x}\right) \end{array} \right] \end{array} \right] d\mu_x \\
= & \sum_{x \in X} \int E_x^2(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - E_x^2(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0)) \, d\mu_x \\
& -2 \sum_{x \in X} \int \left(F_{A|X=x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - F_{A|X=x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))\right) \left(\widehat{F}_{A|X=x} - F_{A|X=x}\right) d\mu_x \\
& + R^{-1/2} \sum_{x \in X} \int \left[\begin{array}{c} [\nu_{R,x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - \nu_{R,x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))] \\ \times \left[\begin{array}{c} F_{A|X=x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) \\ + F_{A|X=x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0)) - 2\widehat{F}_{A|X=x} \end{array} \right] \end{array} \right] d\mu_x \\
& + R^{-1/2} \sum_{x \in X} \int \left[\begin{array}{c} [\nu_{R,x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) + \nu_{R,x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))] \\ \times \left[F_{A|X=x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - F_{A|X=x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))\right] \end{array} \right] d\mu_x \\
& + R^{-1} \sum_{x \in X} \int \left[\begin{array}{c} [\nu_{R,x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) - \nu_{R,x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))] \\ \times [\nu_{R,x}(\theta, \partial_a \widehat{g}_x(\cdot, \theta)) + \nu_{R,x}(\theta_0, \partial_a \widehat{g}_x(\cdot, \theta_0))] \end{array} \right] d\mu_x \\
= & M(\theta, \widehat{g}(\cdot, \theta)) - M(\theta_0, \widehat{g}(\cdot, \theta_0)) + B_1 + B_2 + B_3 + B_4.
\end{aligned}
$$

We now show that, out of $\{B_i\}_{i=1}^4$, B_1 is the leading term that contains C_N in (22), the rest are of smaller stochastic order. Since we are only interested in what happens as $\|\theta - \theta_0\| \to 0$, in what follows, the little 'o' and big 'O' terms will be implicitly assumed to hold with $\|\theta-\theta_0\| \to 0$ and $N \to \infty$.

For B_1 :

By mean value expansion

$$\begin{aligned}
B_1 &= -2(\theta-\theta_0)' \sum_{x\in X} \int D_\theta F_{A|X=x}\left(\overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right)\left(\widehat{F}_{A|X=x} - F_{A|X=x}\right) d\mu_x \\
&= -2(\theta-\theta_0)' \sum_{x\in X} \int D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\left(\widehat{F}_{A|X=x} - F_{A|X=x}\right) d\mu_x \\
&\quad -2(\theta-\theta_0)' \sum_{x\in X} \int \left[\begin{array}{c} \left[D_\theta F_{A|X=x}\left(\overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right) - D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\right] \\ \times \left[\widehat{F}_{A|X=x} - F_{A|X=x}\right] \end{array}\right] d\mu_x \\
&= B_{11} + B_{12},
\end{aligned}$$

where for each $x, \overline{\theta}_x$ is some intermediate (vector) value between θ and θ_0 that corresponds to the MVT w.r.t. the $x-th$ summand. We first show that B_{11} is the leading term that is equal to $(\theta-\theta_0)' C_N$ in (22) and that $\sqrt{N} C_N$ converges to a normal random variable. By Lemma 9 $\sqrt{N}\left(\widehat{F}_{A|X=x} - F_{A|X=x}\right) \rightsquigarrow \mathbb{F}_x$ where $\mathbb{F}_x$ is a tight Gaussian process that belongs to $l^\infty(A)$ for all x, by Slutsky theorem and a similar argument used in the proof of Lemma 9, it is easy to show that $D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right) \sqrt{N}(\widehat{F}_{A|X=x} - F_{A|X=x})$ also converges weakly to a tight Gaussian process. To see the latter, note that for any $\partial_a g_x(\cdot, \theta) \in \mathcal{G}_x^{(1)}$

$$\begin{aligned}
& D_\theta F_{A|X}\left(a|x; \theta, \partial_a g_x(\cdot, \theta)\right) \\
= \; & q\left(\rho_x\left(a, \theta, \partial_a g_x(\cdot, \theta)\right)\right)\left(\partial_\theta \rho_x\left(a, \theta, \partial_a g_x(\cdot, \theta)\right) + D_{\partial_a g} \rho_x\left(a, \theta, \partial_a g_x(\cdot, \theta)\right)\left[\partial_\theta \partial_a g(\cdot, \theta)\right]\right),
\end{aligned}$$

where, ∂_θ denotes the ordinary $L-$dimensional partial derivative, $\partial/\partial\theta$, w.r.t. in the argument θ. This is continuous on A for any x. Now, if we define a linear continuous map $\mathbb{T}_x : l^\infty(A) \to \mathbb{R}$ (w.r.t. sup-norm) so that $\mathbb{T}_x f = \int D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}(\cdot, \theta_0)\right) f d\mu$ for any $f \in l^\infty(A)$ then the map is linear and continuous, the boundedness follows from the observation that $\sup_{a\in A}\left\|D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}(\cdot, \theta_0)\right)\right\| < \infty$. Then, by continuous mapping theorem (CMT)

$$\int D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}(\cdot, \theta_0)\right) \sqrt{N}\left(\widehat{F}_{A|X=x} - F_{A|X=x}\right) d\mu_x \rightsquigarrow \int D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}(\cdot, \theta_0)\right) \mathbb{F}_x d\mu_x.$$

Furthermore, the limit is also Gaussian since we know Gaussianity is preserved for any tight Gaussian process that is transformed by a linear continuous map, see Lemma 3.9.8 of VW. So we let

$$\sqrt{N} C_N = \sum_{x\in X} \int D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}(\cdot, \theta_0)\right) \sqrt{N}\left(\widehat{F}_{A|X=x} - F_{A|X=x}\right) d\mu_x, \tag{23}$$

then $\sqrt{N} C_N$ also converges a Gaussian variable.

For B_{12}, for each x, by Cauchy Schwarz inequality we have

$$
\begin{aligned}
& \left|(\theta-\theta_0)' \int \left(D_\theta F_{A|X=x}\left(\overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right) - D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\right)\left(\widehat{F}_{A|X=x} - F_{A|X=x}\right) d\mu_x\right| \\
\leq & \left[(\theta-\theta_0)' \int \left[\begin{array}{c} \left(D_\theta F_{A|X=x}\left(\overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right) - D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\right) \\ \times\left(D_\theta F_{A|X=x}\left(\overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right) - D_\theta F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\right)' \end{array}\right] d\mu_x (\theta-\theta_0)\right]^{1/2} \\
& \times\left[\int \left[\widehat{F}_{A|X=x} - F_{A|X=x}\right]^2 d\mu_x\right]^{1/2},
\end{aligned}
$$

where for each $x, \overline{\theta}_x$ is some intermediate value between θ_x and $\theta_{0,x}$ that corresponds to the MVT w.r.t. the $x-th$ summand. Let ∂_{θ_l} denotes the $l-$th element of ∂_θ then

$$
\begin{aligned}
& \left|D_{\theta_l} F_{A|X=x}\left(\overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right) - D_{\theta_l} F_{A|X=x}\left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\right| \\
\leq & \left|\begin{array}{c} q\left(\rho_x\left(a, \overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right)\right) \partial_{\theta_l} \rho_x\left(a, \overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right) \\ -q\left(\rho_x\left(a, \theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\right) \partial_{\theta_l} \rho_x\left(a, \theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right) \end{array}\right| \\
& +\left|\begin{array}{c} q\left(\rho_x\left(a, \overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right)\right) D_{\partial_a g} \rho_x\left(a, \overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}\right)\right)\left[\partial_{\theta_l} \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right] \\ -q\left(\rho_x\left(a, \theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\right) D_{\partial_a g} \rho_x\left(a, \theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\left[\partial_{\theta_l} \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right] \end{array}\right|.
\end{aligned}
$$

First note that the terms on the RHS are uniformly bounded, it is easy to see that the terms on the RHS of the inequality are $o(1)$ as $\left\|\overline{\theta}_x - \theta_0\right\| \rightarrow 0$ since $\left(\overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right) \xrightarrow{p} \left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)$ by Lemma 5 and continuity in θ of $\partial_a g_{0,x}(\cdot, \theta)$. Then it will follow by DCT that

$$
\left[(\theta-\theta_0)' \int \left[\begin{array}{c} \left[D_\theta F_{A|X}\left(\overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right) - D_\theta F_{A|X}\left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\right] \\ \times\left[D_\theta F_{A|X}\left(\overline{\theta}_x, \partial_a \widehat{g}_x\left(\cdot, \overline{\theta}_x\right)\right) - D_\theta F_{A|X}\left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\right]' \end{array}\right] d\mu (\theta-\theta_0)\right]^{1/2} = o_p\left(\|\theta-\theta_0\|\right).
$$

From Lemma 9 and CMT, $\left[\int \left[\widehat{F}_{A|X=x} - F_{A|X=x}\right]^2 d\mu\right]^{1/2} = O_p\left(N^{-1/2}\right)$. Since we have finite x then $|B_{12}| = o_p\left(N^{-1/2}\|\theta-\theta_0\|\right)$.

For B_2 :

For each x

$$
\begin{aligned}
& F_{A|X=x}\left(\theta, \partial_a \widehat{g}_x(\cdot, \theta)\right) + F_{A|X=x}\left(\theta_0, \partial_a \widehat{g}_x\left(\cdot, \theta_0\right)\right) - 2\widehat{F}_{A|X=x} \\
= & \left(F_{A|X=x}\left(\theta, \partial_a \widehat{g}_x(\cdot, \theta)\right) - F_{A|X=x}\left(\theta_0, \partial_a \widehat{g}_x\left(\cdot, \theta_0\right)\right)\right) \\
& +2\left(F_{A|X=x}\left(\theta_0, \partial_a \widehat{g}_x\left(\cdot, \theta_0\right)\right) - F_{A|X}\left(\theta_0, \partial_a g_{0,x}\left(\cdot, \theta_0\right)\right)\right) \\
& -2\left(\widehat{F}_{A|X=x} - F_{A|X=x}\right),
\end{aligned}
$$

then we can write B_2 as

$$\begin{aligned}
B_2 &= R^{-1/2}\sum_{x\in X}\int\left[\begin{array}{c}\left[\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x\left(\cdot,\theta\right)\right)-\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right)\right]\\ \times\left[F_{A|X=x}\left(\theta,\partial_a\widehat{g}_x\left(\cdot,\theta\right)\right)-F_{A|X=x}\left(\theta_0,\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right)\right]\end{array}\right]d\mu_x\\
&+2R^{-1/2}\sum_{x\in X}\int\left[\begin{array}{c}\left[\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x\left(\cdot,\theta\right)\right)-\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right)\right]\\ \times\left[F_{A|X=x}\left(\theta_0,\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right)-F_{A|X}\left(\theta_0,\partial_a g_{0,x}\left(\cdot,\theta_0\right)\right)\right]\end{array}\right]d\mu_x\\
&-2R^{-1/2}\sum_{x\in X}\int\left(\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x\left(\cdot,\theta\right)\right)-\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right)\right)\left(\widehat{F}_{A|X=x}-F_{A|X=x}\right)d\mu_x\\
&= B_{21}+B_{22}+B_{23}.
\end{aligned}$$

We first show $\int\left[\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x\left(\cdot,\theta\right)\right)-\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right)\right]^2 d\mu_x = o_p\left(1\right)$ for any x. By Lemma 6 $\partial_a\widehat{g}_x\in\mathcal{G}_x^{(1)}$ w.p.a. 1, and by Lemma 8 it suffices to show that $\left\|\partial_a\widehat{g}_x\left(\cdot,\theta\right)-\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right\|\xrightarrow{p}0$ as $\left\|\theta-\theta_0\right\|$. This follows from the triangle inequality since $\left\|\partial_a\widehat{g}_x\left(\cdot,\theta\right)-\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right\|$ is bounded above by $\left\|\partial_a\widehat{g}_x\left(\cdot,\theta\right)-\partial_a g_{0,x}\left(\cdot,\theta\right)\right\|+\left\|\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)-\partial_a g_{0,x}\left(\cdot,\theta_0\right)\right\|+\left\|\partial_a g_{0,x}\left(\cdot,\theta\right)-\partial_a g_{0,x}\left(\cdot,\theta_0\right)\right\|$, and the fact that the first two terms of the majorant converge to zero by Lemma 5 and the last term converges to zero by the continuity of $\partial_a g_{0,x}\left(\cdot,\theta\right)$ in θ. For B_{21}

$$\begin{aligned}
B_{21} &= 2R^{-1/2}\sum_{x\in X}\int\left[\begin{array}{c}\left[\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x\left(\cdot,\theta\right)\right)-\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right)\right]\\ \times\left[F_{A|X=x}\left(\theta,\partial_a\widehat{g}_x\left(\cdot,\theta\right)\right)-F_{A|X=x}\left(\theta_0,\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right)\right]\end{array}\right]d\mu\\
&= 2R^{-1/2}\sum_{x\in X}\int\left[\begin{array}{c}\left[\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x\left(\cdot,\theta\right)\right)-\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x\left(\cdot,\theta_0\right)\right)\right]\\ \times D_\theta F_{A|X=x}\left(\overline{\theta}_x,\partial_a\widehat{g}_x\left(\cdot,\overline{\theta}_x\right)\right)'\end{array}\right]d\mu\left(\theta-\theta_0\right),
\end{aligned}$$

by Cauchy Schwarz inequality

$$\begin{aligned}
\left|B_{21}\right| &\leq o_p\left(R^{-1/2}\right)\times\\
&\quad\max_{1\leq x\leq J}\left[\left(\theta-\theta_0\right)'\int\left[D_\theta F_{A|X=x}\left(\overline{\theta}_x,\partial_a\widehat{g}_x\left(\cdot,\overline{\theta}_x\right)\right)D_\theta F_{A|X=x}\left(\overline{\theta}_x,\partial_a\widehat{g}_x\left(\cdot,\overline{\theta}_x\right)\right)'\right]d\mu\left(\theta-\theta_0\right)\right]^{1/2}\\
&= o_p\left(R^{-1/2}\right)O_p\left(\left\|\theta-\theta_0\right\|\right)\\
&= o_p\left(N^{-1/2}\left\|\theta-\theta_0\right\|\right),
\end{aligned}$$

the first inequality follows from the stochastic equicontinuity condition of Lemma 8, then it is easy to show the outer product term inside the integral is also bounded in probability and the last equality follows from $N=o\left(R\right)$. This same argument using Cauchy Schwarz inequality again be applied for B_{22} and B_{23}, in particular, it follows from Lemma 10 and Lemma 9 respectively that $\left|B_{22}\right|=o\left(N^{-1}\right)$ and $\left|B_{23}\right|=o\left(N^{-1}\right)$.

For B_3 :

For each x

$$\begin{aligned}\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)+\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right) &= 2\nu_R\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\\ &\quad+\left(\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)-\nu_R\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\right)\\ &\quad+\left(\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)-\nu_R\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\right),\end{aligned}$$

then we can write B_3 as

$$\begin{aligned}B_3 &= 2S^{-1/2}\sum_{x\in X}\int\nu_{R,x}\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\left(F_{A|X=x}\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)-F_{A|X=x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)\right)d\mu_x\\ &\quad+R^{-1/2}\sum_{x\in X}\int\left[\begin{array}{c}\left[\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)-\nu_R\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\right]\\ \times\left[F_{A|X=x}\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)-F_{A|X=x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)\right]\end{array}\right]d\mu_x\\ &\quad+R^{-1/2}\sum_{x\in X}\int\left[\begin{array}{c}\left[\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)-\nu_R\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\right]\\ \times\left[F_{A|X=x}\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)-F_{A|X=x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)\right]\end{array}\right]d\mu_x\\ &= B_{31}+B_{32}+B_{33}.\end{aligned}$$

For each x: we have $\left[\int\left[F_{A|X=x}\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)-F_{A|X=x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)\right]^2d\mu_x\right]^{1/2}=O_p\left(\|\theta-\theta_0\|\right)$ by Cauchy Schwarz inequality; from Donsker theorem and CMT, $\left[\int\left[\nu_R\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\right]^2d\mu_x\right]^{1/2}=O_p(1)$. Then it follows that $|B_{31}|\leq o_p\left(N^{-1/2}\|\theta-\theta_0\|\right)$. By a similar argument, using Cauchy Schwarz inequality, continuity of $\partial_a g(\cdot,\theta)$ in θ, Lemma 5, 6 and 8, $|B_{32}|$ and $|B_{33}|$ are also $o_p\left(N^{-1/2}\|\theta-\theta_0\|\right)$, in particular as we can use the triangle inequality to show $\left\|\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)-\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\right\|_\nu$ and $\left\|\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)-\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\right\|_\nu$ converge in probability to 0 as $\|\theta-\theta_0\|\to 0$ for all x.

For B_4 :

By the same argument above, we can re-express B_4

$$\begin{aligned}B_4 &= 2S^{-1}\sum_{x\in X}\int\nu_{R,x}\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\left(\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)-\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)\right)d\mu_x\\ &\quad+R^{-1}\sum_{x\in X}\int\left[\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)-\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)\right]^2d\mu_x\\ &\quad+R^{-1}\sum_{x\in X}\int\left[\begin{array}{c}\left[\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)-\nu_R\left(\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)\right)\right]\\ \times\left[\nu_{R,x}\left(\theta,\partial_a\widehat{g}_x(\cdot,\theta)\right)-\nu_{R,x}\left(\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)\right)\right]\end{array}\right]d\mu_x\\ &= B_{41}+B_{42}+B_{43}.\end{aligned}$$

By repeatedly using Cauchy Schwarz inequality, continuity of $\partial_a g(\cdot,\theta)$ in θ, and Lemma 5,6 and 8, as seen in the analysis of B_2 and B_3, it follows easily that $|B_{4i}|=o_p\left(N^{-1}\right)$ for $i=1,2,3$.

G6 then follows from Lemma 10.■

Proof of Theorem 2. We proceed by obtaining the pointwise distribution theory for $\widehat{g}_{\theta,x}(a)$

for any non-stochastic $\theta \in \Theta$ and $a \in int\,(A)$; more specifically

$$\sqrt{Nh}\left(\widehat{g}_{\theta,x}(a) - g_{\theta,x}(a) - B_{N,x}(a;m_\theta)\right) \Rightarrow \mathcal{N}\left(0, \frac{\kappa_2(K)}{Tp_{X,A}(x,a)}\text{var}\left(m_\theta(x_{it+1})\,|x_{it}=x, a_{it}=a\right)\right).$$

This immediately follows from Lemma 4. For the asymptotic distribution, we only have to calculate the variance of (28), the rest follows by standard CLT.

From (13) we have

$$\begin{aligned}
\widehat{g}_{\widehat{\theta}} - \widehat{g}_{\theta_0} &= \widehat{\mathcal{H}}\left(I-\widehat{\mathcal{L}}\right)^{-1}\left(\widetilde{r}_{\widehat{\theta}} - \widetilde{r}_{\theta_0}\right) \\
&= \widehat{\mathcal{H}}\left(I-\widehat{\mathcal{L}}\right)^{-1}\left(\left(\widehat{\theta}-\theta_0\right)' D_\theta \widetilde{r}_{\overline{\theta}}\right)
\end{aligned}$$

where the expansion above follows from MVT and $\overline{\theta}$ denotes some intermediate value between $\widehat{\theta}$ and θ_0. It is easy to see that, for $x \in X$

$$\begin{aligned}
\left\|\widehat{g}_x\left(\cdot,\widehat{\theta}\right) - \widehat{g}_x(\cdot,\theta_0)\right\|_\infty &= O_p\left(\left\|\widehat{\theta}-\theta_0\right\|\right) \\
&= O_p\left(N^{-1/2}\right),
\end{aligned}$$

since $\left\|\widehat{\mathcal{H}}\left(I-\widehat{\mathcal{L}}\right)^{-1}\right\| = O_p(1)$, $\|\widetilde{r}_\theta\| = O_p(1)$ and $\sqrt{Nh} = o\left(N^{1/2}\right)$, then $\sqrt{Nh}\left|\widehat{g}_x\left(a,\widehat{\theta}\right) - \widehat{g}_x(a,\theta_0)\right|$ $= o_p(1)$. It remains to show the asymptotic independence between any pair $\left(\widehat{g}_x\left(a,\widehat{\theta}\right), \widehat{g}_{x'}\left(a',\widehat{\theta}\right)\right)$ for any $x' \neq x$ and $a' \neq a$. Since

$$\begin{aligned}
& \text{cov}\left(\widehat{g}_x\left(a,\widehat{\theta}\right), \widehat{g}_{x'}\left(a',\widehat{\theta}\right)\right) \\
= \; & \text{cov}\left(\widehat{g}_x(a,\theta_0), \widehat{g}_{x'}(a',\theta_0)\right) + \text{cov}\left(\widehat{g}_x(a,\theta_0), \widehat{g}_{x'}\left(a',\widehat{\theta}\right) - \widehat{g}_{x'}(a',\theta_0)\right) \\
& + \text{cov}\left(\widehat{g}_{x'}(a',\theta_0), \widehat{g}_x\left(a,\widehat{\theta}\right) - \widehat{g}_x(a,\theta_0)\right) + \text{cov}\left(\widehat{g}_x\left(a,\widehat{\theta}\right) - \widehat{g}_x(a,\theta_0), \widehat{g}_{x'}\left(a',\widehat{\theta}\right) - \widehat{g}_{x'}(a',\theta_0)\right),
\end{aligned}$$

by Cauchy-Schwarz inequality, it suffices to show $\text{var}\left(\sqrt{Nh}\left(\widehat{g}_{x'}\left(a',\widehat{\theta}\right) - \widehat{g}_{x'}(a',\theta_0)\right)\right) = o(1)$; this follows since $\left\|\widehat{g}_x\left(\cdot,\widehat{\theta}\right) - \widehat{g}_x(\cdot,\theta_0)\right\|_\infty = O_p\left(N^{-1/2}\right)$.$\blacksquare$

7.3 Proofs of Theorem G and Lemmas 1-10

Proof of Theorem G. Our argument proceeds in a similar fashion to the case with no preliminary estimates of Newey and McFadden (1994, Theorem 7.1), see also Pollard (1985), by first showing that $\widehat{\theta}$ converge to θ_0 at rate $N^{-1/2}$. By definition of the estimator, we have $M_N\left(\widehat{\theta}, \widehat{g}\left(\cdot,\widehat{\theta}\right)\right) -$

$M_N(\theta_0, \widehat{g}(\cdot,\theta_0)) \leq o_p(N^{-1})$, and

$$
\begin{aligned}
& M_N\left(\widehat{\theta}, \widehat{g}\left(\cdot, \widehat{\theta}\right)\right) - M_N(\theta_0, \widehat{g}(\cdot,\theta_0)) \\
= \;& M\left(\widehat{\theta}, \widehat{g}\left(\cdot, \widehat{\theta}\right)\right) - M(\theta_0, \widehat{g}) + C_N'\left(\widehat{\theta} - \theta_0\right) + N^{-1/2}\left\|\widehat{\theta} - \theta_0\right\| \mathcal{D}_N\left(\widehat{\theta}, \widehat{g}\left(\cdot, \widehat{\theta}\right)\right) \\
\geq \;& \left(C_N + S\left(\theta_0, \widehat{g}\left(\cdot, \widehat{\theta}\right)\right)\right)'\left(\widehat{\theta} - \theta_0\right) + C_0\left\|\widehat{\theta} - \theta_0\right\|^2 (1 + o_p(1)) + N^{-1/2}\left\|\widehat{\theta} - \theta_0\right\| \mathcal{D}_N\left(\widehat{\theta}, \widehat{g}\left(\cdot, \widehat{\theta}\right)\right) \\
= \;& O_p\left(N^{-1/2}\right)'\left(\widehat{\theta} - \theta_0\right) + C_0\left\|\widehat{\theta} - \theta_0\right\|^2 + o_p\left(N^{-1/2}\left\|\widehat{\theta} - \theta_0\right\| + \left\|\widehat{\theta} - \theta_0\right\|^2\right).
\end{aligned}
$$

The first equality follows from the definition of $\mathcal{D}_N$ in (20). For the inequality, we expand $M\left(\widehat{\theta}, \widehat{g}\left(\cdot, \widehat{\theta}\right)\right)$ around θ_0, since $H(\theta, g)$ is continuous around (θ_0, g_0) and H_0 is positive definite by G3, there exists some $C_0 > 0$ such that, w.p.a. 1, $(\theta - \theta_0)' H(\theta_0, \widehat{g}(\cdot,\theta_0))(\theta - \theta_0) + o_p\left(\|\theta - \theta_0\|^2\right) \geq C_0 \|\theta - \theta_0\|^2$. Notice that $C_N + S(\theta_0, \widehat{g}(\cdot,\theta_0)) = O_p\left(N^{-1/2}\right)$, the first term follows from assumption G6 and the latter by G3 and G6 since

$$
\begin{aligned}
\|S(\theta_0, \widehat{g}(\cdot,\theta_0))\| \;&\leq\; \|S(\theta_0, \widehat{g}(\cdot,\theta_0)) - D_g S(\theta_0, g_0(\cdot,\theta_0))[\widehat{g}(\cdot,\theta_0) - g_0(\cdot,\theta_0)]\| \\
&\quad + \|D_g S(\theta_0, g_0(\cdot,\theta_0))[\widehat{g}(\cdot,\theta_0) - g_0(\cdot,\theta_0)]\| \\
&\leq\; o_p\left(N^{-1/2}\right) + O_p\left(N^{-1/2}\right) \\
&=\; O_p\left(N^{-1/2}\right).
\end{aligned}
$$

By completing the square

$$
\left(\left\|\widehat{\theta} - \theta_0\right\| + O_p\left(N^{-1/2}\right)\right)^2 + o_p\left(N^{-1/2}\left\|\widehat{\theta} - \theta_0\right\| + \left\|\widehat{\theta} - \theta_0\right\|^2\right) \leq o_p\left(N^{-1}\right),
$$

thus $\left\|\widehat{\theta} - \theta_0\right\| = O_p\left(N^{-1/2}\right)$. To obtain the asymptotic distribution we define the following related criterion, $J_N(\theta) = D_N(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)' H_0 (\theta - \theta_0)$, note that $J_N(\theta)$ is defined for each $\widehat{g}(\cdot,\theta)$ that satisfies the conditions of Theorem G2, implicit in D_N. $J_N(\theta)$ is a quadratic approximation of $M_N(\theta, \widehat{g}(\cdot,\theta)) - M_N(\theta_0, \widehat{g}(\cdot,\theta_0))$, whose unique minimizer is $\widetilde{\theta} = \theta_0 - H_0^{-1} D_N$, and $\sqrt{N}\left(\widetilde{\theta} - \theta_0\right) \Rightarrow \mathcal{N}\left(0, H_0^{-1}\Omega H_0^{-1}\right)$. Next, we show the approximation error of $J_N(\theta)$ from $M_N(\theta, \widehat{g}(\cdot,\theta)) - M_N(\theta_0, \widehat{g}(\cdot,\theta))$ is small. For any $\theta_N = \theta_0 + O_p\left(N^{-1/2}\right)$ in Θ_{δ_N},

$$
\begin{aligned}
& M_N(\theta_N, \widehat{g}(\cdot,\theta_N)) - M_N(\theta_0, \widehat{g}(\cdot,\theta_0)) \\
= \;& M(\theta_N, \widehat{g}(\cdot,\theta_N)) - M(\theta_0, \widehat{g}(\cdot,\theta_0)) + C_N'(\theta_N - \theta_0) + \frac{\|\theta_N - \theta_0\|}{\sqrt{N}} \mathcal{D}_N(\theta_N, \widehat{g}(\cdot,\theta_N)) \\
= \;& (C_N + S(\theta_0, \widehat{g}(\cdot,\theta_0)))'(\theta_N - \theta_0) + \frac{1}{2}(\theta_N - \theta_0)' H\left(\overline{\theta}, \widehat{g}\left(\cdot, \overline{\theta}\right)\right)(\theta_N - \theta_0) + \frac{\|\theta_N - \theta_0\|}{\sqrt{N}} \mathcal{D}_N(\theta_N, \widehat{g}(\cdot,\theta_N)) \\
= \;& D_N'(\theta_N - \theta_0) + \frac{1}{2}(\theta_N - \theta_0)' H_0 (\theta_N - \theta_0) + o_p\left(\frac{\|\theta_N - \theta_0\|}{\sqrt{N}} + \|\theta_N - \theta_0\|^2\right) \\
= \;& J_N(\theta_N) + o_p\left(\frac{1}{N}\right).
\end{aligned}
$$

The equalities in the display follow straightforwardly from the definition of the $\mathcal{D}_N$, G3, G4 and G5. In particular, this implies that $M_N\left(\theta_N, \widehat{g}\left(\cdot, \theta_N\right)\right) - M_N\left(\theta_0, \widehat{g}\left(\cdot, \theta_0\right)\right) = J_N\left(\theta_N\right) + o_p\left(\frac{1}{N}\right)$ when $\theta_N = \widehat{\theta}$ or $\widetilde{\theta}$, hence we have

$$\begin{aligned} J_N\left(\widehat{\theta}\right) &= \left(J_N\left(\widehat{\theta}\right) - \left(M_N\left(\theta_N, \widehat{g}\left(\cdot, \theta_N\right)\right) - M_N\left(\theta_0, \widehat{g}\left(\cdot, \theta_0\right)\right)\right)\right) \\ &\quad + \left(M_N\left(\theta_N, \widehat{g}\left(\cdot, \theta_N\right)\right) - M_N\left(\theta_0, \widehat{g}\left(\cdot, \theta_0\right)\right)\right) \\ &\leq J_N\left(\widetilde{\theta}\right) + o_p\left(\frac{1}{N}\right), \end{aligned}$$

where the inequality follows from the relation derived from the previous display and G1. Since $J_N\left(\widetilde{\theta}\right) \leq J_N\left(\widehat{\theta}\right)$,

$$\begin{aligned} o_p\left(\frac{1}{N}\right) &= J_N\left(\widehat{\theta}\right) - J_N\left(\widetilde{\theta}\right) \\ &= D_N\left(\widehat{\theta} - \theta_0\right) + \frac{1}{2}\left(\widehat{\theta} - \theta_0\right)' H_0\left(\widehat{\theta} - \theta_0\right) - D_N\left(\widetilde{\theta} - \theta_0\right) - \frac{1}{2}\left(\widetilde{\theta} - \theta_0\right)' H_0\left(\widetilde{\theta} - \theta_0\right) \\ &= -\left(\widetilde{\theta} - \theta_0\right)' H_0\left(\widehat{\theta} - \theta_0\right) + \frac{1}{2}\left(\widehat{\theta} - \theta_0\right)' H_0\left(\widehat{\theta} - \theta_0\right) + \frac{1}{2}\left(\widetilde{\theta} - \theta_0\right)' H_0\left(\widetilde{\theta} - \theta_0\right) \\ &= \frac{1}{2}\left(\widehat{\theta} - \widetilde{\theta}\right)' H_0\left(\widehat{\theta} - \widetilde{\theta}\right), \end{aligned}$$

this implies that $\left\|\widehat{\theta} - \widetilde{\theta}\right\|^2 = o_p\left(\frac{1}{N}\right)$. Since $\sqrt{N}\left(\widetilde{\theta} - \theta_0\right)$ has the desired asymptotic distribution, this completes the proof.■

Proof of Lemma 1. We can write, for any $1 \leq x', x \leq J$

$$\widehat{p}_{X'|X}\left(x'|x\right) - p_{X'|X}\left(x'|x\right) = \frac{\widehat{p}_{X',X}\left(x', x\right) - p_{X',X}\left(x', x\right)}{p_X\left(x\right)} - \frac{\widehat{p}_{X'|X}\left(x'|x\right)}{p_X\left(x\right)\widehat{p}_X\left(x\right)}\left(\widehat{p}_X\left(x\right) - p_X\left(x\right)\right).$$

Under M4, by standard CLT and LLN, we have $\widehat{p}_{X',X}\left(x', x\right) - p_{X',X}\left(x', x\right) = O_p\left(N^{-1/2}\right), \widehat{p}_X\left(x\right) - p_X\left(x\right) = O_p\left(N^{-1/2}\right)$ and $\widehat{p}_X\left(x\right)^{-1} = O_p\left(1\right)$, so it follows that $\widehat{p}_{X'|X}\left(x'|x\right) - p_{X'|X}\left(x'|x\right) = O_p\left(N^{-1/2}\right)$ for any x' and x. Since $\mathcal{L}$ is a linear map on $\mathbb{R}^J$ to $\mathbb{R}^J$, for any vector $m \in \mathbb{R}^J$ we have $\left(\left(\widehat{\mathcal{L}} - \mathcal{L}\right)m\right)_x = \beta\sum_{x' \in X}(\widehat{p}\left(x'|x\right) - p\left(x'|x\right))m_i = O_p\left(N^{-1/2}\right)$ for all x then it follows from the definition of an operator norm that $\left\|\widehat{\mathcal{L}} - \mathcal{L}\right\| = O_p\left(N^{-1/2}\right)$.■

Proof of Lemma 2. For any $x \in X$ and $\theta \in \Theta$, $\widetilde{r}_\theta\left(x\right)$ is defined in Section 2, and define $\widehat{r}_\theta\left(x\right) = \frac{1}{NT}\sum_{i=1,t=1}^{N,T}\frac{1}{\widehat{p}_X(x)}u_\theta(a_{it}, x_{it}, \varepsilon_{it})$. Then we write

$$\widetilde{r}_\theta\left(x\right) - r_\theta\left(x\right) = \left(\widehat{r}_\theta\left(x\right) - r_\theta\left(x\right)\right) + \left(\widetilde{r}_\theta\left(x\right) - \widehat{r}_\theta\left(x\right)\right), \tag{24}$$

the first term is the usual term had we observed $\{\varepsilon_{it}\}$, the latter term arises due to the use of generated residuals. Treating them separately, for the first term

$$
\begin{aligned}
\widehat{r}_\theta(x) - r_\theta(x) &= \frac{1}{\widehat{p}_X(x)} \frac{1}{NT} \sum_{i=1,t=1}^{N,T} \mathbf{1}[x_{it} = x] \left(u_\theta(a_{it}, x_{it}, \varepsilon_{it}) - r_\theta(x) \right) \\
&= \frac{1}{\widehat{p}_X(x)} \frac{1}{NT} \sum_{i=1,t=1}^{N,T} \upsilon_{\theta,it} \mathbf{1}[x_{it} = x],
\end{aligned}
$$

where for each θ, $\upsilon_{\theta,it} = u_\theta(a_{it}, x_{it}, \varepsilon_{it}) - r_\theta(x_{it})$ is a zero mean random variable, note that $\mathbf{1}[x_{it} = x] \times (r_\theta(x_{it}) - r_\theta(x)) = 0$ for all i, x and t. Define $\Upsilon_{N,x}(\theta)$ as the sample average of i.i.d. random variables $\left\{ \sum_{t=1}^{T} \frac{1}{T} \upsilon_{\theta,it} \mathbf{1}[x_{it} = x] \right\}_{i=1}^{N}$, given the assumptions on the DGP, in particular on the second moments, $\Upsilon_{N,x}(\theta) = O_p\left(N^{-1/2}\right)$ for any θ by standard CLT. We want to obtain the uniform rate of convergence of $\Upsilon_{N,x}(\theta)$ over Θ. This can be achieved by using the arguments along the line of Masry (1996). We first obtain the uniform bound for the variance of $\Upsilon_{N,x}(\theta)$, some exponential inequality is then applied to get the rate of decay on the tail probability for any θ. The pointwise rate can then be made uniform by Lipschitz continuity of $\upsilon_{\theta,it}$ (in θ) and compactness of Θ. More precisely, we first show that $\sup_{\theta \in \Theta} \mathrm{var}(\Upsilon_{N,x}(\theta)) = O(N^{-1})$. Since $\mathrm{var}(\Upsilon_{N,x}(\theta))$ is just a variance of $\sum_{t=1}^{T} \frac{1}{T} \upsilon_{\theta,it} \mathbf{1}[x_{it} = x]$ by divided by N, the numerator takes the following form

$$
\begin{aligned}
\mathrm{var}\left(\frac{1}{T} \sum_{t=1}^{T} \upsilon_{\theta,it} \mathbf{1}[x_{it} = x] \right) &= \frac{1}{T} \sum_{t=1}^{T} \mathrm{var}\left(\upsilon_{\theta,it} \mathbf{1}[x_{it} = x] \right) \\
&\quad + \frac{2}{T} \sum_{s=1}^{T-1} \left(1 - \frac{s}{T} \right) Cov\left(\upsilon_{\theta,i0} \mathbf{1}[x_{i0} = x], \upsilon_{\theta,is} \mathbf{1}[x_{is} = x] \right), \\
&= Y_{\theta,1,x} + Y_{\theta,2,x}.
\end{aligned}
$$

The covariance structure in $Y_{\theta,2,x}$ follows from the strict stationarity assumption, which also implies we can write $Y_{\theta,1,x} = E\left[\left[\upsilon_{\theta,it}^2 | x_{it} \right] \mathbf{1}[x_{it} = x] \right]$. Since $u_\theta(a, x, \varepsilon)$ is continuous in θ for all a, x and ε, it follows that $\sup_{\theta \in \Theta} Y_{\theta,1,x} < \infty$. For the covariance term, by Cauchy-Schwarz inequality, $Cov\left(\upsilon_{\theta,i0} \mathbf{1}[x_{i0} = x], \upsilon_{\theta,is} \mathbf{1}[x_{is} = x] \right) \leq E\left[\upsilon_{\theta,i0}^2 \mathbf{1}[x_{i0} = x] \right]^2 < \infty$, since $\sup_{\theta \in \Theta} \left| \upsilon_{\theta,i0,x}^2 \right| < \infty$, it follows that $\sup_\theta Y_{\theta,2} < \infty$ for any finite T. Since $\Upsilon_{N,x}(\theta)$ is an average of $N-$i.i.d. sequence of random variables that, for each θ, it satisfies the Cramér' conditions (since u is uniformly bounded over all its arguments), then Bernstein's inequality, e.g. see Bosq (1998), can be used to obtain the following bound

$$
\Pr\left[|N \Upsilon_{N,x}(\theta)| > N \delta_N \right] \leq 2 \exp\left\{ -\frac{N^2 \delta_N^2}{4 Var\left(N \Upsilon_{N,x}(\theta) \right) + 2CN\delta_N} \right\}. \tag{25}
$$

Let $\delta_N = N^{(-1+\xi)/2}$, simple calculation of the display above yields $\Pr\left[|\Upsilon_{N,x}(\theta)| > \delta_N \right] = O\left(\exp\left(-N^\xi \right) \right)$. By compactness of Θ, let $\{L_N\}_{N=1}^{\infty}$ be an increasing sequence of natural number, we can define a sequence $\{\theta_{iL_N}\}_{i=1}^{L_N}$ to be the centres of open balls, $\{\Theta_{iL_N}\}_{i=1}^{L_N}$, of radius $\{\epsilon_{L_N}\}_{i=1}^{L_N}$ such that $\Theta \subset \bigcup_{i=1}^{L_N} \Theta_{iL_N}$ and $L_N \times \epsilon_{L_N} = O(1)$, then it follows that

$$\begin{aligned}
\Pr\left[\sup_{\theta} |\Upsilon_{N,x}(\theta)| > \delta_N\right] &\leq \Pr\left[\max_{1\leq i\leq L_N} |\Upsilon_{N,x}(\theta_{iL_N})| > \delta_N\right] \\
&\quad + \Pr\left[\max_{1\leq i\leq N} \sup_{\theta\in\Theta_{iL_N}} |\Upsilon_{N,x}(\theta) - \Upsilon_{N,x}(\theta_{iL_N})| > \delta_N\right] \\
&\leq C_0 L_N \exp\left(-N^{\xi}\right) + \Pr\left[\epsilon_{L_N} > \delta_N\right] \\
&= o(1).
\end{aligned}$$

The second inequality from the display above follows from, Bonferroni inequality and (25) for the first term, and by Lipschitz continuity of $\Upsilon_{N,x}$ for the latter. Then the equality holds if we take $\epsilon_{L_N} = o(\delta_N)$ such that L_N grows at some power rate. It then follows that that $\sup_\theta |\Upsilon_{N,x}(\theta)| = o_p\left(N^{-\lambda}\right)$. Then w.p.a. 1

$$\begin{aligned}
\sup_{\theta\in\Theta} |\widehat{r}_\theta(x) - r_\theta(x)| &\leq \frac{\max_{1\leq x\leq J} \sup_{\theta\in\Theta} |\Upsilon_{N,x}(\theta)|}{\min_{1\leq x\leq J} p_X(x)} \\
&= o_p\left(N^{-\lambda}\right).
\end{aligned}$$

The procedure to obtain the uniform rate of convergence is shown above in detail to avoid repetition later since we will require to show many zero mean processes converge uniformly (either over the compact parameter space or the state space) to zero faster than some rates. The argument above can also be applied to nonparametric estimates, as well as some other appropriately (weakly) dependent zero mean process, see Linton and Mammen (2005), and especially Srisuma and Linton (2009) for such usages in closely related context. We comment here that, our paper along with the papers mentioned in the previous sentence, unlike Masry (1996), are not interested in sharp rate of uniform convergence so our proofs are comparatively more straightforward.

For the generated residuals, by definition

$$\widetilde{r}_\theta(x) - \widehat{r}_\theta(x) = \frac{1}{NT}\sum_{i=1,t=1}^{N,T} w_{itN}(x)\left(u_\theta(a_{it}, x_{it}, \widehat{\varepsilon}_{it}) - u_\theta(a_{it}, x_{it}, \varepsilon_{it})\right),$$

where $\widehat{\varepsilon}_{it} = \chi\left(\widehat{F}_{A|X}(a_{it}|x_{it})\right)$ with $\chi \equiv Q^{-1}$. Using mean value expansion, $u_\theta(a_{it}, x_{it}, \widehat{\varepsilon}_{it}) - u_\theta(a_{it}, x_{it}, \varepsilon_{it}) = \frac{\partial}{\partial\varepsilon} u_\theta(a_{it}, x_{it}, \overline{\varepsilon}_{it}) \chi'\left(\overline{F}_{A|X}(a_{it}|x_{it})\right)\left(\widehat{F}_{A|X}(a_{it}|x_{it}) - F_{A|X}(a_{it}|x_{it})\right)$, where $\overline{\varepsilon}_{it}$ and $\overline{F}_{A|X}(a_{it}|x_{it})$ are some intermediate points between $\widehat{\varepsilon}_{it}$ and ε_{it}, and, $\widehat{F}_{A|X}(a_{it}|x_{it})$ and $F_{A|X}(a_{it}|x_{it})$, respectively. Then it follows that

$$\begin{aligned}
\widetilde{r}_\theta(x)-\widehat{r}_\theta(x) &= \frac{1}{NT}\sum_{i=1,t=1}^{N,T} w_{itN}(x)\left(u_\theta\left(a_{it},x_{it},\widehat{\varepsilon}_{it}\right)-u_\theta\left(a_{it},x_{it},\varepsilon_{it}\right)\right) \\
&= \frac{1}{NT}\sum_{i=1,t=1}^{N,T}\frac{\mathbf{1}\left[x_{it}=x\right]}{p_X(x)}\varkappa_\theta\left(a_{it},x_{it},\varepsilon_{it}\right)\left(\widehat{F}_{A|X}\left(a_{it}|x_{it}\right)-F_{A|X}\left(a_{it}|x_{it}\right)\right)+O_p\left(N^{-1}\right),
\end{aligned}$$

where $\varkappa_\theta\left(a_{it},x_{it},\varepsilon_{it}\right)=\frac{\partial}{\partial\varepsilon}u_\theta\left(a_{it},x_{it},\varepsilon_{it}\right)\chi'\left(F_{A|X}\left(a_{it}|x_{it}\right)\right)$. In addition, the $O_p\left(N^{-1}\right)-$term holds uniformly over θ and x, this follows from Markov inequality since $\frac{\partial^2}{\partial\varepsilon^2}u$ and χ'' are uniformly bounded over all of their arguments, $\max_{1\leq x\leq J}\left|\widehat{p}_X(x)-p_X(x)\right|=O_p\left(N^{-1/2}\right)$, and, $\max_{1\leq x\leq J}\sup_{a\in A}\left|\widehat{F}_{A|X}(a|x)-F_{A|X}(a|x)\right|=O_p\left(N^{-1/2}\right)$ by Lemma 9. By a similar argument, using the leave one out estimator for $\widehat{F}_{A|X}$, the leading term for $\widetilde{r}_\theta(x)-\widehat{r}_\theta(x)$ can be simplified further to

$$\frac{1}{NT(NT-1)}\sum_{i=1,t=1}^{N,T}\sum_{x,s,(-it)}^{N,T}\varkappa_\theta\left(a_{it},x_{it},\varepsilon_{it}\right)\frac{\mathbf{1}\left[x_{it}=x\right]}{p_X(x)}\frac{\mathbf{1}\left[x_{xs}=x_{it}\right]}{p\left(x_{it}\right)}\left(\mathbf{1}\left[a_{xs}\leq a_{it}\right]-F_{A|X}\left(a_{it}|x_{it}\right)\right),$$

where $\sum_{x,s,(-it)}^{N,T}$ sums over the indices $x=1,\ldots,N$ and $s=1,\ldots,T$ but omits the $it^{th}-$summand. Subsequently, the term in the display above can be written as the following second order U-statistic

$$\begin{pmatrix} NT \\ 2 \end{pmatrix}^{-1}\frac{1}{2}\sum_{C((it),(xs))}\begin{pmatrix} \varkappa_\theta\left(a_{it},x_{it},\varepsilon_{it}\right)\frac{\mathbf{1}[x_{it}=x]}{p_X(x)}\frac{\mathbf{1}[x_{xs}=x_{it}]}{p(x_{it})}\left(\mathbf{1}\left[a_{xs}\leq a_{it}\right]-F_{A|X}\left(a_{it}|x_{it}\right)\right) \\ +\varkappa_\theta\left(a_{xs},x_{xs},\varepsilon_{xs}\right)\frac{\mathbf{1}[x_{xs}=x]}{p_X(x)}\frac{\mathbf{1}[x_{xs}=x_{it}]}{p(x_{xs})}\left(\mathbf{1}\left[a_{it}\leq a_{xs}\right]-F_{A|X}\left(a_{xs}|x_{xs}\right)\right)\end{pmatrix},$$

where $\sum_{C((it),(xs))}$ sums over all distinct combinations of $C\left((it),(xs)\right)$. Note that $\mathbf{1}\left[a_{it}\leq a\right]=F_{A|X}\left(a|x_{it}\right)+\omega\left(x_{it};a\right)$ where $E\left[\omega\left(x_{it};a\right)|x_{it}\right]=0$, so $\omega\left(x_{it};\cdot\right)$ is a random element in $L^2(A)$. Then it is straightforward to obtain the leading term of the Hoeffding decomposition of our U-statistic, see Powell, Stock and Stoker (1989) and Lee (1990), in particular we have for all x

$$\widetilde{r}_\theta(x)-\widehat{r}_\theta(x)=\frac{1}{NT}\sum_{i=1,t=1}^{N,T}\zeta_\theta\left(\omega\left(x_{it};\cdot\right),x_{it};x\right)+o_p\left(N^{-1/2}\right),$$

where $\zeta_\theta\left(\omega\left(x_{it};\cdot\right),x_{it};x\right)=\frac{1}{p_X(x)}\int\omega\left(x_{it};a_{xs}\right)\left[\int\varkappa_\theta\left(a_{xs},x_{it},\varepsilon_{xs}\right)\mathbf{1}\left[x_{it}=x\right]\frac{f_{A,X,\mathcal{E}}\left(a_{xs},x_{it},\varepsilon_{xs}\right)}{p(x_{it})}d\varepsilon_{xs}\right]da_{xs}$ and $f_{A,X,\mathcal{E}}$ denotes the joint continuous-discrete density of $\left(a_{it},x_{it},\varepsilon_{it}\right)$. Note that ζ_θ is random with respect to ω_{it} and x_{it}, and $E\left[\omega\left(x_{it};\cdot\right)|x_{it}\right]=0$, so ζ_θ has zero mean. Given the boundedness and smoothness conditions on $\varkappa_\theta$, then $\frac{1}{NT}\sum_{i=1,t=1}^{N,T}\zeta_\theta\left(\omega\left(x_{it};\cdot\right),x_{it};x\right)$ can be shown to converge uniformly in probability to zero faster than the rate $N^{-\lambda}$ as shown above. In sum, we have shown for $x \in X$ that $\widetilde{r}_\theta(x) = r_\theta(x) + \widetilde{r}_\theta^R(x)$ with

$$\begin{aligned} \widetilde{r}_\theta^R(x) &= \frac{1}{p_X(x)} \frac{1}{NT} \sum_{i=1,t=1}^{N,T} \mathbf{1}[x_{it} = x] \left(u_\theta(a_{it}, x_{it}, \varepsilon_{it}) - r_\theta(x)\right) \\ &\quad + \frac{1}{NT} \sum_{i=1,t=1}^{N,T} \zeta_\theta(\omega(x_{it}; \cdot), x_{it}; x) + o_p\left(N^{-\lambda}\right) \\ &= o_p\left(N^{-\lambda}\right), \end{aligned}$$

where the smaller order term holds uniformly over x and θ.■

Proof of Lemma 3. Since $0 < \|\mathcal{L}\| < 1$ and $0 < \left\|\widehat{\mathcal{L}}\right\| < 1$, the argument used in Linton and Mammen (2005) can be used to show

$$\left\| \left(I - \widehat{\mathcal{L}}\right)^{-1} - (I - \mathcal{L})^{-1} \right\| = O_p\left(N^{-1/2}\right).$$

We note that, using the contraction property, $(I - \mathcal{L})^{-1}$ and $\left(I - \widehat{\mathcal{L}}\right)^{-1}$ are bounded linear operators since $\left\|(I - \mathcal{L})^{-1}\right\| \leq (1 - \|\mathcal{L}\|)^{-1} < \infty$ and similarly $\left\|\left(I - \widehat{\mathcal{L}}\right)^{-1}\right\| \leq \left(1 - \left\|\widehat{\mathcal{L}}\right\|\right)^{-1} < \infty$, this can be shown from the respective Neumann series representation of the inverses and by the basic properties of operator norms. Then for each $x \in X$ and $\theta \in \Theta$, $\widehat{m}_\theta(x)$ is defined in (12), we write $\widehat{m}_\theta(x) = \left(I - \widehat{\mathcal{L}}\right)^{-1}\left(r_\theta(x) + \widetilde{r}_\theta^R(x)\right)$, given the results from Lemma 2, it follows that $\max_{1 \leq x \leq J} \sup_{\theta \in \Theta} \left|\left(I - \widehat{\mathcal{L}}\right)^{-1} \widetilde{r}_\theta^R(x)\right| = o_p\left(N^{-\lambda}\right)$, since $\left\|\left(I - \widehat{\mathcal{L}}\right)^{-1}\right\| = O_p(1)$. For first term, we can write $\left(I - \widehat{\mathcal{L}}\right)^{-1} r_\theta(x) = m_\theta(x) + \widehat{m}_\theta^A(x)$ where $\widehat{m}_\theta^A(x) = \left(I - \widehat{\mathcal{L}}\right)^{-1}\left(\widehat{\mathcal{L}} - \mathcal{L}\right) m_\theta(x)$. Since we know $\left\|\left(I - \widehat{\mathcal{L}}\right)^{-1}\right\| = O_p(1)$ from earlier, from Lemma 1 $\left\|\widehat{\mathcal{L}} - \mathcal{L}\right\| = O_p\left(N^{-1/2}\right)$, and, $\max_{1 \leq x \leq J} \sup_{\theta \in \Theta} |m_\theta(x)| = O(1)$ as $m_\theta(x)$ is a continuous function on a compact set Θ any x, this completes the proof with $\widehat{m}_\theta^R = \widehat{m}_\theta^A + \left(I - \widehat{\mathcal{L}}\right)^{-1} \widetilde{r}_\theta^R$.■

Proof of Lemma 4. The empirical analogue of (9) is

$$\widehat{g}_\theta = \widehat{\mathcal{H}} \widehat{m}_\theta,$$

where $\widehat{\mathcal{H}}$ is a linear operator that uses local constant approximation to estimate the conditional expectation operator $\mathcal{H}$. Then we proceed, similarly to the proof of Lemma 3, by writing $\widehat{g}_x(a, \theta) = g_x(a, \theta) + \widehat{g}_x^A(a, \theta) + \widehat{\mathcal{H}} \widehat{m}_\theta^R(x, a)$ where $\widehat{g}_x^A(a, \theta) = \left(\widehat{\mathcal{H}} - \mathcal{H}\right) m_\theta(x, a)$ for any x. The approach taken here is similar to that found in Srisuma and Linton (2009), we decompose $\widehat{g}_x^A(a, \theta)$ into variance+bias terms, note that the presence of discrete regressor only leads to a straightforward sample splitting in the local regression for each x. Since A is a compact set, the bias term near the boundary for Nadaraya-Watson estimator has a slower rate of convergent there than in the interior, for this reason we will need to trim out values near the boundary of A. For the ease of notation we proceed by assuming that the support of a_{it} is A_N, where $\{A_n\}_{n=1}^{N}$ is a sequence of increasing sets such that $\bigcup_{n=1}^{\infty} A_n = int\,(A)$, here the boundary of the set A has zero measure w.r.t. any relevant measure to our problem so we can ignore the difference between A and $int\,(A)$. In our case $A = [\underline{a}, \overline{a}]$ and $A_N = [\underline{a} + \gamma_N, \overline{a} - \gamma_N]$ such that $\gamma_N = o\,(1)$ and $h = o\,(\gamma_N)$. So we only need the trimming factor to converge to zero (at any rate) slower than the bandwidth, the reason behind this is fact that, for large N, the boundary only effect exists within a neighborhood of a single bandwidth. Then for any $m = (m\,(1) \ldots m\,(J))' \in \mathbb{R}^J$, a and x

$$\begin{aligned}
\left(\widehat{\mathcal{H}} - \mathcal{H}\right) m\,(x,a) &= \sum_{x' \in X} m\,(x') \left(\frac{\widehat{p}_{X',X,A}\,(x',x,a)}{\widehat{p}_{X,A}\,(x,a)} - \frac{p_{X',X,A}\,(x',x,a)}{p_{,X,A}\,(x,a)} \right) \qquad (26) \\
&= \sum_{x' \in X} m\,(x') \left(\frac{\widehat{p}_{X',X,A}\,(x',x,a) - p_{X',X,A}\,(x',x,a)}{p_{X,A}\,(x,a)} \right) \\
&\quad - \sum_{x' \in X} m\,(x') \left(\frac{\widehat{p}_{X',X,A}\,(x',x,a)}{\widehat{p}_{X,A}\,(x,a)\, p_{X,A}\,(x,a)} \left(\widehat{p}_{X,A}\,(x,a) - p_{X,A}\,(x,a)\right) \right),
\end{aligned}$$

where

$$\begin{aligned}
\widehat{p}_{X',X,A}\,(x',x,a) &= \frac{1}{NT} \sum_{i=1,t=1}^{N,T} \mathbf{1}\,[x_{it+1} = x', x_{it} = x]\, K_h\,(a_{it} - a)\,, \\
\widehat{p}_{X,A}\,(x,a) &= \frac{1}{NT} \sum_{i=1,t=1}^{N,T} \mathbf{1}\,[x_{it} = x]\, K_h\,(a_{it} - a)\,.
\end{aligned}$$

For any x, x', then

$$\begin{aligned}
& \widehat{p}_{X',X,A}\,(x',x,a) - p_{X',X,A}\,(x',x,a) \\
= \; & \left(\widehat{p}_{X',X,A}\,(x',x,a) - E\,[\widehat{p}_{X',X,A}\,(x',x,a)]\right) + \left(E\,[\widehat{p}_{X',X,A}\,(x',x,a)] - p_{X',X,A}\,(x',x,a)\right) \\
= \; & I_{11}\,(x',x,a) + I_{12}\,(x',x,a)\,,
\end{aligned}$$

where $I_{11}\,(x',x,a)$ has zero mean and $I_{12}\,(x',x,a)$ is nonstochastic for any $a \in A_N$. Under stationarity, by the standard change of variable and differentiability of $p_{X',X,A}\,(x',x,a)$ (w.r.t. a)

$$I_{12}\,(x',x,a) = \frac{1}{2} h^4 \mu_2\,(K)\, \frac{\partial^2}{\partial a^2} p_{X',X,A}\,(x',x,a) + o\left(h^2\right).$$

It then follows that $\max_{1 \leq x,x' \leq J} \sup_{a \in A_N} |I_{12}\,(x',x,a)| = O\,(h^4)$ since $\frac{\partial^4}{\partial a^4} p_{X',X,A}\,(x',x,a)$ is a continuous function on a for any x and x'. It is also straightforward to show by using the same arguments as in Lemma 2 that $\max_{1 \leq x,x' \leq J} \sup_{a \in A_N} |I_{11}\,(x',x,a)| = o_p\left(\frac{N^{\xi}}{\sqrt{Nh}}\right)$. In particular, this follows since

$$\mathrm{var}\left(\sqrt{NTh} I_{11}\,(x',x,a)\right) = p_{X',X,A}\,(x',x,a)\, \kappa_2\,(K) + o\,(1)\,,$$

where the display above for any x and x' uniformly over A_N. Combining terms we have

$$\begin{aligned}
& \max_{1\leq x,x'\leq J}\sup_{a\in A_N}\left|\sum_{x'\in X} m(x')\left(\frac{\widehat{p}_{X',X,A}(x',x,a)-p_{X',X,A}(x',x,a)}{p_{X,A}(x,a)}\right)\right| \\
\leq\ & J\frac{\max_{1\leq x\leq J}|m_x|}{\min_{1\leq x\leq J}\inf_{a\in A_N}|p_{X,A}(x,a)|}\times\max_{1\leq x,x'\leq J}\sup_{a\in A_N}|\widehat{p}_{X',X,A}(x',x,a)-p_{X',X,A}(x',x,a)| \\
=\ & O_p\left(h^4+\frac{N^{\xi}}{\sqrt{Nh}}\right),
\end{aligned}$$

where the inequality holds w.p.a. 1 since we know (to be shown next) $\widehat{p}_{X,A}$ converges to $p_{X,A}$ uniformly over $X\times A_N$. By the same type of argument as above, write for each x

$$\begin{aligned}
& \widehat{p}_{X,A}(x,a)-p_{X,A}(x,a) \\
=\ & (\widehat{p}_{X,A}(x,a)-E[\widehat{p}_{X,A}(x,a)])+(E[\widehat{p}_{X,A}(x,a)]-p_{X,A}(x,a)) \\
=\ & I_{21}(x,a)+I_{22}(x,a),
\end{aligned}$$

then it is straightforward to show the followings hold uniformly over its arguments

$$\begin{aligned}
I_{22}(x,a) &= \frac{1}{2}h^4\mu_4(K)\frac{\partial^4}{\partial a^2}p_{X,A}(x,a)+o\left(h^2\right), \\
\operatorname{var}\left(\sqrt{NTh}I_{21}(x',x,a)\right) &= p_{X,A}(x,a)\kappa_2(K)+o(1),
\end{aligned}$$

then we have

$$\begin{aligned}
& \max_{1\leq x,x'\leq J}\sup_{a\in A_N}\left|\sum_{x'\in X} m(x')\left(\frac{\widehat{p}_{X',X,A}(x',x,a)}{\widehat{p}_{X,A}(x,a)\,p_{X,A}(x,a)}(\widehat{p}_{X,A}(x,a)-p_{X,A}(x,a))\right)\right| \\
\leq\ & J\frac{\max_{1\leq x\leq J}|m_x|}{\min_{1\leq x\leq J}\inf_{a\in A_N}|p_{X,A}(x,a)|^2}\times\max_{1\leq x\leq J}\sup_{a\in A_N}|\widehat{p}_{X,A}(x,a)-p_{X,A}(x,a)| \\
=\ & O_p\left(h^4+\frac{N^{\xi}}{\sqrt{Nh}}\right).
\end{aligned}$$

So we can write for each x

$$\begin{aligned}
\left(\widehat{\mathcal{H}}-\mathcal{H}\right)m(x,a) &= \sum_{x'\in X} m(x')\left(\frac{\widehat{p}_{X',X,A}(x',x,a)-p_{X',X,A}(x',x,a)}{p_{X,A}(x,a)}\right) \\
&\quad -\sum_{x'\in X} m(x')\left(\frac{p_{X',X,A}(x',x,a)}{p^2_{X,A}(x,a)}(\widehat{p}_{X,A}(x,a)-p_{X,A}(x,a))\right)+W_{N,x}(a;m) \\
&= B_{N,x}(a;m)+V_{N,x}(a;m)+W_{N,x}(a;m),
\end{aligned}$$

where

$$B_{N,x}(a;m) = \frac{1}{2}h^4\mu_4(K)\sum_{x'\in X} m(x')\left(\frac{\frac{\partial^4}{\partial a^4}p_{X',X,A}(x',x,a)}{p_{X,A}(x,a)} + \frac{p_{X',X,A}(x',x,a)\frac{\partial^4}{\partial a^4}p_{X,A}(x,a)}{p^2_{X,A}(x,a)}\right), \quad (27)$$

$$V_{N,x}(a;m) = \sum_{x'\in X} m(x')\left(\begin{array}{c} \frac{1}{p_{X,A}(x,a)}\frac{1}{NT}\sum_{i=1,t=1}^{N,T}\left(\begin{array}{c}\mathbf{1}[x_{it+1}=x', x_{it}=x]K_h(a_{it}-a) \\ -E[\mathbf{1}[x_{it+1}=x', x_{it}=x]K_h(a_{it}-a)]\end{array}\right) \\ -\frac{p_{X',X,A}(x',x,a)}{p^2_{X,A}(x,a)}\frac{1}{NT}\sum_{i=1,t=1}^{N,T}\left(\begin{array}{c}\mathbf{1}[x_{it}=x]K_h(a_{it}-a) \\ -E[\mathbf{1}[x_{it}=x]K_h(a_{it}-a)]\end{array}\right)\end{array}\right) \quad (28)$$

$$W_{N,x}(a;m) = \sum_{x'\in X} m(x')\left(\begin{array}{c}\frac{1}{p_{X,A}(x,a)}\left(\frac{\widehat{p}_{X',X,A}(x',x,a)}{\widehat{p}_{X,A}(x,a)} - \frac{p_{X',X,A}(x',x,a)}{p_{X,A}(x,a)}\right) \\ \times\left(\widehat{p}_{X,A}(x,a) - p_{X,A}(x,a)\right)\end{array}\right). \quad (29)$$

Note that $B_{N,x}$ is a deterministic term, $V_{N,x}$ is the zero mean process that will deliver CLT whilst, using the same arguments as above, it is straightforward to show that $\max_{1\leq x\leq J}\sup_{a\in A_N} W_{N,x}(a;m) = o_p(B_{N,x}(a;m) + V_{N,x}(a;m))$ for any $m \in \mathbb{R}^J$. Then we can conclude $\left\|\widehat{\mathcal{H}} - \mathcal{H}\right\| = O_p\left(h^4 + \frac{N^\xi}{\sqrt{Nh}}\right)$. Using the decomposition of $\widehat{\mathcal{H}} - \mathcal{H}$ above we have

$$\widehat{g}^A_x(a,\theta) = \widehat{g}^B_x(a,\theta) + \widehat{g}^S_x(a,\theta) + W_{N,x}(a;m_\theta),$$

where, from (27) - (28), $\widehat{g}^B_x(a,\theta) = B_{N,x}(a;m_\theta)$ and $\widehat{g}^S_x(a,\theta) = V_{N,x}(a;m_\theta)$. It also follows that these terms have the desired rate of convergence that holds uniformly over Θ as well since $\mathcal{H}$ is independent of θ and m_θ is a vector of $J-$real value functions that are continuous on Θ. Finally, we define $\widehat{g}^R_x(a,\theta)$ to be $W_{N,x}(a;m_\theta) + \widehat{\mathcal{H}}\widehat{m}^R_\theta(x,a)$. By the previous reasoning $W_{N,x}(a;m_\theta)$ already has the desired stochastic order so the proof of Lemma 4 will be complete if we can show, generally, that $\max_{1\leq x\leq J}\sup_{\theta,a\in\Theta\times A_N}\left|\widehat{\mathcal{H}}\widehat{m}^R_\theta(x,a)\right| = o_p\left(h^4 + \frac{N^\xi}{\sqrt{Nh}}\right)$. This is indeed true, since we have already shown that $\left\|\widehat{\mathcal{H}} - \mathcal{H}\right\| = o_p\left(h^4 + \frac{N^\xi}{\sqrt{Nh}}\right)$ and given that $\mathcal{H}$ is a conditional expectation operator, this implies that $\|\mathcal{H}\| \leq 1$, it follows from triangle inequality and the definition of operator norm that $\max_{1\leq x\leq J}\sup_{\theta,a\in\Theta\times A_N}\left|\widehat{\mathcal{H}}\widehat{m}^R_\theta(x,a)\right| = o_p\left(N^{-\lambda}\right)$.■

Proof of Lemma 5. When $l = 0$, this follows from Lemma 4 with $h = O\left(N^{-1/7}\right)$. Other values of l can also be shown very similarly, only more tedious. Since $\dim(A) = 1$ then $\partial^{|l|}_a = \frac{\partial^l}{\partial a^l}$, when $l=1$, taking a derivative w.r.t. a on (26) we obtain

$$\begin{aligned}
\frac{\partial}{\partial a}\left(\widehat{\mathcal{H}}-\mathcal{H}\right)m(x,a) &= \sum_{x'\in X} m(x')\frac{\partial}{\partial a}\left(\frac{\widehat{p}_{X',X,A}(x',x,a)-p_{X',X,A}(x',x,a)}{p_{X,A}(x,a)}\right) \\
&\quad -\sum_{x'\in X} m(x')\frac{\partial}{\partial a}\left(\frac{\widehat{p}_{X',X,A}(x',x,a)}{\widehat{p}_{X,A}(x,a)\,p_{X,A}(x,a)}\left(\widehat{p}_{X,A}(x,a)-p_{X,A}(x,a)\right)\right) \\
&= \sum_{x'\in X} m(x')\left(\begin{array}{c}\frac{1}{p_{X,A}(x,a)}\frac{\partial}{\partial a}\left(\widehat{p}_{X',X,A}(x',x,a)-p_{X',X,A}(x',x,a)\right) \\ -\frac{\frac{\partial}{\partial a}p_{X,A}(x,a)}{p^2_{X,A}(x,a)}\left(\widehat{p}_{X',X,A}(x',x,a)-p_{X',X,A}(x',x,a)\right)\end{array}\right) \\
&\quad -\sum_{x'\in X} m(x')\left(\begin{array}{c}\frac{\widehat{p}_{X',X,A}(x',x,a)}{\widehat{p}_{X,A}(x,a)p_{X,A}(x,a)}\frac{\partial}{\partial a}\left(\widehat{p}_{X,A}(x,a)-p_{X,A}(x,a)\right) \\ +\left(\frac{\partial}{\partial a}\frac{\widehat{p}_{X',X,A}(x',x,a)}{\widehat{p}_{X,A}(x,a)p_{X,A}(x,a)}\right)\left(\widehat{p}_{X,A}(x,a)-p_{X,A}(x,a)\right)\end{array}\right).
\end{aligned}$$

As seen in the proof of Lemma 4, it will be sufficient to show that $\max_{1\leq x,x'\leq J}\sup_{a\in A_N}|\frac{\partial}{\partial a}\widehat{p}_{X',X,A}(x',x,a)$ $-\frac{\partial}{\partial a}p_{X',X,A}(x',x,a)|=o_p(1)$, and, $\max_{1\leq x,x'\leq J}\sup_{a\in A_N}|\frac{\partial}{\partial a}\widehat{p}_{X,A}(x,a)-\frac{\partial}{\partial a}p_{X,A}(x,a)|=o_p(1)$ since we assume that $\frac{\partial}{\partial a}p_{X',X,A}(x',x,a)$ and $\frac{\partial}{\partial a}p_{X,A}(x,a)$ are continuous functions on a compact set A for any x,x'. Proceeding as in the proof of Lemma 4, first note that for any x,x'

$$\begin{aligned}
E\left[\frac{\partial}{\partial a}\widehat{p}_{X',X,A}(x',x,a)\right] &= -\frac{1}{h}\int p_{X',X,A}(x',x,a+wh)\,dK(w) \\
&= \int\frac{\partial}{\partial a}p_{X',X,A}(x',x,a+wh)\,K(w)\,dw \\
&= \frac{\partial}{\partial a}p_{X',X,A}(x',x,a)+O\left(h^4\right).
\end{aligned}$$

The first line in the display follows from a standard change of variable argument, then using integration by parts and Taylor's expansion, the last equality above holds uniformly over A. It is easy to verify that uniformly over A

$$\mathrm{var}\left(\sqrt{NTh^3}\frac{\partial}{\partial a}\widehat{p}_{X',X,A}(x',x,a)\right)=O(1).$$

As seen in Lemma 2, it then follows that $\max_{1\leq x,x'\leq J}\sup_{a\in A_N}\left|\frac{\partial}{\partial a}\widehat{p}_{X',X,A}(x',x,a)-\frac{\partial}{\partial a}p_{X',X,A}(x',x,a))\right|$ $=O_p(h^4+\frac{N^\xi}{\sqrt{Nh^3}})$. Similarly one can show $\max_{1\leq x,x'\leq J}\sup_{a\in A_N}\left|\frac{\partial}{\partial a}\widehat{p}_{X,A}(x,a)-\frac{\partial}{\partial a}p_{X,A}(x,a)\right|=O_p(h^4$ $+\frac{N^\xi}{\sqrt{Nh^3}})$. It is easy to see that choosing $h=O\left(N^{-1/7}\right)$ will imply $\max_{1\leq x\leq J}\sup_{\theta,a\in\Theta\times A_N}|\frac{\partial}{\partial a}\widehat{g}_x(a,\theta)$ $-\frac{\partial}{\partial a}g_x(a,\theta)|=o_p(1)$.■

Proof of Lemma 6. When $p=0$ the result follows from Lemma 5. Consider the case when $p=1$ and $l=0$, for all $1\leq x\leq J, 1\leq k\leq L$ and $\lambda<1/2$, the exact same arguments used in proofing Lemma 2 can then be used to show $\frac{\partial}{\partial\theta_k}\widetilde{r}_\theta(x)=\frac{\partial}{\partial\theta_k}r_\theta(x)+\frac{\partial}{\partial\theta_k}\widetilde{r}^R_\theta(x)$ with $\max_{1\leq x\leq J}\sup_{\theta,a\in\Theta\times A_N}\left|\frac{\partial}{\partial\theta_k}\widetilde{r}^R_\theta(x)\right|=o_p\left(N^{-\lambda}\right)$, and since $\mathcal{L}$ is independent of θ, the same arguments found in Lemma 3 can be used to show $\frac{\partial}{\partial\theta_k}\widehat{m}_\theta(x)=\frac{\partial}{\partial\theta_k}m_\theta(x)+\frac{\partial}{\partial\theta_k}\widehat{m}^R_\theta(x)$ with

$\max_{1\leq x\leq J}\sup_{\theta,a\in\Theta\times A_N}\left|\frac{\partial}{\partial\theta_k}\widehat{m}_\theta^R(x)\right| = o_p\left(N^{-\lambda}\right)$. Apart from replacing (r_θ, m_θ) everywhere by $\left(\frac{\partial}{\partial\theta_k}r_\theta, \frac{\partial}{\partial\theta_k}m_\theta\right)$, we note that it is here that we need $\frac{\partial^2}{\partial\varepsilon\partial\theta_k}u_\theta(a,x,\varepsilon)$ to be continuous on all a, x and θ. Since $\mathcal{H}$ is independent of θ, the arguments used in Lemma 4 can be directly applied to show

$$\frac{\partial}{\partial\theta_k}\widehat{g}_x(a,\theta) = \frac{\partial}{\partial\theta_k}g_x(a,\theta) + \frac{\partial}{\partial\theta_k}\widehat{g}_x^B(a,\theta) + \frac{\partial}{\partial\theta_k}\widehat{g}_x^S(a,\theta) + \frac{\partial}{\partial\theta_k}\widehat{g}_x^R(a,\theta),$$

such that

$$\begin{aligned}
\max_{1\leq x\leq J}\sup_{\theta,a\in\Theta\times A_N}\left|\frac{\partial}{\partial\theta_k}\widehat{g}_x^B(a,\theta)\right| &= O_p\left(h^2\right), \\
\max_{1\leq x\leq J}\sup_{\theta,a\in\Theta\times A_N}\left|\frac{\partial}{\partial\theta_k}\widehat{g}_x^S(a,\theta)\right| &= o_p\left(\frac{N^\xi}{\sqrt{Nh}}\right), \\
\max_{1\leq x\leq J}\sup_{\theta,a\in\Theta\times A_N}\left|\frac{\partial}{\partial\theta_k}\widehat{g}_x^R(a,\theta)\right| &= o_p\left(h^2 + \frac{N^\xi}{\sqrt{Nh}}\right),
\end{aligned}$$

where $\frac{\partial}{\partial\theta_k}\widehat{g}_x^B(a,\theta) = B_{N,x}\left(a;\frac{\partial}{\partial\theta_k}m_\theta\right)$, $\frac{\partial}{\partial\theta_k}\widehat{g}_x^S(a,\theta) = V_{N,x}\left(a;\frac{\partial}{\partial\theta_k}m_\theta\right)$ and $\frac{\partial}{\partial\theta_k}\widehat{g}_x^R(a,\theta) = W_{N,x}\left(a;\frac{\partial}{\partial\theta_k}m_\theta\right)$ $+\widehat{\mathcal{H}}\frac{\partial}{\partial\theta_k}\widehat{m}_\theta^R(x,a)$ and these terms are defined in (27) - (29). For $l=2$ and $1\leq k,d\leq L$, we simply replace $\frac{\partial}{\partial\theta_k}$ by $\frac{\partial^2}{\partial\theta_k\partial\theta_d}$ and the exact same reasoning used when $p=1$ can be applied directly. All other cases of $0\leq l,p\leq 2$ can be shown similarly.■

Proof of Lemma 7. We first show that $\mathbf{1}[\cdot\leq\rho_x(a,\theta,\partial_a g_x)]$ is locally uniformly $L^2(Q)$ −continuous for all x with respect to $a,\theta,\partial_a g_x$. More precisely, we need to show for a positive sequence $\delta_N = o(1)$ and any $(a,\theta,\partial_a g_x)\in A\times\Theta\times\mathcal{G}_x^{(1)}$ that

$$\lim_{N\to\infty}\left(E\left[\sup_{\|(a'-a,\theta'-\theta,\partial_a g_x'-\partial_a g_x)\|<\delta_N}\left|\mathbf{1}\left[\varepsilon_i\leq\rho_x(a',\theta',\partial_a g_x')\right]-\mathbf{1}\left[\varepsilon_i\leq\rho_x(a,\theta,\partial_a g_x)\right]\right|^2\right]\right)^{1/2} = 0. \quad (30)$$

To do this, take any $\|(a'-a,\theta'-\theta,\partial_a g_x'-\partial_a g_x)\|<\delta_N$, then we have for all x

$$\begin{aligned}
\left|\rho_x(a',\theta',\partial_a g_x')-\rho_x(a,\theta,\partial_a g_x)\right| &\leq C_0\left\{\|(a'-a,\theta'-\theta)\| + \|\partial_a g_x'-\partial_a g_x\|_{\mathcal{G}}\right\} \\
&\quad + o\left(\|(a'-a,\theta'-\theta)\|^2 + \|\partial_a g_x'-\partial_a g_x\|_{\mathcal{G}}^2\right) \\
&\leq C_0\delta_N + o(\delta_N),
\end{aligned}$$

this follows from Taylor's theorem in Banach Space since ρ_x is twice Fréchet differentiable, see Chapter 4 of Zeidler (1986). Ignoring the smaller order term, this implies

$$\begin{aligned}
\rho_x(a,\theta,\partial_a g_x) - C_0\delta_N &\leq \rho_x(a',\theta',\partial_a g_x') \leq \rho_x(a,\theta,\partial_a g_x) + C_0\delta_N, \\
\rho_x(a,\theta,\partial_a g_x) - C_0\delta_N &\leq \rho_x(a,\theta,\partial_a g_x) \leq \rho_x(a,\theta,\partial_a g_x) + C_0\delta_N.
\end{aligned}$$

Combining the inequalities above, it follows that $\sup_{\|(a'-a,\theta'-\theta,\partial_a g_x'-\partial_a g_x)\|<\delta_N}|\mathbf{1}\left[\varepsilon_i\leq\rho_x(a',\theta',\partial_a g_x')\right]$ $-\mathbf{1}\left[\varepsilon_i\leq\rho_x(a,\theta,\partial_a g_x)\right]|$ is bounded above by $\mathbf{1}\left[\rho_x(a,\theta,\partial_a g_x)-C_0\delta_N<\varepsilon_i\leq\rho_x(a,\theta,\partial_a g_x)+C_0\delta_N\right]$.

This majorant takes value 1 with probability $Q(\rho_x(a,\theta,\partial_a g_x)+C_0\delta_N)-Q(\rho_x(a,\theta,\partial_a g_x)-C_0\delta_N)$ and zero otherwise, then by Lipschitz continuity of Q, (30) holds as required. Since $A\times\Theta$ is a compact Euclidean set it has a known covering number. For $\mathcal{G}_x^{(1)}$, since $\mathcal{G}_x\subset C^2(A)$, a class of continuously twice differentiable function on A, we can apply Corollary 2.7.3 of VW so that $\int_0^\infty\sqrt{\log N\left(\varepsilon,\mathcal{G}_x^{(1)},\|\cdot\|_{\mathcal{G}}\right)}d\varepsilon<\infty$, together with $L^2(Q)-$continuity of $\mathbf{1}[\cdot\leq\rho_x(a,\theta,\partial_a g_x)]$, as shown in the proof of Theorem 3 (part (ii)) in Chen et al. (2003), $\mathcal{F}_x$ is $Q-$Donsker for each x.■

Proof of Lemma 8. For all x, $\mathcal{F}_x$ is $Q-$Donsker and is locally uniformly $L^2(Q)-$continuous with respect to $a,\theta,\partial_a g_x$, as described in (30), Lemma 1 of Chen et al. (2003) implies that the stochastic equicontinuity also holds with respect to the parameters that index the functions in $\mathcal{F}_x$.■

Proof of Lemma 9. For any a and x write

$$\sqrt{N}\left(\widehat{F}_{A|X}(a|x)-F_{A|X}(a|x)\right)=\mathfrak{F}_{1,N}(a,x)+\mathfrak{F}_{2,N}(a,x),$$

where

$$\begin{aligned}\mathfrak{F}_{1,N}(a,x)&=\frac{1}{T\widehat{p}_X(x)}\times\frac{1}{\sqrt{N}}\sum_{i=1,t=1}^{N,T}\left(\mathbf{1}[a_{it}\leq a,x_{it}=x]-F_{A,X}(a,x)\right),\\ \mathfrak{F}_{2,N}(a,x)&=-\frac{\sqrt{T}F_{A|X}(a|x)}{\widehat{p}_X(x)}\times\sqrt{N}\left(\widehat{p}_X(x)-p_X(x)\right).\end{aligned}$$

Define $\mathcal{C}_a=\{y_a\in\mathbb{R}:y_a\leq a\}$, then $\mathcal{C}=\bigcup_{a\in A}\mathcal{C}_a$ a classical VC-class of sets, for the definition VC-class of sets see Pollard (1990). Since X is finite, it is also necessarily a VC-class of sets. Then for each x, $\frac{1}{\sqrt{NT}}\sum_{i=1,t=1}^{N,T}(\mathbf{1}[a_{it}\leq\cdot,x_{it}=x]-F_{A,X}(\cdot,x))$ converges weakly to some tight Gaussian process in $l^\infty(A)$ since $\mathcal{C}\times X$ is VC in $A\times X$, by Lemma 2.6.17 in VW, and VC-classes of functions is a Donsker class, see also Type I classes of Andrews (1994). With an abuse of notation, for each x let $\frac{1}{\widehat{p}_X(x)}$ $(\frac{1}{p_X(x)})$ also denote a random element that takes value in $l^\infty(A)$ such that the sample path of $\frac{1}{\widehat{p}_X(x)}$ $(\frac{1}{p_X(x)})$ is constant over A. By standard LLN $\frac{1}{\widehat{p}_X(x)}\xrightarrow{p}\frac{1}{p_X(x)}$ and it follows by Slutsky's theorem that $\mathfrak{F}_{1,N}(\cdot,x)$ converges weakly to a random element in $l^\infty(A)$. In particular, the limit of $\mathfrak{F}_{1,N}(\cdot,x)$ is also a tight Gaussian process. From the finite dimensional (fidi) weak convergence, Gaussianity is clearly preserved if we replace $\frac{1}{\widehat{p}_X(x)}$ by $\frac{1}{p_X(x)}$, but since $\widehat{p}_X(x)-p_X(x)=o_p(1)$ the remainder term from the expansion $\frac{1}{\widehat{p}_X(x)}-\frac{1}{p_X(x)}$ can be used to construct a random element that converges to zero in probability on A, so by an application of Slutsky's theorem Gaussianity is preserved. Tightness trivially follow since the multiplication of $\frac{1}{\widehat{p}_X(x)}$ does not affect the asymptotic tightness of $\{\frac{1}{\sqrt{NT}}\sum_{i=1,t=1}^{N,T}(\mathbf{1}[a_{it}\leq\cdot,x_{it}=x]-F_{A,X}(\cdot,x))\}$. Since the only random component of $\mathfrak{F}_{2,N}(\cdot,x)$ is from $\sqrt{NT}(\widehat{p}_X(x)-p_X(x))$, which is a finite dimensional random variable, then a similar argument to the one used previously can trivially show that $\mathfrak{F}_{2,N}(\cdot,x)$ must also converge to a Gaussian process which is tight $l^\infty(A)$, where tightness follows from the (equi-)continuity of

$F_{A|X}(a|x)$ on A. Therefore $\sqrt{N}\left(\widehat{F}_{A|X=x} - F_{A|X=x}\right)$ must converge to a tight Gaussian process in $l^{\infty}(A)$ for all x since asymptotic tightness is closed under finite addition and, in this case, it is easy to see that Gaussianity is also closed under the sum.■

PROOF OF LEMMA 10. By MVT, for all a and x

$$\begin{aligned} & F_{A|X}(a|x;\theta_0,\partial_a\widehat{g}(\cdot,\theta_0)) - F_{A|X}(a|x;\theta_0,\partial_a g_{0,x}(\cdot,\theta_0)) \\ = \;& q\left(\overline{\rho}_x(a,\theta_0,\partial_a g_{0,x}(\cdot,\theta_0))\right)\left(\rho_x(a,\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0)) - \rho_x(a,\theta_0,\partial_a g_{0,x}(\cdot,\theta_0))\right), \end{aligned}$$

where $\overline{\rho}_x(a,\theta_0,\partial_a g_{0,x}(\cdot,\theta_0))$ is some intermediate value between $\rho_x(a,\theta_0,\partial_a\widehat{g}_x(\cdot,\theta_0))$ and $\rho_x(a,\theta_0,\partial_a g_{0,x}(\cdot,\theta_0))$. Since $\rho_x(a,\theta_0,\partial_a g_x)$ is twice Fréchet continuously differentiable on A at $\partial_a g_{0,x}(\cdot,\theta_0)$, using the linearization assumption, the argument analogous to Lemma 9 with Slutsky theorem can be used to complete the proof.■

References

[1] Ackerberg, D., L. Benkard, S. Berry, and A. Pakes (2005): "Econometric Tools for Analyzing Market Outcome," *Handbook of Econometrics*, vol. 6, eds. J. Heckman and E. Leamer. North-Holland.

[2] Aguirregabiria, V. (2008): "Another Look at the Identification of Dynamic Discrete Decision Processes: An Application to Retirement Behavior," *Journal of Business and Economic Statistics*, forthcoming.

[3] Aguirregabiria, V., and P. Mira (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, **75**, 1-53.

[4] Aguirregabiria, V., and P. Mira (2008): "Dynamic Discrete Choice Structural Models: A Survey," *Journal of Econometrics,* forthcoming.

[5] Andrews, D.W.K. (1994): "Empirical Processes Methods in Econometrics," *Handbook of Econometrics*, vol. 4, eds. R.F. Engle and D. McFadden. North-Holland.

[6] Andrews, D.W.K. (2005): "Higher-order Improvements of the Parametric Bootstrap for Markov Processes," *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, eds. D.W.K. Andrews and J.H. Stock. Cambridge Press.

[7] Arcidiacono, P. and R.A. Miller (2008): "CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," *Working Paper.*

[8] Bajari, P., C. L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, **75**, 1331-1370.

[9] Bajari, P., V. Chernozhukov, H. Hong and D. Nekipelov (2009): "Semiparametric Estimation of a Dynamic Game of Incomplete Information," *Working paper.*

[10] Berry, S., and A. Pakes (2002): "Estimation from the Optimality Conditions for Dynamic Controls," *Mimeo.*

[11] Bosq, D. (1996): *Nonparametric Statistics for Stochastic Processes*, Springer.

[12] Carrasco, M., and J.P. Florens (2000): "Generalization Of Gmm To A Continuum Of Moment Conditions," *Econometric Theory*, **16**, 797-834.

[13] Chen, X., O.B. Linton and I. van Keilegom (2003): "Estimation of Semiparametric Models when the Criterion Function is not Smooth," *Econometrica*, **71**, 1591-1608.

[14] Gowrisankaran, G., C. Lucarelli, P. Schmidt-Dengler and R. Town (2010): "*Government policy and the dynamics of market structure: Evidence from Critical Access Hospitals*, " *Working Paper.*

[15] Hansen, L.P. and K.J. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, **50**, 1269-1286.

[16] Heckman, J. and S. Navarro (2007): "Dynamic Discrete Choice and Dynamic Treatment Effects," *Journal of Econometrics*, **136**, 341-396.

[17] Hong, H. and M. Shum (2009): "Pairwise-Difference Estimation of a Dynamic Optimization Model," forthcoming, *Review of Economics Studies.*

[18] Hotz, V., and R.A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, **60**, 497-531.

[19] Hotz, V., R.A. Miller, S. Sanders and J. Smith (1994): "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies*, **61**, 265-289.

[20] Hu. Y. and M. Shum (2009): "Nonparametric Identification of Dynamic Models with Unobserved State Variables," *Working Paper.*

[21] Jofre-Bonet, M. and M. Pesendorfer (2003): "Estimation of a Dynamic Auction Game," *Econometrica*, **71**, 1443-1489.

[22] Kasahara, H. and K. Shimotsu (2008a): "Pseudo-Likelihood Estimation and Bootstrap Inference for Structural Discrete Markov Decision Models," *Journal of Econometrics*, **146**, 92-106.

[23] Kasahara, H. and K. Shimotsu (2008b): "Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices," *Econometrica*, **77**, 135-175..

[24] Koul, H.L. (2002): *Weighted Empiricals in Dynamic Nonlinear Models*, Lecture Notes in Statistics, Springer, New York.

[25] Lee, A.J. (1990): *U-Statistics: Theory and Practice*, Marcel Dekker, New York.

[26] Linton, O.B., and E. Mammen (2005): "Estimating Semiparametric Arch(∞) Models by Kernel Smoothing Methods," *Econometrica*, **73**, 771-836.

[27] Magnac, T. and D. Thesmar (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, **70**, 801-816.

[28] Masry, E. (1996): "Multivariate Local Polynomial Regression for Time Series: Uniform Strong Consistency and Rates," *Journal of Time Series Analysis* **17**, 571–599.

[29] Newey, W.K. & D. McFadden, 1994. "Large Sample Estimation and Hypothesis Testing," *Handbook of Econometrics*, vol. 4, eds. R.F. Engle and D. McFadden. North-Holland.

[30] Olley, S. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263-1297 .

[31] Pakes, A.(1994a): "Estimation of Dynamic Structural Models: Problems and Prospects Part II: Mixed Continuous-Discrete Control Models and Market Interactions," *Proceedings of the 6th World Congress of the Econometric Society, Cambridge University Press.*

[32] Pesendorfer, M., and P. Schmidt-Dengler (2008): "Asymptotic Least Squares Estimator for Dynamic Games," *Reviews of Economics Studies*, **75**, 901-928.

[33] Pollard, D. (1985): "New Ways to Prove Central Limit Theorems," *Econometric Theory*, **1**, 295-314.

[34] Pollard, D. (1990): *Empirical Processes: Theory and Applications*, NSF-CBMS Regional Conference Series in Probability and Statistics.

[35] Powell, J.L., J.H. Stock and T.M. Stoker (1989): "Semiparametric Estimation of Index Coefficients," *Econometrica,* **57**, 1403-1430.

[36] Ranga Rao, R. (1962): "Relations Between Weak and Uniform Convergence of Measures with Applications," Annals of Mathematical Statistics, **33**, 659-680.

[37] Robinson, P.M. (1988): "Root-N-consistent semiparametric regression," *Econometrica,* **56**, 931-954.

[38] Rust, J. (1987): "Optimal Replacement of GMC bus engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999-1033.

[39] Rust, J. (1994): "Estimation of Dynamic Structural Models: Problems and Prospects Part I: Discrete Decision Processes," *Proceedings of the 6th World Congress of the Econometric Society, Cambridge University Press.*

[40] Srisuma, S. (2010): "Estimation of Structural Optimization Models: A Note on Identification," *STICERD Discussion Paper Series, EM/2010/547.*

[41] Srisuma, S., and O. B. Linton (2009): "Semiparametric Estimation of Markov Decision Processes with Continuous State Space," *LSE, Working Paper.*

[42] Taussky, O. (1949): "A Recurring Theorem on Determinants," *American Mathematical Monthly*, **56**, 672–676.

[43] van der Vaart, A. W. and J. A. Wellner (1996): *Weak Convergence and Empirical Processes,* Springer-Verlag, New York.

[44] Wolfowitz, J. (1953): "Estimation by Minimum Distance Method," *Ann. Inst. Stat. Math.*, **5**, 9–23.

[45] Zeidler, E. (1986): *Nonlinear Functional Analysis and its Applications I: Fixed-Point Theorem,* Springer-Verlag, New York.

		$\widehat{\theta}_1^{UM}$					$\widehat{\theta}_1^{EM}$				
NT	ς	bias	mbias	std	iqr	mse	bias	mbias	std	iqr	mse
100	1/6	-0.0052	0.0126	0.1609	0.1337	0.0259	-0.0005	0.0180	0.1663	0.1349	0.0276
	1/7	0.0056	0.0262	0.1456	0.1270	0.0212	0.0143	0.0345	0.1448	0.1262	0.0212
	1/8	0.0138	0.0326	0.1370	0.1201	0.0190	0.0214	0.0402	0.1355	0.1181	0.0188
	static	0.2610	0.2610	0.0179	0.0176	0.0685	0.2666	0.2664	0.0179	0.0179	0.0714
500	1/6	-0.0034	-0.0030	0.0665	0.0663	0.0044	-0.0014	-0.0015	0.0665	0.0651	0.0044
	1/7	0.0024	0.0028	0.0627	0.0643	0.0039	0.0037	0.0039	0.0632	0.0627	0.0040
	1/8	0.0090	0.0094	0.0598	0.0619	0.0037	0.0103	0.0100	0.0590	0.0612	0.0036
	static	0.2608	0.2609	0.0073	0.0077	0.0681	0.2619	0.2619	0.0073	0.0076	0.0687
1000	1/6	-0.0016	-0.0009	0.0487	0.0443	0.0024	-0.0016	-0.0004	0.0485	0.0442	0.0024
	1/7	0.0020	0.0019	0.0464	0.0443	0.0022	0.0022	0.0017	0.0463	0.0451	0.0021
	1/8	0.0065	0.0062	0.0436	0.0432	0.0019	0.0066	0.0063	0.0438	0.0433	0.0020
	static	0.2607	0.2606	0.0052	0.0052	0.0680	0.2612	0.2611	0.0053	0.0052	0.0683

Table 1: $h_\varsigma = 1.06s(NT)^{-\varsigma}$ are the bandwidths used in the nonparametric estimation, s denotes the standard deviation of $\{a_{it}\}_{i=1,t=1}^{N,T}$.

		$\widehat{\theta}_2^{UM}$					$\widehat{\theta}_2^{EM}$				
NT	ς	bias	mbias	std	iqr	mse	bias	mbias	std	iqr	mse
100	1/6	0.0875	0.0413	0.2187	0.1708	0.0555	0.0901	0.0474	0.2122	0.1738	0.0532
	1/7	0.0787	0.0411	0.1998	0.1643	0.0461	0.0782	0.0399	0.1954	0.1621	0.0443
	1/8	0.0702	0.0357	0.1964	0.1614	0.0435	0.0709	0.0324	0.1970	0.1638	0.0438
	static	0.1288	0.1297	0.0319	0.0316	0.0176	0.1306	0.1333	0.0329	0.0328	0.0181
500	1/6	0.0412	0.0348	0.0933	0.0826	0.0104	0.0418	0.0340	0.0929	0.0851	0.0104
	1/7	0.0410	0.0323	0.0901	0.0787	0.0098	0.0405	0.0319	0.0902	0.0793	0.0098
	1/8	0.0391	0.0286	0.0870	0.0800	0.0091	0.0391	0.0296	0.0867	0.0792	0.0090
	static	0.1306	0.1308	0.0138	0.0137	0.0172	0.1310	0.1315	0.0138	0.0138	0.0173
1000	1/6	0.0334	0.0322	0.0653	0.0625	0.0054	0.0334	0.0304	0.0646	0.0604	0.0053
	1/7	0.0330	0.0307	0.0628	0.0586	0.0050	0.0322	0.0302	0.0635	0.0612	0.0051
	1/8	0.0335	0.0315	0.0604	0.0578	0.0048	0.0331	0.0303	0.0609	0.0564	0.0048
	static	0.1302	0.1304	0.0096	0.0095	0.0170	0.1304	0.1308	0.0096	0.0097	0.0171

Table 2: $h_\varsigma = 1.06s(NT)^{-\varsigma}$ are the bandwidths used in the nonparametric estimation, s denotes the standard deviation of $\{a_{it}\}_{i=1,t=1}^{N,T}$.