Journal of Econometrics 166 (2012) 320–341

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Semiparametric estimation of Markov decision processes with continuous state space[☆]

Sorawoot Srisuma [*], Oliver Linton

Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, CB3 9DD, United Kingdom

ARTICLE INFO

Article history:
Received 2 August 2010
Received in revised form
13 October 2011
Accepted 18 October 2011
Available online 24 October 2011

JEL classification:
C13
C14
C51

Keywords:
Discrete Markov decision models
Kernel smoothing semiparametric estimation
Well-posed inverse problem

ABSTRACT

We propose a general two-step estimator for a popular Markov discrete choice model that includes a class of Markovian games with continuous observable state space. Our estimation procedure generalizes the computationally attractive methodology of Pesendorfer and Schmidt-Dengler (2008) that assumed finite observable states. This extension is non-trivial as the policy value functions are solutions to some type II integral equations. We show that the inverse problem is well-posed. We provide a set of primitive conditions to ensure root-T consistent estimation for the finite dimensional structural parameters and the distribution theory for the value functions in a time series framework.

1. Introduction

The inadequacy of static frameworks to model economic phenomena led to the development of recursive methods in economics. The mathematical theory underlying discrete time modeling is dynamic programming developed by Bellman (1957); for a review of its prevalence in modern economic theory, see Stokey and Lucas (1989). In this paper we study the estimation of structural parameters and functionals thereof that underlie a class of Markov decision processes (MDP) with discrete controls and time in the infinite horizon setting. Such models are popular in applied work, in particular in labor and industrial organization. The econometrics involved can be seen as an extension of the classical discrete choice analysis to a dynamic framework.

Discrete choice modeling has a long established history in the structural analysis of behavioral economics. McFadden (1974) pioneered the theory and methods of analyzing discrete choice in a static framework. Rust (1987), using *additive separability* and *conditional independence* assumptions, shows that a class of dynamic discrete choice models can naturally preserve the familiar structure of discrete choice problems of the static framework. In particular, Rust proposed the Nested Fixed Point (NFP) algorithm to estimate his parametric model by the maximum likelihood method. However, in practice, this method can pose a considerable obstacle due to its requirement to repeatedly solve for the fixed point of some nonlinear map to obtain the value functions. The two-step approach of Hotz and Miller (1993) avoids the full solution method by relying on the existence of an inversion map between the normalized value functions and the (conditional) choice probabilities, which significantly reduces the computational burden relative to the NFP algorithm.

The two-step estimator of Hotz and Miller is central to several methodologies that followed, especially in the recent development of the estimation of dynamic games. A class of stationary infinite horizon Markovian games can be defined to include the MDP of interest as a special case. Various estimation procedures have been proposed to estimate the structural parameters of dynamic discrete action games: Pakes et al. (2004) and Aguirregabiria and Mira (2007), the latter builds on Aguirregabiria and Mira (2002), consider the two-step method of moments and pseudo maximum likelihood estimators respectively (these are included in the general class of asymptotic least squares

☆ We are grateful to the co-editor and three anonymous referees whose comments help improve the paper. We thank Xiaohong Chen, Philipp Schmidt-Dengler, and seminar participants at the 19th EC-Squared Conference on "Recent Advances in Structural Microeconometrics" in Rome, and Workshop in "Semiparametric and Nonparametric Methods in Econometrics" in Banff, for helpful comments. We thank the ERC for financial support. .

* Corresponding author. Tel.: +44 1223 335261; fax: +44 1223 335475.
E-mail addresses: sts25@cam.ac.uk (S. Srisuma), obl20@cam.ac.uk (O. Linton).

0304-4076/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2011.10.003

estimator defined by Pesendorfer and Schmidt-Dengler (2008)); Bajari et al. (2007) generalize the simulation-based estimators of Hotz et al. (1994) to the multiple agent setting. In both single and multiple agent settings, the aforementioned work and most research in the literature assumed that the observed state space is finite whenever the transition distribution of the observed state variables is not specified parametrically. One notable exception can be found in Altug and Miller (1998) who rely on a finite dependence assumption to estimate their problem semiparametrically. Although finite dependence can sometimes be motivated empirically (see Arcidiacono and Miller, 2010, for a definition and a discussion) it imposes a non-trivial restriction on the stochastic process that simplifies the theoretical and practical aspects of the estimation problem considerably. We do not make this assumption here.

In this paper we propose a simple two-step estimator that falls in the general class of semiparametric estimation discussed in Pakes and Olley (1995), and Chen et al. (2003). The criterion function will be based on some conditional moment restrictions that depend on the value function to be estimated in the preliminary step. Altug and Miller (1998) show that the value functions can be estimated by some finite-step ahead choice probabilities under finite dependence. However, without it, value functions are generally defined as solutions to linear integral equations of type II. The study of the statistical properties of solutions to integral equations falls under the growing research area on inverse problems in econometrics, see Carrasco et al. (2007) for a survey. We make use of the contraction property that occurs naturally in dynamic optimization problems to show that our problem is generally well-posed and obtain the uniform expansion for our kernel estimator that satisfies a type II integral equation. Our approach to estimate the solution to the integral equation is similar to the work of Linton and Mammen (2005) in their study of nonparametric ARCH models.

Our estimation strategy can be seen as a generalization of the unifying method of Pesendorfer and Schmidt-Dengler (2008) that allows for continuous components in the observable state space. The novel approach of Pesendorfer and Schmidt-Dengler relies on the attractive feature of the infinite time stationary model, where they write their ex-ante value function as the solution to a matrix equation. We show that the solving of an analogous linear equation, in an infinite dimensional space, is also a well-posed problem for both population and empirical versions (at least for large sample size). We note that an independent working paper of Bajari et al. (2008) also propose a sieve estimator for a closely related Markovian games, which allows for continuous observable state space. Therefore our methods are complementary in filling this gap in the literature. We use the local approach of kernel smoothing, under some easily interpretable primitive conditions, to provide explicit pointwise distribution theory of the infinite dimensional parameters that would otherwise be elusive with the series or spline expansion. Since the infinite dimensional parameters in MDP are the value functions, they may be of considerable interest in themselves. In addition, we explicitly work under time series framework and provide the type of primitive conditions required for the validity of the methodology.

The paper is organized as follows. Section 2 defines the MDP of interest, motivates and discusses the estimation strategy and the related linear inverse problem. Section 3 describes in detail the practical implementation of the procedure to obtain the feasible conditional choice probabilities. In Section 4, primitive conditions and the consequent asymptotic distribution are provided, the semiparametric profiled likelihood estimator is illustrated as a special case. Section 5 presents a small scale Monte Carlo experiment to study the finite sample performance of our estimator. Section 6 concludes.

2. Markov decision processes

We define our time homogeneous MDP and introduce the main model assumptions and notation used throughout the paper. The sources of the computational complexity for estimating MDP are briefly reviewed, there we focus on the representation of the value function as a solution to the policy value equation that can generally be written as an integral equation, in 2.2. We discuss the inverse problem associated with solving such integral equations in 2.3.

2.1. Definitions and assumptions

We consider a decision process of a forward looking agent who solves the following infinite horizon intertemporal problem. The random variables in the model are the control and state variables, denoted by a_t and s_t respectively. The control variable, a_t, belongs to a finite set of alternatives $A = \{1, \ldots, K\}$. The state variables, s_t, have support $S \subset \mathbb{R}^{L+K}$. At each period t, the agent observes s_t and chooses an action a_t in order to maximize her discounted expected utility. The present period utility is time separable and is represented by $u(a_t, s_t)$. The agent's action in period t affects the uncertain future states according to the (first order) Markovian transition density $p(ds_{t+1}|s_t, a_t)$. The next period utility is subjected to discounting at the rate $\beta \in (0, 1)$. Formally, for any time t, the agent is represented by a triple of primitives (u, β, F), who is assumed to behave according to an optimal decision rule, $\{\alpha_\tau(s_\tau)\}_{\tau=t}^{\infty}$, in solving the following *sequential problem*

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

Under some regularity conditions, see Bertsekas and Shreve (1978) and Rust (1994), Blackwell's Theorem and its generalization ensure the following important properties. First, there exists a stationary (time invariant) Markovian *optimal policy function* $\alpha : S \to A$ so that $\alpha(s_t) = \alpha(s_{t+\tau})$ for any $s_t = s_{t+\tau}$ and any t, τ, where

$$\alpha(s_t) = \arg\max_{a\in A} \{u(a, s_t) + \beta E[V(s_{t+1})|s_t, a_t = a]\}.$$

Secondly, the value function, defined in (1), is the unique solution to the *Bellman's equation*

$$V(s_t) = \max_{a\in A} \{u(a, s_t) + \beta E[V(s_{t+1})|s_t, a_t = a]\}. \tag{2}$$

We now introduce the following set of modeling assumptions.

Assumption M1 (*Conditional Independence*)**.** The transitional density has the following factorization: $p(dx_{t+1}, d\varepsilon_{t+1}|x_t, \varepsilon_t, a_t) = q(d\varepsilon_{t+1}|x_{t+1}) f_{X'|X,A}(dx_{t+1}|x_t, a_t)$, where the first moment of ε_t exists and its conditional distribution is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^K$, we denote its density by q.

The conditional independence assumption of Rust (1987) is fundamental in the current literature. It is a subject of current research on how to find a practical methodology that can relax this assumption; see Arcidiacono and Miller (2010) for an example. The continuity assumption on the distribution of ε_t ensures we can apply Hotz and Miller's inversion theorem.

Assumption M2. The support of $s_t = (x_t, \varepsilon_t)$ is $X \times \mathcal{E}$, where X is a compact subset of $\mathbb{R}^L$, in particular, $x_t = \left(x_t^c, x_t^d\right) \in X^C \times X^D$, and $\mathcal{E} = \mathbb{R}^K$.

In order to avoid a degenerate model, we assume that the state variables $s_t = (x_t, \varepsilon_t) \in X \times \mathbb{R}^K$ can be separated into two parts, which are observable and unobservable respectively to the econometrician; see Rust (1994) for various interpretations of the unobserved heterogeneity. Compactness of X is assumed for simplicity, in particular to X^C can be unbounded.

Assumption M3 (*Additive Separability*)**.** The per period payoff function $u : A \times X \times \mathcal{E} \to \mathbb{R}$ is additive separable w.r.t. unobservable state variables, $u(a_t, x_t, \varepsilon_t) = \pi(a_t, x_t) + \sum_{k=1}^{K} \varepsilon_{a_k,t} \mathbf{1}[a_t = k]$.

The combination of M1 and M3 allows us to set our model in the familiar framework of static discrete choice modeling.

Condition M2 relaxes the usual finite X assumption when no parametric assumption is assumed on $f_{X'|X,A}(dx_{t+1}|x_t, a_t)$. Otherwise Conditions M1–M3 are standard in the literature. For departures of this framework see the discussion in the survey of Aguirregabiria and Mira (2010) and the references therein. Henceforth Conditions M1–M3 will be assumed and later strengthened as appropriate.

2.2. Value functions

Similarly to the static discrete choice models, the choice probabilities play a central role in the analysis of the controlled process. There are two numerical aspects that we need to consider in the evaluation of the choice probabilities. The first are the multiple integrals that also arise in the static framework, where in practice many researchers avoid this issue via the use of conditional logit assumption of McFadden (1974). The second is regarding the value function—this is unique to the dynamic setup. To see precisely the problem we face, we first update the Bellman's equation (2) under the Assumptions M1–M3,

$$V(s_t) = \max_{a \in A} \left\{ \pi(a, x_t) + \varepsilon_{a,t} + \beta E[V(s_{t+1}) | x_t, a_t = a] \right\}.$$

Denoting the future expected payoff $E[V(s_{t+1})|x_t, a_t]$ by $g(a_t, x_t)$, and the choice specific value, net of $\varepsilon_{a,t}$, $\pi(a_t, x_t) + \beta g(a_t, x_t)$ by $v(a_t, x_t)$, the optimal policy function must satisfy

$$\alpha(x_t, \varepsilon_t) = a \Leftrightarrow v(a, x_t) + \varepsilon_{a,t} \geq v(a', x_t) + \varepsilon_{a',t} \text{ for } a' \neq a. \quad (3)$$

The conditional choice probabilities, $\{P(a|x)\}$, are then defined by

$$\begin{aligned} P(a|x) &= \Pr\left[v(a, x_t) + \varepsilon_{a,t} \geq v(a', x_t) + \varepsilon_{a',t} \text{ for } a' \neq a | x_t = x\right] \\ &= \int \mathbf{1}[\alpha(x, \varepsilon_t) = a]\, q(d\varepsilon_t | x). \end{aligned} \quad (4)$$

Even if we knew v, (4) will generally not have a closed form and the task of performing multiple integrals numerically can be non-trivial, see Hajivassiliou and Ruud (1994) for an extensive discussion on an alternative approach to approximating integrals. For some specific distributional assumptions on ε_t, for example using the popular i.i.d. extreme value of type I—we can avoid the multiple integrals as (4) has the well-known multinomial logit form

$$P(a|x) = \frac{\exp(v(a, x))}{\sum_{\widetilde{a} \in A} \exp(v(\widetilde{a}, x))}.$$

Note that, unlike in static models, the conditional logit model does not generally impose the undesirable I.I.A. property in the dynamic framework. The problem we want to focus on is the fact that we generally do not know v, as it depends on g that is defined through some nonlinear functional equation that we need to solve for. Next, we outline a characterization of the value function that motivates our approach to estimate g (and v).

The main insight to the simplicity of our methodology is motivated from the geometric series representation for the value function that is commonly used in dynamic programming theory (see Bertsekas and Shreve, 1978, Chapter 9). This type of representation has been frequently exploited, in one way or another, in the estimation of Markov decision problems with finite states, for example, see the survey of Miller (1997) for a discussion. Formally, one can define the value function corresponding to a particular stationary Markovian policy μ by

$$V(s_t; \mu) = E\left[\sum_{\tau=t}^{\infty} \beta^{\tau} u(\mu(s_\tau), s_\tau) \middle| s_\tau = s_t \right],$$

which is the solution to the following *policy value equation*

$$V(s_t; \mu) = u(\mu(s_t), s_t) + \beta E[V(s_{t+1}; \mu) | s_t].$$

In this paper we only consider values corresponding to the optimal policy, to reduce the notation, so we suppress the explicit dependence on the policy. Therefore, by definition of the optimal policy, the solution to (2) is also the solution to the following policy value equation

$$V(s_t) = u(\alpha(s_t), s_t) + \beta E[V(s_{t+1}) | s_t]. \quad (5)$$

If the state space S is finite, then V is a solution of a matrix equation above since the conditional expectation operator here can be represented by a stochastic transitional matrix. By the dominant diagonal theorem, the matrix representing $(I - \beta E[\cdot|s])_{s \in S}$ is invertible and (5) has a unique solution, solvable by direct matrix inversion or approximated by a geometric series (see the Neumann series below). The notion of simply inverting a matrix has an obvious appeal over Rust's fixed point iterations. In the infinite dimensional case, the matrix equation generalizes to an integral equation. In the presence of some unobserved state variables, we can also define the conditional value function as a solution to the following conditional policy value equation, taking conditional expectation on (5) w.r.t. x_t yields

$$\begin{aligned} &E[V(s_t) | x_t] \\ &= E[u(\alpha(s_t), s_t) | x_t] + \beta E[E[V(s_{t+1}) | s_t] | x_t] \\ &= E[u(\alpha(s_t), s_t) | x_t] + \beta E[E[V(s_{t+1}) | x_{t+1}] | x_t], \end{aligned}$$

where the last equality follows from the law of iterated expectations and M1. Noting that, again by M1, $g(a_t, x_t)$ can be written as $E[m(x_{t+1}) | x_t, a_t]$, where $m(x_t) = E[V(s_t) | x_t]$, then we have m as a solution to some particular integral equation of type II; more succinctly, m satisfies

$$m = r + \mathcal{L}m, \quad (6)$$

where r is the ex-ante expected immediate payoff given state x_t, namely $E[u(\alpha(s_t), s_t) | x_t = \cdot]$; and the integral operator $\mathcal{L}$ generates discounted expected next period values of its operands, e.g. $\mathcal{L}m(x) = \beta E[m(x_{t+1}) | x_t = x]$ for any $x \in X$. If we could solve (6) then we need another level of smoothing on m to obtain the choice specific value v. In particular, we can define g through the following linear transform

$$g = \mathcal{H}m, \quad (7)$$

where $\mathcal{H}$ is an integral operator that generates the choice specific expected next period values of its operands operator, e.g. $\mathcal{H}m(x, a) = \beta E[m(x_{t+1}) | x_t = x, a_t = a]$ for any $(x, a) \in X \times A$. Therefore we can write the choice specific value net of unobserved states in a linear functional notation as

$$v = \pi + \beta \mathcal{H}m. \quad (8)$$

In Section 3 we discuss in details on how to use the policy value approach to estimate the model implied transformed of the value functions and choice probabilities.

2.3. Linear inverse problems

Before we consider the estimation of v, we need to address some issues regarding the solution of integral equation (6). It is natural to ask the fundamental question whether our problem is well-posed, more specifically, whether the solution of such equation exists and if so, whether it is unique and stable. The study of the solution to such integral equations falls in the general framework of linear inverse problems.

The study of inverse problems is an old problem in applied mathematics. The type of inverse problems one commonly encounters in econometrics are integral equations. Carrasco et al. (2007) focused their discussion on ill-posed problems of integral equations of type I where recent works often needed regularizations in Hilbert Spaces to stabilize their solutions. Here we face an integral equation of type II, which is easier to handle, and in addition, the convenient structure of the policy value equations allows us to easily show that the problem is well-posed in a familiar Banach Space. We now define the normed linear space and the operator of interest, and proof this claim. We shall simply state relevant results from the theory of integral equations. For definitions, proofs and further details on integral equations, readers are referred to Kress (1999) and the references therein.

From the Riesz Theory of operator equations of the second kind with compact operators on a normed space, say $A : X \rightarrow X$, we know that $I - A$ is injective if and only if it is surjective, and if it is bijective, then the inverse operator $(I-A)^{-1} : X \rightarrow X$ is bounded. For the moment, suppose that X^D is empty, we will be working on the Banach space $(B, \|\cdot\|)$, where $B = C(X)$ is a space of continuous functions defined on the compact subset of $\mathbb{R}^L$, equipped with the sup-norm, i.e. $\|\phi\| = \sup_{x\in X} |\phi(x)|$. $\mathcal{L}$ is a linear map, $\mathcal{L} : C(X) \rightarrow C(X)$, such that, for any $\phi \in C(X)$ and $x \in X$,

$$\mathcal{L}\phi(x) = \beta \int_X \phi\left(x'\right) f_{X'|X}\left(dx'|x\right),$$

where $f_{X|X}(dx_{t+1}|x_t)$ denotes the conditional density of x_{t+1} given x_t.

In this case since we know the existence, uniqueness and stability of the solution to (6) are assured for any $r = \phi \in C(X)$ as we can show $\mathcal{L}$ is a contraction. To see this, take any $\phi \in C(X)$ and $x \in X$,

$$|\mathcal{L}\phi(x)| \leq \beta \int_X \left|\phi\left(x'\right)\right| f_{X'|X}\left(dx'|x\right) \leq \beta \sup_{x\in X} |\phi(x)|,$$

since the discounting factor $\beta \in (0, 1)$,

$$\|\mathcal{L}\phi\| \leq \beta \|\phi\| \Rightarrow \|\mathcal{L}\| \leq \beta < 1.$$

This implies that our inverse is well-posed. Further, the contraction property means we can represent the solution to (6) using the Neumann series:

$$m = (I - \mathcal{L})^{-1} r = \lim_{\mathcal{T}\longrightarrow\infty} \sum_{\tau=1}^{\mathcal{T}} \mathcal{L}^{\tau} r. \tag{9}$$

Therefore the infinite series representation of the inverse suggests one obvious way of approximating the solution to the integral equation which will converge geometrically fast to the true function. If X is countable, then $\mathcal{L}^{\tau}$ would be represented by a τ-step ahead transition matrix (scaled by β^{τ}). Note that the operator for the infinite dimensional case shares the analogous interpretation of τ-step ahead transition operator with discounting.

Since our problem is well-posed, then it is reasonable to expect that with sufficiently good estimates of $(r, \mathcal{L}, \mathcal{H})$, our estimated integral equation is also well-posed and will lead to (uniform) consistent estimators for (m, g, v). Our strategy is to use nonparametric methods to generate the empirical versions of (6) and (7), then use them to provide an approximate for v necessary for computing the choice probabilities.

3. Estimation

Given a time series $\{a_t, x_t\}_{t=1}^{T}$ generated from the controlled process of an economic agent represented by $\left(u_{\theta_0}, \beta, p\right)$, for some $\theta_0 \in \Theta$, where u_θ reflects the parameterization of π by θ. In this section we provide in details the procedure to estimate θ_0 as well as their corresponding conditional value functions. We based our estimation on the conditional choice probabilities. We define the model implied choice probabilities from a family of value functions, $\{V_\theta\}_{\theta\in\Theta}$, induced by underlying optimal policy that generates the data. In particular, for each θ, V_θ satisfies (cf. Eq. (5))

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

The policy value V_θ has the interpretation of a discounted expected value for an economic agent whose payoff function is indexed by θ but behaves optimally as if her structural parameter is θ_0. By definition of the optimal policy, V_θ coincides with the solution of a Bellman's equation in (2) when $\theta = \theta_0$. We then define the following (optimal) policy-induced equations to analogous to (6)–(8), respectively for each θ:

$$m_\theta = r_\theta + \mathcal{L}m_\theta, \tag{10}$$

$$g_\theta = \mathcal{H}m_\theta, \tag{11}$$

$$v_\theta = \pi_\theta + \beta\mathcal{H}m_\theta, \tag{12}$$

where r_θ is the ex-ante expected payoff given state x_t, namely $E[u_\theta(\alpha(s_t), s_t)|x_t = \cdot]$; and the integral operators $\mathcal{L}$ and $\mathcal{H}$ are the same as in Section 2.2. The functions m_θ, g_θ and v_θ are defined to satisfy the linear equation and transforms respectively. Naturally, for each $(a, x) \in A \times X$, $P_\theta(a|x)$ is then defined to satisfy

$$P_\theta(a|x) = \Pr\left[v_\theta(a, x_t) + \varepsilon_{a,t} \geq v_\theta\left(a', x_t\right) + \varepsilon_{a',t} \text{ for } a' \neq a | x_t = x\right],$$

which is analogous to (4).

The estimation procedure proceeds in two steps. In the first step, we nonparametrically compute estimates of the kernels of $\mathcal{L}$, $\mathcal{H}$ and, for each θ, estimate r_θ, which are then used to estimate m_θ by solving the empirical version of the integral equation (10) and estimate g_θ analogously from an empirical version of (11). The second step is the optimization stage, the model implied choice specific value functions are used to compute the choice probabilities that can be used to construct various objective functions to estimate the structural parameter θ_0.

3.1. Estimation of r_θ, $\mathcal{L}$ and $\mathcal{H}$

There are several decisions to be made to solve the empirical integral equation in (10). We need to first decide on the nonparametric method. We will focus on the method of kernel smoothing due to its simplicity of use as well as its well established theoretical grounding. Our nonparametric estimation of the conditional expectations will be based on the Nadaraya–Watson estimator. However, since we will be working on bounded sets, it is necessary to address the boundary effects. The treatment of the boundary issues is straightforward, the precise trimming condition is described in Section 4. So we will assume to work on a smaller space $X_T \subset X$ where $X_T = \left(X_T^C, X^D\right)$ denotes a set where the support of the uncountable component is some strict compact subset of X^C but increases to X^C in T. When allowing for discrete components we simply use the frequency approach, smoothing over the discrete components is also possible, see the monograph by Li and Racine (2006) for a recent update on this literature. We will also need to make a decision on how to define and interpolate the solution to the empirical version of (10) in practice. We discuss two asymptotically equivalent options for this latter choice, whether the size of the empirical integral equation does or does not depend on the sample size, as one may have a preference given the relative size of the number of observations.

We now define the nonparametric estimators, $(\widehat{r}_\theta, \widehat{\mathcal{L}}, \widehat{\mathcal{H}})$, of $(r_\theta, \mathcal{L}, \mathcal{H})$. Any generic density of a mixed continuous–discrete random vector $w_t = (w_t^c, w_t^d), f_W : \mathbb{R}^{l^C} \times \mathbb{R}^{l^D} \to \mathbb{R}^+$ for some positive integers l^C and l^D, is estimated as follows,

$$\widehat{f}_W(w^c, w^d) = \frac{1}{T}\sum_{t=1}^{T} K_h(w_t^c - w^c)\,\mathbf{1}[w_t^d = w^d],$$

where K is some user-chosen symmetric probability density function, h is a positive bandwidth and for simplicity independent of w^c. $K_h(\cdot) = K(\cdot/h)/h$ and if $l^C > 1$ then $K_h(w_t^c - w^c) = \prod_{l=1}^{l^C} K_{h_l}(w_{t,l}^c - w_l^c)$, $\mathbf{1}[\cdot]$ denotes the indicator function, namely $\mathbf{1}[\mathcal{A}] = 1$ if event $\mathcal{A}$ occurs and takes value zero otherwise. Similar to the product kernel, the contribution from a multivariate discrete variable is represented by products of indicator functions. The conditional densities/probabilities are estimated using the ratio of the joint and marginal densities. The local constant estimator of any generic regression function, $E[z_t|w_t = w]$ is defined by,

$$\widehat{E}[z_t|w_t = w] = \frac{\frac{1}{T}\sum_{t=1}^{T} z_t K_h(w_t^c - w^c)\,\mathbf{1}[w_t^d = w^d]}{\widehat{f}_w(w)}. \tag{13}$$

Since the conditional choice probabilities can be defined as a conditional expectation, in this paper we define $\widehat{P}(a|x) = \widehat{E}[\mathbf{1}[a_t = a]|x_t = x]$ for all $(a, x) \in A \times X$.

Estimation of r_θ For any $x \in X_T$,

$$\begin{aligned} r_\theta(x) &= E[u_\theta(a_t, x_t, \varepsilon_t)|x_t = x] \\ &= E[\pi_\theta(a_t, x_t)|x_t = x] + E[\varepsilon_{a_t}|x_t = x] \\ &= \rho_{1,\theta}(x) + \rho_2(x). \end{aligned}$$

The first term can be estimated by

$$\widehat{\rho}_{1,\theta}(x) = \sum_{a\in A}\widehat{P}(a|x)\,\pi_\theta(a, x), \tag{14}$$

or, alternatively, the Nadaraya–Watson estimator $\widetilde{\rho}_{1,\theta}(x) = \widehat{E}[\pi_\theta(a_t, x_t)|x_t = x]$. We also comment that it might be more convenient to use $\widehat{\rho}_{1,\theta}$ over $\widetilde{\rho}_{1,\theta}$, as we shall see, since the nonparametric estimates for the choice probabilities are required to estimate ρ_2.

The conditional mean of the unobserved states, ρ_2, is generally non-zero due to selectivity. By Hotz and Miller's inversion theorem, we know ρ_2 can be expressed as a known smooth function of the choice probabilities. For example, the i.i.d. type I extreme value errors assumption will imply that

$$\rho_2(x) = \gamma + \sum_{a\in A} P(a|x)\log(P(a|x)), \tag{15}$$

where γ is the Euler's constant. An estimator of ρ_2 can therefore be obtained by plugging in a nonparametric estimator of the choice probabilities. Our procedure is not restricted to the conditional logit assumption. Although other distributional assumption will generally not provide a closed form expression for ρ_2 in $\{P(a|x)\}$, it can be computed for any $(a, x) \in A \times X$, for example see Pesendorfer and Schmidt-Dengler (2003) who assume the unobserved states are i.i.d. standard normals. Note also that ρ_2 is independent of θ as the distribution of ε_t is assumed to be known.

Estimation of $\mathcal{L}$ and $\mathcal{H}$

For the ease of notation let us suppose X^D is empty. For the integral operators $\mathcal{L}$ and $\mathcal{H}$, if we would like to use the numerical integration to approximate the integral, we only need to provide the nonparametric estimators of their kernels, respectively, $\widehat{f}_{X'|X}(dx_{t+1}|x_t)$ and $\widehat{f}_{X'|X,A}(dx_{t+1}|x_t, a_t)$.

For any $\phi \in C(X_T)$, the empirical operators are defined as,

$$\widehat{\mathcal{L}}\phi(x) = \int_{X_T}\phi(x')\widehat{f}_{X'|X}(dx'|x), \tag{16}$$

$$\widehat{\mathcal{H}}\phi(x, a) = \int_{X_T}\phi(x')\widehat{f}_{X'|X,A}(dx'|x, a). \tag{17}$$

So $\widehat{\mathcal{L}}$ and $\widehat{\mathcal{H}}$ are linear operators on the Banach space of continuous functions on X_T with range $C(X_T)$ and $C(X_T \times A)$ respectively under sup-norm. Alternatively, we could use the Nadaraya–Watson estimator, defined in (13), to estimate the operators, i.e, $\widetilde{\mathcal{L}}\phi(x) = \widehat{E}[\phi(x_{t+1})|x_t = x]$ and $\widetilde{\mathcal{H}}\phi(x, a) = \widehat{E}[\phi(x_{t+1})|x_t = x, a_t = a]$. This approach may be more convenient when sample size is relatively small, and we want to solve the empirical version of (10) by using purely nonparametric methods for interpolation, where we could use the local linear estimator to address the boundary effects.

Note that, if X is finite then the integrals in (16) and (17) will be defined with respect to discrete measures, then $(\widehat{\mathcal{L}}, \widehat{\mathcal{H}})$ and $(\widetilde{\mathcal{L}}, \widetilde{\mathcal{H}})$ can be equivalently represented by the same stochastic matrices.

3.2. Estimation of m_θ, g_θ and v_θ

We first describe the procedure used in Linton and Mammen (2005), by using $(\widehat{\mathcal{L}}, \widehat{\mathcal{H}})$, to solve the empirical integral equation. We define $\widehat{m}_\theta$ as any sequence of random functions defined on X_T that approximately solves $\widehat{m}_\theta = \widehat{r}_\theta + \widehat{\mathcal{L}}\widehat{m}_\theta$. Formally,

We assume that $\widehat{m}_\theta$ is any random sequence of functions that satisfy

$$\sup_{\theta\in\Theta, x\in X_T}\left|\left(I - \widehat{\mathcal{L}}\right)\widehat{m}_\theta(x) - \widehat{r}_\theta(x)\right| = o_p\left(T^{-1/2}\right). \tag{18}$$

Therefore $\widehat{m}_\theta$ only has to "nearly" solve the integral equation. Pakes and Pollard (1985) allow for such flexibility in defining their simulation estimator and show the approximation error is negligible asymptotically. We shall make use of this approximation approach to defining many of our parameters in this paper. In practice, we solve the integral equation on a finite grid of points, which reduces it to a large linear system. Next we use $\widehat{m}_\theta$ to define $\widehat{g}_\theta$, specifically we define $\widehat{g}_\theta$ as any random sequence of functions that satisfy

$$\sup_{\theta\in\Theta, a\in A, x\in X_T}\left|\widehat{g}_\theta(a, x) - \widehat{\mathcal{H}}\widehat{m}_\theta(x, a)\right| = o_p\left(T^{-1/2}\right). \tag{19}$$

Once we obtain $\widehat{g}_\theta$, the estimator of v_θ is defined by

$$\sup_{\theta\in\Theta, a\in A, x\in X_T}\left|\widehat{v}_\theta(a, x) - \pi_\theta(a, x) - \beta\widehat{g}_\theta(a, x)\right| = o_p\left(T^{-1/2}\right). \tag{20}$$

For illustrational purposes, ignoring the trimming factors, we will assume that $X = [\underline{x}, \overline{x}] \subset \mathbb{R}$.

For any integrable function ϕ on X, define $J(\phi) = \int \phi(x)\,dx$. Given an ordered sequence of n nodes $\{x_{j,n}\} \subset [a, b]$, and a corresponding sequence of weights $\{\varpi_{j,n}\}$ such that $\sum_{j=1}^{n}\varpi_{j,n} = b - a$, a valid integration rule would satisfy

$$\lim_{n\to\infty} J_n(\phi) = J(\phi) \quad \text{where } J_n(\phi) = \sum_{j=1}^{n}\varpi_{j,n}\phi(x_{j,n}),$$

for example Simpson's rule and Gaussian quadrature both satisfy this property for smooth ϕ. Therefore the empirical version of (10) can be approximated for any $x \in [a, b]$ by

$$\widehat{m}_\theta(x) = \widehat{r}_\theta(x) + \beta\sum_{j=1}^{n}\varpi_{j,n}\widehat{f}_{X'|X}(x_{j,n}|x)\,\widehat{m}_\theta(x_{j,n}). \tag{21}$$

So the desired solution that approximately solves the empirical integral equation will satisfy the following equation at each node $\{x_{j,n}\}$,

$$\widehat{m}_\theta\left(x_{i,n}\right)=\widehat{r}_\theta\left(x_{i,n}\right)+\beta\sum_{j=1}^{n}\varpi_{j,n}\widehat{f}_{X'|X}\left(x_{j,n}|x_{i,n}\right)\widehat{m}_\theta\left(x_{j,n}\right).$$

This is equivalent to solving a system of n equations with n variables, the linear system above can be written in a matrix notation as

$$\widehat{\mathbf{m}}_\theta=\widehat{\mathbf{r}}_\theta+\widehat{\mathbf{L}}\widehat{\mathbf{m}}_\theta, \tag{22}$$

where $\widehat{\mathbf{m}}_\theta = \left(\widehat{m}_\theta\left(x_{1,n}\right),\ldots,\widehat{m}_\theta\left(x_{n,n}\right)\right)^\top, \widehat{\mathbf{r}}_\theta = (\widehat{r}_\theta(x_{1,n}),\ldots,\widehat{r}_\theta(x_{n,n}))^\top$, I_n is an identity matrix of order n and $\widehat{\mathbf{L}}$ is a square n matrix such that $(\widehat{\mathbf{L}})_{ij} = \beta\varpi_{j,n}\widehat{f}_{X'|X}\left(x_{j,n}|x_{i,n}\right)$. Since $\widehat{f}_{X'|X}(\cdot|x)$ is a proper density for any x, with a sufficiently large n, $(I_n-\widehat{\mathbf{L}})$ is invertible by the dominant diagonal theorem. So there is a unique solution to the system (22) for a given $\widehat{\mathbf{r}}_\theta$. In practice we have a variety of ways to solve for $\widehat{\mathbf{m}}_\theta$ with one obvious candidate being the successive approximation implied by (9). Once we obtain $\widehat{\mathbf{m}}_\theta$, we can approximate $\widehat{m}_\theta(x)$ for any $x\in X$ by substituting $\widehat{\mathbf{m}}_\theta$ into the RHS of (21). This is known as the Nyström interpolation. We need to approximate another integral to estimate g_θ. This could be done using the conventional method of kernel regression as discussed in Section 3.1, or by appropriately selecting sequences of n' nodes $\left\{x'_{j,n'}\right\}$ and weights $\left\{\varpi'_{j,n'}\right\}$ so that

$$\widehat{g}_\theta(a,x)=\sum_{j=1}^{r}\varpi'_{j,n'}\widehat{f}_{X'|X,A}\left(x'_{j,n'}|x,a\right)\widehat{m}_\theta\left(x'_{j,n'}\right),$$

where the computation for this last linear transform is trivial. See Judd (1998) for a more extensive review of the methods and issues of approximating integrals and also the discussion of iterative approaches in Linton and Mammen (2003) for large grid sizes.

Alternatively, we can form a matrix equation of size $T-1$,

$$\widetilde{\mathbf{m}}_\theta=\widetilde{\mathbf{r}}_\theta+\widetilde{\mathbf{L}}\widetilde{\mathbf{m}}_\theta,$$

to estimate Eq. (10) at the observed points with the t-th element. For each t, let

$$\widetilde{m}_\theta\left(x_t\right)=\widetilde{r}_\theta\left(x_t\right)+\beta\frac{\frac{1}{T-1}\sum_{\tau=1}^{T-1}\widetilde{m}_\theta\left(x_{t+1}\right)K_h\left(x_\tau-x_t\right)}{\frac{1}{T-1}\sum_{\tau=1}^{T-1}K_h\left(x_\tau-x_t\right)}.$$

By the dominant diagonal theorem, the matrix equation above always has a unique solution for any $T\geq 2$. Once solved, the estimators of $\widetilde{m}_\theta$ can be interpolated by

$$\widetilde{m}_\theta(x)=\widetilde{r}_\theta(x)+\beta\widehat{E}\left[m\left(x_{t+1}\right)|x_t=x\right],$$

for any $x\in X_T$. Similarly, $\widetilde{g}_\theta$ and $\widetilde{v}_\theta$ can be estimated nonparametrically without introducing any additional numerical error. Clearly, the more observation we have, the latter method will be more difficult as dimension of the matrix representing $\widetilde{\mathbf{L}}$ is large whilst the grid points for the former empirical equation are user-chosen.

3.3. Estimation of θ

By construction, when $\theta=\theta_0$, the model implied conditional choice probability P_θ coincides with the underlying choice probabilities defined in (4). Therefore one natural estimator for the finite dimensional structural parameters can be obtained by maximizing a likelihood criterion. Define

$$\begin{aligned}Q_T(\theta)&=\frac{1}{T}\sum_{t=1}^{T}\mathfrak{c}_{t,T}\ell\left(a_t,x_t;\theta,g_\theta\right);\\ \widehat{Q}_T(\theta)&=\frac{1}{T}\sum_{t=1}^{T}\mathfrak{c}_{t,T}\ell\left(a_t,x_t;\theta,\widehat{g}_\theta\right),\end{aligned}\tag{23}$$

where $\ell\left(a_t,x_t;\theta,g_\theta\right)$ and $\ell\left(a_t,x_t;\theta,\widehat{g}_\theta\right)$ denote $\log P_\theta\left(a_t|x_t\right)$ and $\log\widehat{P}_\theta\left(a_t|x_t\right)$ respectively. Here $\left\{\mathfrak{c}_{t,T}\right\}$ is a triangular array of trimming factors, more discussion on this can be found in Section 4. In practice, we replace $P_\theta(a|x)$ by

$$\begin{aligned}\widehat{P}_\theta(a|x) = \Pr\big[&\widehat{v}_\theta\left(a,x_t\right)+\varepsilon_{a,t}\geq\widehat{v}_\theta\left(a',x_t\right)+\varepsilon_{a',t}\\ &\text{for } a'\neq a|x_t=x\big],\end{aligned}$$

where $\widehat{v}_\theta$ satisfies condition (20). Of particular interest is the special case of the conditional logit framework, as discussed in Section 2, where we have

$$\widehat{P}_\theta(a|x)=\frac{\exp\left(\widehat{v}_\theta(a,x)\right)}{\sum_{\widetilde{a}\in A}\exp\left(\widehat{v}_\theta(\widetilde{a},x)\right)}.$$

Therefore $\widehat{Q}_T$ denotes the feasible objective function, which is identical to Q_T when the infinite dimensional component $\widehat{v}_\theta$ is replaced by v_θ. We define our maximum likelihood estimator, $\widehat{\theta}$, to be any sequence that satisfies the following inequality

$$\widehat{Q}_T(\widehat{\theta})\geq\sup_{\theta\in\Theta}\widehat{Q}_T(\theta)-o_p\left(T^{-1/2}\right). \tag{24}$$

Alternatively, a class of criterion functions can be generated from the following conditional moment restrictions

$$E\left[\mathbf{1}\left[a_t=a\right]-P_\theta\left(a|x_t\right)|x_t\right]=0\quad\text{for all } a\in A \text{ when } \theta=\theta_0.$$

Note that these moment conditions are the infinite dimensional counterparts (with respect to the observable states) of Eq. (18) in Pesendorfer and Schmidt-Dengler (2008) for a single agent problem.

There are general large sample theories of profiled semiparametric estimators available that treat the estimators defined in our models. In particular, the work of Pakes and Olley (1995) and Chen et al. (2003) provide high level conditions for obtaining root$-T$ consistent estimators are directly applicable. The latter is a generalization of the work by Pakes and Pollard (1985), who provided the asymptotic theory when the criterion function is allowed to be non-smooth, which may arise if we use simulation methods to compute the multiple integral of (4), to the semiparametric framework. In Section 4, as an illustration, we derive the asymptotic distribution of the semiparametric likelihood estimator under a set of weak conditions in the conditional logit framework.

We end this section with some brief comments regarding the computational aspects. As in the case when X is finite (cf. Pesendorfer and Schmidt-Dengler, 2008), the nonparametric estimators of $(r_\theta,\mathcal{L},\mathcal{H})$ have closed form and are easy to compute even with large dimensions. Therefore the solving of the empirical integral equation, in Eq. (22), to obtain $\widehat{m}_\theta$ reduces to inverting a large matrix that approximates $(I-\mathcal{L})$ that only needs to be done once since it is independent of θ. The estimators of $(m_\theta,g_\theta,v_\theta)$ can then be obtained trivially for any θ by simple matrix multiplications. There is also a computational advantage in specifying π_θ to be linear in θ for estimating the conditional value function. Bajari et al. (2007, Section 3.3.1, p. 1343) discuss this for their forward simulation methodology. However, their idea is not methodology specific and is relevant for any value function that satisfies a linear equation like (22). More specifically, if $\pi_\theta=\theta^\top\pi_0$ for a vector of known functions π_0 then $r_\theta=\theta^\top r_0+\rho_2$, where $r_0(\cdot)=\sum_{a\in A}P(a|\cdot)\pi_0(a,\cdot)$. Utilizing the fact that the inverse of $(I-\mathcal{L})$ is a linear operator we have $m_\theta=\theta^\top(I-\mathcal{L})^{-1}r_0+(I-\mathcal{L})^{-1}\rho_2$, hence the estimates of $(I-\mathcal{L})^{-1}r_0$ and $(I-\mathcal{L})^{-1}\rho_2$ only need to be computed once for the optimization procedure.

4. Distribution theory

In this section we provide a set of primitive conditions and derive the distribution theory for the estimators $\widehat{\theta}$, as defined in

(24), and $(\widehat{m}_\theta, \widehat{g}_\theta)$ as defined in (18) and (19) respectively when the unobserved state variables is distributed as i.i.d. extreme value of type I. This distributional assumption is the most commonly used in practice as it yields closed-form expressions for the choice probabilities. We also restrict the dimensionality of X^C to be a subset of $\mathbb{R}$, the reason being this is the scenario that applied researchers may prefer to work with. These specifics do not limit the usefulness of the primitives provided. For other estimation criteria, since two-step estimation problems of this type can be compartmentalized into nonparametric first stage and optimization in the second stage, the primitives below will be directly applicable. In particular, the discussions and results in 4.1 are independent of the choice of the objective functions chosen in the second stage. There might be other intrinsically continuous observable state variables that require discretizing but with increasing dimension in X^C, the practitioners will need to employ higher order kernels and/or undersmooth in order to obtain the parametric rate of convergence for the finite structural parameters, adaptation of the primitives are straightforward and will be discussed accordingly.

4.1. Infinite dimensional parameters

The relevant large sample properties for the nonparametric first stage, under the time series framework, for the pointwise results see the results of Roussas (1967, 1969), Rosenblatt (1956) and Robinson (1983). Roussas first provided central limit results for kernel estimates of Markov sequences, Rosenblatt established the asymptotic independence and Robinson generalized such results to the α-mixing case. The uniform rates have been obtained for the class of polynomial estimators by Masry (1996), in particular, our method is closely related to the recent framework of Linton and Mammen (2005) who obtained the uniform rates and pointwise distribution theory for the solution of a linear integral equation of type II.

We begin with some primitives. In addition to M1–M3, they are not necessary and only sufficient but they are weak enough to accommodate most of the existing empirical works in applied labor and industrial organization involving estimation of MDP.

We denote the strong mixing coefficient as

$$\alpha(k) = \sup_{t\in\mathbb{N}} \sup_{A\in\mathcal{F}_{t+k}^{\infty},\mathcal{F}_{-\infty}^{t}} |\Pr(A\cap B) - \Pr(A)\Pr(B)| \quad \text{for } k\in\mathbb{Z},$$

where $\mathcal{F}_a^b$ denotes the sigma-algebra generated by $\{a_t, x_t\}_{t=a}^b$. Our regularity conditions are listed below:

B1 *$X\times\Theta$ is a compact subset of $\mathbb{R}^L\times\mathbb{R}^J$ with $X^C = [\underline{x}, \overline{x}]$.*

B2 *The process $\{a_t, x_t\}_{t=1}^T$ is strictly stationary and strongly mixing, with a mixing coefficient $\alpha(k)$, such that for some $C \geq 0$ and some, possibly large $\chi > 0$, $\alpha(k) \leq Ck^{-\chi}$.*

B3 *The density of x_t is absolutely continuous $f_{X^C,X^D}(dx_t, x_t^d)$ for each $x_t^d \in X^D$. The joint density of (a_t, x_t) is bounded away from zero on X^C and is twice continuously differentiable over X^C for each $(x_t^d, a_t) \in X^D \times A$. The joint density of (x_{t+1}, x_t, a_t) is twice continuously differentiable over $X^C\times X^C$ for each $(x_{t+1}^d, x_t^d, a_t) \in X^D\times X^D\times A$.*

B4 *The mean of the per period payoff function $u_\theta(a_t, x_t)$ is twice continuously differentiable on $X^C\times\Theta$ for each $(x_t^d, a_t)\in X^D\times A$.*

B5 *The kernel function is a symmetric probability density function with bounded support such that for some constant C, $|K(u) - K(v)| \leq C|u - v|$. Define $\mu_j(K) = \int u^j K(u)\,du$ and $\kappa_j(K) = \int K^j(u)\,du$.*

B6 *The bandwidth sequence h_T satisfies $h_T = \gamma_0(T)T^{-1/5}$ and $\gamma_0(T)$ bounded away from zero and infinity.*

B7 *The triangular array of trimming factors $\{\mathfrak{c}_{t,T}\}$ is defined such that $\mathfrak{c}_{t,T} = \mathbf{1}[x_t^c \in X_T^C]$ where $X_T = [\underline{x} + c_T, \overline{x} - c_T]$ and $\{c_T\}$ is any positive sequence converging monotonically to zero such that $h_T < c_T$.*

B8 *The distribution of ε_t is known to be distributed as i.i.d. extreme value of type I across K alternatives, and is an independent of x_t and is i.i.d. across t.*

The compactness of the parameter space in B1 is standard. Compactness of the continuous component of the observable state space can be relaxed by using an increasing sequence of compact sets that cover the whole real line, see Linton and Mammen (2005) for the modeling in the tails of the distribution. The dimension of X^C is assumed to be 1 for expositional simplicity, discussion on this is follows the theorems below. On the other hand, it is a trivial matter to add arbitrary (finite) number of discrete components to X^D.

Condition B2 is quite weak despite the value of χ can be large.

The assumptions of B3, B4 and B5 are standard in the kernel smoothing literature using second order kernel.

Here in B6 we use the bandwidth with the optimal MSE rate for a regular 1-dimensional nonparametric estimates.

The trimming factor in B7 provides the necessary treatment of the boundary effects. This would ensure all the uniform convergence results on the expanding compact subset $\{X_T\}$ whose limit is X. In practice we will want to minimize the trimming out of the data, we can choose c_T close enough to h_T to do this.

Condition B8 is not necessary for consistency and asymptotic normality for any of the parameters below. The only requirement on the distribution of ε_t is that it allows us to employ Hotz and Miller's inversion theorem. For other distribution of ε_t, as discussed in Section 2, this will result in the use of a more complicated inversion map.

Let $\mu_2 = \int u^2K(u)\,du$ and $\kappa_2 = \int K^2(u)\,du$ denote the kernel constants. Next we provide the pointwise distribution theory for the nonparametric estimators obtained in the first stage.

Theorem 1. *Suppose* B1–B8 *hold. Then for each $\theta\in\Theta$, there exists deterministic functions $\eta_{m,\theta}$ and $\omega_{m,\theta}$ such that for each $x\in\operatorname{int}(X)$,*

$$\sqrt{Th_T}\left(\widehat{m}_\theta(x) - m_\theta(x) - \frac{1}{2}\mu_2 h_T^2\eta_{m,\theta}(x)\right) \Longrightarrow \mathcal{N}\left(0, \omega_{m,\theta}(x)\right),$$

where $\widehat{m}_\theta(x)$ is defined as in (18) and

$$\eta_{m,\theta}(x) = (I-\mathcal{L})^{-1}\left(\eta_{r,\theta} + \beta\eta_{\mathcal{L},\theta}\right)(x), \tag{25}$$

$$\omega_{m,\theta}(x) = \frac{\kappa_2}{f_X(x)}\left(\omega_{r,\theta}(x) + \beta^2\operatorname{var}(m_\theta(x_{t+1})|x_t = x)\right). \tag{26}$$

The explicit forms of $\eta_{r,\theta}$, $\eta_{\mathcal{L},\theta}$ and $\omega_{r,\theta}$ can be found at the beginning of Appendix A.2 in the Appendix A. The estimators $\widehat{m}_\theta(x)$ and $\widehat{m}_\theta(x')$ are also asymptotically independent for any $x\neq x'$. Furthermore,

$$\sup_{(x,\theta)\in X_T\times\Theta} |\widehat{m}_\theta(x) - m_\theta(x)| = o_p\left(T^{-1/4}\right).$$

We do not provide full expressions of $(\eta_{r,\theta}, \eta_{\mathcal{L},\theta}, \omega_{r,\theta})$ to save space. However, the determinants of the bias and variance terms are very intuitive; they are derived from the estimators of the intercept (r_θ) and operator $(\mathcal{L})$ of the integral equations respectively. In particular the pair $(\eta_{r,\theta}, \omega_{r,\theta})$ is precisely the (scaled) bias and variance terms for $\widehat{r}_\theta$, and $(\eta_{\mathcal{L},\theta}, \operatorname{var}(m_\theta(x_{t+1})|x_t = x))$ corresponds to the bias and variance of $(\widehat{\mathcal{L}} - \mathcal{L})m_\theta$. These terms have familiar expressions for our local constant estimators that are easy to estimate.

Theorem 2. *Suppose* B1–B8 *hold. Then for each $\theta\in\Theta$, $x\in\operatorname{int}(X)$ and $a\in A$,*

$$\sqrt{Th_T}\left(\widehat{g}_\theta(a,x)-g_\theta(a,x)-\frac{1}{2}\mu_2h_T^2\eta_{g,\theta}(a,x)\right)$$
$$\Longrightarrow \mathcal{N}\left(0,\omega_{g,\theta}(a,x)\right),$$

where $\widehat{g}_\theta(a,x)$ is defined as in (19) and

$$\eta_{g,\theta}(a,x)=\mathcal{H}\eta_{m,\theta}(a,x)+\eta_{\mathcal{H},\theta}(a,x),\tag{27}$$

$$\omega_{g,\theta}(a,x)=\frac{\kappa_2}{f_{X,A}(a,x)}\mathrm{var}\left(m_\theta(x_{t+1})\,|x_t=x,a_t=a\right).\tag{28}$$

The explicit forms of $\eta_{r,\theta}$, $\eta_{\mathcal{L},\theta}$ and $\eta_{\mathcal{H},\theta}$ can be found at the beginning of Appendix A.2 in the Appendix A. $\widehat{g}_\theta(a,x)$ and $\widehat{g}_\theta(a',x')$ are also asymptotically independent for any $x\neq x'$ and any a. Furthermore,

$$\sup_{(x,a,\theta)\in X_T\times A\times\Theta}\left|\widehat{g}_\theta(a,x)-g_\theta(a,x)\right|=o_p\left(T^{-1/4}\right).$$

Similar to Theorem 1, $\eta_{g,\theta}$ is a sum of the bias terms from $\widehat{m}_\theta$ and $(\widehat{\mathcal{H}}-\mathcal{H})m_\theta$. The asymptotic variance of $\widehat{g}_\theta$ is simply the asymptotic variance of $(\widehat{\mathcal{H}}-\mathcal{H})m_\theta$ since, unlike a solution to an integral equation (as seen in Theorem 1), the variance effect from estimating m_θ is of a smaller order.

We end with a brief discussion of the change in primitives required to accommodate the case when the dimension of X^C is higher than 1. Clearly, using the optimal (MSE) rates for h_T, $\dim(X^C)$ cannot exceed 3 with second order kernel if we were to have the uniform rate of convergence for our nonparametric estimates to be faster than $T^{-1/4}$ that is necessary for $\sqrt{T}$-consistency of the finite dimensional parameters. It is possible to overcome this by exploiting additional smoothness (if available) in the joint distribution of the random variables. This can be done by using higher order kernels to control the order of the bias, for details of their constructions and usages see Robinson (1988) and also Powell et al. (1989).

4.2. Finite dimensional parameters

In order to obtain consistency result and the parametric rate of convergence for $\widehat{\theta}$, we need to adjust some assumptions described in the previous subsection and add an identification assumption. Consider:

B6′ *The bandwidth sequence h_T satisfies $Th_T^4\to 0$ and $Th_T^2\to\infty$*

B9 *The value $\theta_0\in\mathrm{int}(\Theta)$ is defined by, for any $\varepsilon>0$*

$$\sup_{\|\theta-\theta_0\|\geq\varepsilon}Q(\theta_0)-Q(\theta)>0,$$

where $Q(\theta)$ denotes the limiting objective function of Q_T (defined in (23)), namely $Q(\theta)=\lim_{T\to\infty}EQ_T(\theta)$.

B10 *The matrix $E[\frac{\partial^2\log\ell(a_t,x_t;\theta,g_\theta)}{\partial\theta\partial\theta^\top}]$ is positive definite at θ_0.*

The rate of undersmoothing (relative to B6) in Condition B6′ ensures that the bias from the nonparametric estimation disappears sufficiently quickly to obtain parametric rate of convergence for $\widehat{\theta}$. To accommodate for higher dimension of X^C, we generally cannot just proceed by undersmoothing but combining this with the use higher order kernels, again, see Robinson (1988) and also Powell et al. (1989).

Condition B9 assumes the identification of the parametric part. This is a high level assumption that might not be easy to verify due to the complication with the value function. In practice we will have to check for local maxima for robustness. We note that this is the only assumption concerning the criterion function, for other type of objective functions, obvious analogous identification conditions will be required.

The properties of $\widehat{\theta}$ can be obtained by application of the asymptotic theory for semiparametric profile estimators. This requires the uniform expansion of $\widehat{g}_\theta$ (and hence $\widehat{m}_\theta$) and their derivatives with respect to θ. Let

$$\mathcal{I}=\lim_{T\to\infty}\mathrm{var}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T\frac{\partial\ell\left(a_t,x_t;\theta_0,g_{\theta_0}\right)}{\partial\theta}\right.$$
$$\left.+\frac{1}{\sqrt{T}}\sum_{t=1}^T\mathfrak{c}_{t,T}\left(\frac{\partial\ell\left(a_t,x_t;\theta_0,\widehat{g}_{\theta_0}\right)}{\partial\theta}-\frac{\partial\ell\left(a_t,x_t;\theta_0,g_{\theta_0}\right)}{\partial\theta}\right)\right),$$
$$\mathcal{J}=E\left[\frac{\partial^2\ell\left(a_t,x_t;\theta_0,g_{\theta_0}\right)}{\partial\theta\partial\theta^\top}\right].$$

Theorem 3. *Suppose* B1–B5, B6′ *and* B7–B10 *hold. Then*

$$\sqrt{T}\left(\widehat{\theta}-\theta_0\right)\Longrightarrow\mathcal{N}\left(0,\mathcal{J}^{-1}\mathcal{I}\mathcal{J}^{-1}\right).$$

Asymptotic normality follows, as suggested by $\mathcal{I}$, from applying central limit theorem to the sum of the sample scores and a corresponding term that takes into account of the nonparametric estimation in the first stage; averaging the latter improves on the nonparametric rate of convergence. Unlike $\widehat{m}_\theta$ and $\widehat{g}_\theta$, where inference can be performed based on obvious plug-in estimators, the asymptotic variance of $\widehat{\theta}$ is more complicated due to $\mathcal{I}$. This is not an uncommon problem for many semiparametric estimators. One popular approach for inference is through the bootstrap. Although there are some general theorems available on bootstrapping semiparametric estimators, e.g. see Chen et al. (2003) and Chen and Pouzo (2009), these results are derived under i.i.d. assumption and, to our knowledge, no analogous results are currently available for time series data. However, there are well-known positive results on the bootstrap of dependent processes with additional parametric and/or Markovian structures. Given that we know the primitives of the decision problem upto an estimation error, we suggest a bootstrap procedure that can be seen as a combination of (a semiparametric version of) Andrews' (2005) and Horowitz's (2003). In particular, for a given initial state x_1, a bootstrap sample can be obtained as follows: (Step 1) a vector ε_1^* is drawn from $q(\cdot|x_1)$; (Step 2) a_1^* is the maximizer of $\left\{\widehat{v}_{\widehat{\theta}}(a,x_1)+\varepsilon_{a,1}^*\right\}_{a\in A}$, i.e. a_1^* is the estimated policy $\widehat{\alpha}\left(x_1,\varepsilon_1^*\right)$ (see (3)); (Step 3) x_2^* is drawn from $f_{X'|X,A}\left(\cdot|x_1^*,a_1^*\right)$; (Step 4) repeat Steps 1–3 $(T-1)$-times to obtain $\left\{a_t^*,x_t^*\right\}_{t=1}^T$.[1] The formal proof of the semiparametric bootstrap is beyond the scope of this paper. We provide some Monte Carlo results below that show our suggested algorithm appears to work well even at small sample sizes.

Lastly, we present the results for the feasible estimators of m_{θ_0} and g_{θ_0}, which follow from $\sqrt{T}$-consistency of $\widehat{\theta}$.

Theorem 4. *Suppose* B1–B5, B6′ *and* B7–B10 *hold. Then for any arbitrary estimator $\widetilde{\theta}$ such that $\widetilde{\theta}-\theta_0=O_p\left(T^{-1/2}\right)$ and $x\in\mathrm{int}(X)$,*

$$\sqrt{Th_T}\left(\widehat{m}_{\widetilde{\theta}}(x)-m_{\theta_0}(x)\right)\Longrightarrow\mathcal{N}\left(0,\omega_{m,\theta_0}(x)\right),$$

where $\widehat{m}_\theta$ and $\omega_{m,\theta}$ are defined as those in Theorem 1 and, $\widehat{m}_{\widetilde{\theta}}(x)$ and $\widehat{m}_{\widetilde{\theta}}(x')$ are asymptotically independent for any $x\neq x'$.

Theorem 5. *Suppose* B1–B5, B6′ *and* B7–B10 *hold. Then for any arbitrary estimator $\widetilde{\theta}$ such that $\widetilde{\theta}-\theta_0=O_p\left(T^{-1/2}\right)$, $x\in\mathrm{int}(X)$ and $a\in A$,*

[1] In a closely related problem, Kasahara and Shimotsu (2008) assumes a parametric functional form for the transition density, $f_{X'|X,A}$, and propose a bootstrap algorithm based on Andrews' (2005) parametric bootstrap in a large cross-sectional framework.

Table 1

Summary statistics of various estimators for θ_{01}. $h_T = 1.06sT^{-1/5}$ is the bandwidth used in the nonparametric estimation, s denotes the sample standard deviation of $\{x_t\}_{t=1}^{T}$.

T	Bandwidth	Bias	mbias	std	(b-se)	iqr	mse
100	$h/2$	0.0278	0.0241	0.3951	(0.4172)	0.3941	0.1569
	h	0.0512	0.0262	0.3913	(0.4206)	0.3963	0.1557
	$2h$	0.0830	0.0935	0.3811	(0.4293)	0.3717	0.1521
	infML	0.0187	−0.0234	0.3923	–	0.3707	0.1543
	$D = 10$	−0.4427	−0.4208	0.8114	–	0.6373	0.8543
	$D = 25$	−1.9092	−1.9258	0.3069	–	0.2467	3.7393
	$D = 50$	−2.2054	−2.1755	0.2570	–	0.2422	4.9300
	$D = 100$	−2.4162	−2.3613	0.3114	–	0.2954	5.9350
500	$h/2$	0.0192	0.0110	0.1876	(0.1732)	0.1823	0.0344
	h	0.0216	0.0100	0.1822	(0.1728)	0.1749	0.0337
	$2h$	0.0474	0.0441	0.1832	(0.1755)	0.1840	0.0358
	infML	0.0030	−0.0067	0.1667	–	0.1560	0.0278
	$D = 10$	0.0292	0.0377	0.2000	–	0.2061	0.0409
	$D = 25$	−0.0408	−0.0041	0.3301	–	0.2121	0.1106
	$D = 50$	−0.8063	−0.7632	0.6979	–	1.0223	1.1371
	$D = 100$	−1.8170	−1.8185	0.1050	–	0.1039	3.3125
1000	$h/2$	0.0167	0.0166	0.1236	(0.1227)	0.1142	0.0155
	h	0.0174	0.0219	0.1222	(0.1210)	0.1210	0.0152
	$2h$	0.0381	0.0388	0.1225	(0.1233)	0.1250	0.0165
	infML	−0.0010	−0.0076	0.1161	–	0.1084	0.0135
	$D = 10$	0.0284	0.0251	0.1375	–	0.1278	0.0197
	$D = 25$	0.0102	0.0099	0.1366	–	0.1314	0.0188
	$D = 50$	−0.0314	−0.0113	0.2207	–	0.1354	0.0497
	$D = 100$	−0.9325	−0.9939	0.5878	–	0.8249	1.2151

$$\sqrt{Th_T}\left(\widehat{g_{\widetilde{\theta}}}(a,x) - g_{\theta_0}(a,x)\right) \Longrightarrow \mathcal{N}\left(0, \omega_{g,\theta_0}(a,x)\right),$$

where $\widehat{g_\theta}$ *and* $\omega_{g,\theta}$ *are defined as those in Theorem 2 and,* $\widehat{g_{\widetilde{\theta}}}(a,x)$ *and* $\widehat{g_{\widetilde{\theta}}}(a',x')$ *are asymptotically independent for any* $x \neq x'$ *and any* a.

5. Numerical illustration

In this section we illustrate some finite sample properties of our proposed estimator in a small scale Monte Carlo experiment.

Design and Implementation. We consider the decision process of an agent (say, a mobile vender) who, in each period t, has a choice to operate in either location A or B. The decision variable a_t takes value 1 if location A is chosen, and 0 otherwise. The immediate payoff from the decision is

$$u(a_t, x_t, \varepsilon_t) = \pi_{\theta_0}(a_t, x_t) + a_t\varepsilon_{1,t} + (1 - a_t)\varepsilon_{0,t},$$

where $\pi_\theta(a_t, x_t) = \theta_1 a_t x_t + \theta_2(1 - a_t)(1 - x_t)$. Here x_t denotes a publicly observed measure of the demand determinant that has been normalized to lie between [0, 1]. The vector $(\varepsilon_{1,t}, \varepsilon_{0,t})$ represents some non-persistent idiosyncratic private costs associated with each choice, which are distributed as i.i.d. extreme value of type 1, that are not observed by the econometricians. To capture the most general aspect of the decision processes discussed in the paper, the future value of x_{t+1} evolves stochastically and its conditional distribution is affected by the observables from the previous period (a_t, x_t). We suppose the transition density has the following form

$$f_{X'|X,A}(x'|x,a) = \begin{cases} \delta_{11}(x)x' + \delta_{12}(x) & \text{when } a = 1 \\ \delta_{21}(x)x' + \delta_{22}(x) & \text{when } a = 0. \end{cases}$$

We design our model to be consistent with a plausible scenario that future demand builds on existing demand, which is driven by whether or not the vender was present at a particular location. In particular, if the demand at location A is high and the vender is not present, the demand at location A is more likely to be significantly lower for the next period (and vice versa). We use the following simple specific forms for $\{\delta_{ij}\}$ that display such behavior: $\delta_{11}(x) = 2(2x - 1)$, $\delta_{12}(x) = 2(1 - x)$, $\delta_{21}(x) = 2(1 - 2x)$, and $\delta_{22}(x) = 2x$. To introduce some asymmetry, we impose that the agent has underlying preference toward location A, which is captured by the condition $\theta_{01} > \theta_{02} > 0$.

We set $(\beta, \theta_{01}, \theta_{02})$ to be (0.9, 1, 0.5) and use the fixed point method described in Rust (1996) to generate the controlled Markovian process under the proposed primitives. The initial state is taken as $x_1 = 1/2$, we begin sampling each decision process after 1000 periods and consider $T = 100, 500, 1000$. We conduct 500 replications of each time length. For each T, we obtain our estimators and the corresponding bootstrap standard errors by following the procedures described in Section 3. To approximate the integral equation we partition [0, 1] by using 1000 equally-spaced grid points. Since the support of the observable state is compact, we need to trim of values near the boundary. As an alternative, we employ a simple boundary corrected kernel, see Wand and Jones (1994), based on a Gaussian kernel, namely

$$K_h^b(x_t - x) = \begin{cases} \frac{1}{h}K\left(\frac{x_t - x}{h}\right) \Big/ \int_{-x/h}^{\infty} K(v)\,dv & \text{if } x \in [0, h) \\ \frac{1}{h}K\left(\frac{x_t - x}{h}\right) & \text{if } x \in (h, 1 - h) \\ \frac{1}{h}K\left(\frac{x_t - x}{h}\right) \int_{-\infty}^{(1-x)/h} K(v)\,dv & \text{if } x \in (1 - h, 1], \end{cases}$$

where K is the pdf of a standard normal. We consider three choices of bandwidth $(h_T/2, h_T, 2h_T)$, where $h_T = 1.06sT^{-1/5}$ and s denotes the standard deviation of the observed $\{x_t\}$. For a comparison we also estimate the infeasible parametric maximum likelihood (ML) estimator as well as a series of manually discretized estimators. For the discretization, we partition [0, 1] into D-equally spaced intervals and the support of x_t is reduced to D-points taking the mid-point value of each interval for $D = 10, 25, 50, 100$. There is an empty cell problem with the frequency estimator associated with our discretization scheme. We use the smooth kernel estimator for categorical data analyzed in Ouyang et al. (2006) with the smoothing parameter taking value $1/T$, for a review of smooth nonparametric estimation with discrete data see Li and Racine (2006).

Results. We report the summary statistics for the θ_{01} and θ_{02} in Tables 1 and 2 respectively. The bias (mean and median) and the standard deviation generally improve as the sample size increases for all estimators. We note that the semiparametric

Table 2
Summary statistics of various estimators for θ_{02}. $h_T = 1.06sT^{-1/5}$ is the bandwidth used in the nonparametric estimation, s denotes the sample standard deviation of $\{x_t\}_{t=1}^T$.

T	Bandwidth	Bias	mbias	std	(b-se)	iqr	mse
100	$h/2$	0.0410	0.0416	0.4961	(0.4979)	0.4780	0.2478
	h	0.0590	0.0273	0.4693	(0.4954)	0.4470	0.2237
	$2h$	0.0897	0.0636	0.4525	(0.4834)	0.4471	0.2128
	infML	0.0070	−0.0110	0.4784	–	0.4755	0.2289
	$D = 10$	−0.8634	−0.8237	0.7330	–	0.6011	1.2826
	$D = 25$	−1.8833	−1.9168	0.3148	–	0.2659	3.6459
	$D = 50$	−1.4172	−1.4276	0.2971	–	0.2575	2.0968
	$D = 100$	−1.9162	−1.8613	0.3114	–	0.2954	3.7688
500	$h/2$	0.0364	0.0196	0.2188	(0.2095)	0.2041	0.0492
	h	0.0413	0.0297	0.2106	(0.2099)	0.1956	0.0461
	$2h$	0.0579	0.0525	0.2129	(0.2075)	0.1934	0.0487
	infML	0.0112	0.0054	0.2166	–	0.2174	0.0470
	$D = 10$	−0.0478	−0.0617	0.2252	–	0.2364	0.0530
	$D = 25$	−0.1547	−0.0939	0.3978	–	0.2480	0.1822
	$D = 50$	−1.0590	−1.0052	0.8908	–	1.3362	1.9152
	$D = 100$	−1.3170	−1.3185	0.1050	–	0.1039	1.7455
1000	$h/2$	0.0084	0.0053	0.1461	(0.1479)	0.1344	0.0214
	h	0.0154	0.0248	0.1428	(0.1466)	0.1426	0.0206
	$2h$	0.0303	0.0265	0.1484	(0.1462)	0.1499	0.0229
	infML	0.0100	0.0053	0.1449	–	0.1430	0.0211
	$D = 10$	−0.0248	−0.0239	0.1614	–	0.1649	0.0267
	$D = 25$	−0.0454	−0.0454	0.1616	–	0.1672	0.0282
	$D = 50$	−0.0875	−0.0603	0.2730	–	0.1625	0.0822
	$D = 100$	−0.4325	−0.4939	0.5878	–	0.8249	0.5326

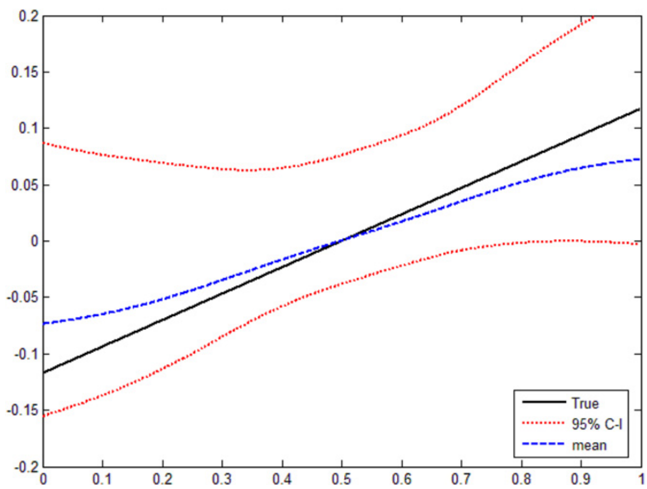

Fig. 1. The difference in the estimated mean estimator of choice specific expected continuation values, $E\left[V_{\theta_0}(s_{t+1})\,|x_t, a_t = 1\right] - E\left[V_{\theta_0}(s_{t+1})\,|x_t, a_t = 0\right]$, with 95% confidence interval for $T = 100$ for $h_T = 1.06sT^{-1/5}$.

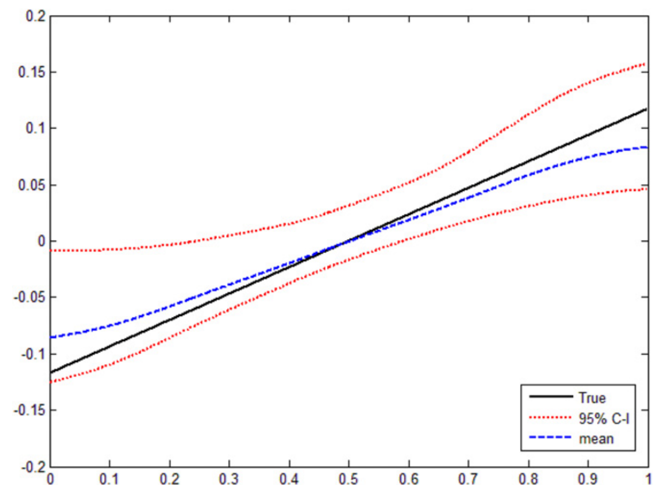

Fig. 2. The difference in the estimated mean estimator of choice specific expected continuation values, $E\left[V_{\theta_0}(s_{t+1})\,|x_t, a_t = 1\right] - E\left[V_{\theta_0}(s_{t+1})\,|x_t, a_t = 0\right]$, with 95% confidence interval for $T = 500$ for $h_T = 1.06sT^{-1/5}$.

bootstrap standard errors provide a very reasonable estimate for our estimators in this particular example. The ML estimators dominate the semiparametric estimators as expected. For these estimators, we find that across all bandwidths: (1) The ratios of the bias and standard deviation remain small at larger sample sizes providing support for consistency; (2) The ratios of the scaled interquartile range (by a factor 1.349) and the standard error are also close to 1, which is a trait of a normal random variable. However, the performance of the discretized estimators is mixed and they appear to generally underperform, based on the MSE criterion, relative to the semiparametric estimators in this study. In particular, the biases of the discretized estimators are very large when sample size is relatively small for a given support size of the discretized states.

We plot the estimated differences between the expected continuation values, $E\left[V_{\theta_0}(s_{t+1})\,|x_t, a_t = 1\right] - E[V_{\theta_0}(s_{t+1})|x_t, a_t = 0]$, along with their 95% confidence bands (see Figs. 1–3). These functional estimates are computed using the estimated $(\widehat{\theta}_1, \widehat{\theta}_2)$. We only provide the plots for a single bandwidth (h_T), other bandwidths lead to qualitatively similar results. The plots clearly show that the bias and the confidence sets reduce as sample size increases. Although our estimates exhibit larger bias near the boundaries, the true function is also covered by the confidence sets almost uniformly over the support across all sample sizes.

6. Conclusion

In this paper we propose a method to estimate a class of Markov discrete decision processes that allows for continuous observable state space based on kernel smoothing. The primitive conditions are provided for the inference of the finite and infinite dimensional parameters in the model. Our estimation technique relies on the convenient well-posed linear inverse problem presented by the policy value equation. It inherits the computational simplicity of Pesendorfer and Schmidt-Dengler (2008) and can be used to estimate the same class of Markovian games that allow for a continuous state space. We conduct a Monte Carlo experiment

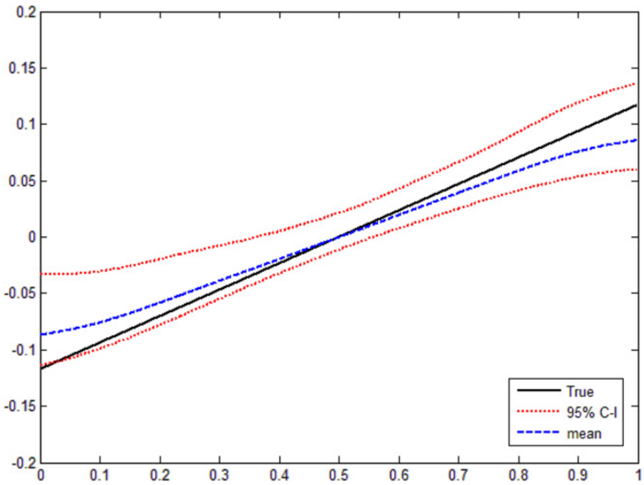

Fig. 3. The difference in the estimated mean estimator of choice specific expected continuation values, $E\left[V_{\theta_0}(s_{t+1})\,|x_t, a_t = 1\right] - E\left[V_{\theta_0}(s_{t+1})\,|x_t, a_t = 0\right]$, with 95% confidence interval for $T = 1000$ for $h_T = 1.06sT^{-1/5}$.

and compare the performance of our estimator to the infeasible maximum likelihood estimator and those obtained from manually discretizing the state space. The results show that our estimator appears to work well generally, and relative to the discretized estimators, in finite sample in this particular exercise.

There are some practical aspects of our estimators worth exploring. Firstly, is the role of numerical error brought upon by approximating the integral in the case that we have large sample size compared to the purely nonparametric approximation. Second, although we provide some Monte Carlo results, is to see how our estimator performs in practice relative to discretization methods. Thirdly, some efficiency bounds should be obtainable in the special case of the conditional logit assumption.

Appendix A

We first provide a set of high level assumptions (A1–A6) and their consequences (C1–C4) of the nonparametric estimators described in Section 3. We outline the stochastic expansions required to obtain the asymptotic properties of $\widehat{m}_\theta$ and $\widehat{g}_\theta$. The high level assumptions and the main theorems are then proved under the primitives of M1–M3 and B1–B10. Consequences are simple and their proofs are omitted. In what follows, we refer frequently to Bosq (1998), Linton and Mammen (2005), Masry (1996) and Robinson (1983), so for brevity, we denote their references by [B], [LM], [M] and [R] respectively.

A.1. Outline of the asymptotic approach

For notational simplicity we work on a Banach space, $(C(\mathcal{X}), \|\cdot\|)$, where $\mathcal{X} = \mathcal{X}^C \times X^D$, the continuous part of $\mathcal{X}$ is a compact set $\left[\underline{x}+\varepsilon, \overline{x}-\varepsilon\right]$ for some arbitrarily small $\varepsilon > 0$. We let $B1'$ be an analogous condition to B1 when we replace X by $\mathcal{X}$. The approach taken here is similar to [LM], who worked on the L^2 Hilbert Space. The main difference between our problem and theirs is, after getting consistent estimates of (10), we require another level of smoothing, see (11), before plugging it into the criterion function. The first part here follows [LM].

Assumption A1. Suppose that for some sequence $\delta_T = o(1)$:

$$\sup_{x\in\mathcal{X}}\left|\left(\widehat{\mathcal{L}}-\mathcal{L}\right)m(x)\right| = o_p(\delta_T),$$

i.e., $\|(\widehat{\mathcal{L}}-\mathcal{L})m\| = o_p(\delta_T)$ for any $m \in C(\mathcal{X})$.

Consequence C1. Under A1:

$$\left\|\left(\left(I-\widehat{\mathcal{L}}\right)^{-1}-(I-\mathcal{L})^{-1}\right)m\right\| = o_p(\delta_T).$$

The rate of uniform approximation of the linear operator gets transferred to the inverse of $(I-\mathcal{L})$. This is summarized by C1 and is proven in [LM].

We suppose that $\widehat{r}_\theta(x) - r_\theta(x)$ can be decomposed into the following terms that satisfy some properties.

Assumption A2. For each $x \in \mathcal{X}$:

$$\widehat{r}_\theta(x) - r_\theta(x) = \widehat{r}_\theta^B(x) + \widehat{r}_\theta^C(x) + \widehat{r}_\theta^D(x), \quad (29)$$

where $\widehat{r}_\theta^C, \widehat{r}_\theta^D$ and $\widehat{r}_\theta^E$ satisfy:

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\widehat{r}_\theta^B(x)\right| = O_p\left(T^{-2/5}\right) \quad \text{with } \widehat{r}_\theta^B \text{ deterministic}, \quad (30)$$

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\widehat{r}_\theta^C(x)\right| = o_p\left(T^{-2/5+\xi}\right) \quad \text{for any } \xi > 0, \quad (31)$$

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\mathcal{L}(I-\mathcal{L})^{-1}\widehat{r}_\theta^C(x)\right| = o_p\left(T^{-2/5}\right), \quad (32)$$

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\widehat{r}_\theta^D(x)\right| = o_p\left(T^{-2/5}\right). \quad (33)$$

This is the standard bias+variance+remainder of local constant kernel estimates of the regression function under some smoothness assumptions. The intuition behind (32), as provided in [LM], is that the operator applies averaging to a local smoother and transforms it into a global average thereby reducing its variance. These terms are used to obtain the components of $\widehat{m}_\theta(x)$, for $j = B, C, D$, where each $\widehat{m}_\theta^j(x)$ satisfies the following integral equation

$$\widehat{m}_\theta^j = \widehat{r}_\theta^j + \widehat{\mathcal{L}}\widehat{m}_\theta^j, \quad (34)$$

and $\widehat{m}_\theta^A$, from writing the solution $m_\theta + \widehat{m}_\theta^A$ to the integral equation

$$\left(m_\theta + \widehat{m}_\theta^A\right) = r_\theta + \widehat{\mathcal{L}}\left(m_\theta + \widehat{m}_\theta^A\right). \quad (35)$$

The existence and uniqueness of the solutions to (34) and (35) are assured, at least w.p.a. 1, under the contraction property of the integral operator, so it follows from the linearity of $(I-\widehat{\mathcal{L}})^{-1}$ that

$$\widehat{m}_\theta = m_\theta + \widehat{m}_\theta^A + \widehat{m}_\theta^B + \widehat{m}_\theta^C + \widehat{m}_\theta^D.$$

These components can be approximated by simpler terms. Define also m_θ^B, as the solution to

$$m_\theta^B = \widehat{r}_\theta^B + \mathcal{L}m_\theta^B. \quad (36)$$

Consequence C2. Under A1–A2:

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\widehat{m}_\theta^B(x) - m_\theta^B(x)\right| = o_p\left(T^{-2/5}\right), \quad (37)$$

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\widehat{m}_\theta^C(x) - r_\theta^C(x)\right| = o_p\left(T^{-2/5}\right), \quad (38)$$

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\widehat{m}_\theta^D(x)\right| = o_p\left(T^{-2/5}\right). \quad (39)$$

(37) and (39) follow immediately from (30), (33) and C1. (38) follows from (32), A1 and C1, as one can easily show that

$$\left\|\widehat{\mathcal{L}}\left(I-\widehat{\mathcal{L}}\right)^{-1} - \mathcal{L}(I-\mathcal{L})^{-1}\right\| = o_p(\delta_T).$$

We next approximate $\widehat{m}_\theta^A$ by simpler terms. Subtracting (10) from (35) yields

$$\widehat{m}_\theta^A = \left(\widehat{\mathcal{L}}-\mathcal{L}\right)m_\theta + \widehat{\mathcal{L}}\widehat{m}_\theta^A. \quad (40)$$

Assumption A3. For any $x \in \mathcal{X}$:

$$\left(\widehat{\mathcal{L}} - \mathcal{L}\right) m_\theta (x) = \widehat{r}_\theta^E (x) + \widehat{r}_\theta^F (x) + \widehat{r}_\theta^G (x),$$

where $\widehat{r}_\theta^E, \widehat{r}_\theta^F$ and $\widehat{r}_\theta^G$ satisfy:

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta} \left|\widehat{r}_\theta^E (x)\right| = O_p\left(T^{-2/5}\right) \quad \text{with } \widehat{r}_\theta^E \quad \text{deterministic},$$

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta} \left|\widehat{r}_\theta^F (x)\right| = o_p\left(T^{-2/5+\xi}\right) \quad \text{for any } \xi > 0,$$

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta} \left|\mathcal{L}(I-\mathcal{L})^{-1}\widehat{r}_\theta^F (x)\right| = o_p\left(T^{-2/5}\right),$$

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta} \left|\widehat{r}_\theta^G (x)\right| = o_p\left(T^{-2/5}\right).$$

These terms are obtained by decomposing the conditional density estimates (cf. A2). For $j = E, F, G$, let the components $\widehat{m}_\theta^j (x)$ solve (40) so that

$$\widehat{m}_\theta^A = \left(I - \widehat{\mathcal{L}}\right)^{-1}\left(\widehat{\mathcal{L}} - \mathcal{L}\right) m_\theta = \widehat{m}_\theta^E + \widehat{m}_\theta^F + \widehat{m}_\theta^G.$$

Define also m_θ^E as the solution to the analogous integral equation of (36), then we have a similar result to C2.

Consequence C3. Under A1–A3:

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta} \left|\widehat{m}_\theta^E (x) - m_\theta^E (x)\right| = o_p\left(T^{-2/5}\right),$$

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta} \left|\widehat{m}_\theta^F (x) - \widehat{r}_\theta^F (x)\right| = o_p\left(T^{-2/5}\right),$$

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta} \left|\widehat{m}_\theta^G (x)\right| = o_p\left(T^{-2/5}\right).$$

C3 can be shown using the same reasoning used to obtain C2.

Combining these assumptions leads to the following Proposition 1.

Proposition 1. *Suppose that [A1–A3] holds for some estimators $\widehat{r}_\theta$ and $\widehat{\mathcal{L}}$. Define $\widehat{m}_\theta$ as any solution of $\widehat{m}_\theta = \widehat{r}_\theta + \widehat{\mathcal{L}}\widehat{m}_\theta$. Then the following expansion holds for $\widehat{m}_\theta$*

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta} \left|\widehat{m}_\theta (x) - m_\theta (x) - m_\theta^B (x) - m_\theta^E (x) - \widehat{r}_\theta^C (x) - \widehat{r}_\theta^F (x)\right|$$
$$= o_p\left(T^{-2/5}\right),$$

where all of the terms above have been defined previously.

The uniform expansion for the nonparametric estimators discussed in [LM] ends here. However, to obtain the uniform expansion of $\widehat{g}_\theta$ defined in (19), we need another level of smoothing. Note that the integral operator $\mathcal{H}$ has a different range $\mathcal{H} : C(\mathcal{X}) \to C(A \times X)$, where $C(A \times X)$ denotes a space of functions, say $g(a, x)$, which are continuous on X for each $a \in A$. So the relevant Banach Space is equipped with the sup-norm over $A \times X$, which we also denote by $\|\cdot\|$ though this should not lead to any confusion. For notational simplicity, we first define: $\overline{m}_\theta^B (x) = m_\theta^B (x) + m_\theta^E (x)$, $\overline{m}_\theta^C (x) = \widehat{r}_\theta^C (x) + \widehat{r}_\theta^F (x)$, $\overline{m}_\theta^D (x) = \widehat{m}_\theta (x) - m_\theta (x) - \overline{m}_\theta^B (x) - \overline{m}_\theta^C (x)$. We next define various components of $\widehat{g}_\theta$ and g_θ, analogously to (34) and (36), for $j = B, C, D$ let $\widehat{g}_\theta^j = \widehat{\mathcal{H}}\overline{m}_\theta^j$ and $g_\theta^j = \mathcal{H}\overline{m}_\theta^j$, and define $\widehat{g}_\theta^A$ to satisfy $\widehat{\mathcal{H}}m_\theta = g_\theta + \widehat{g}_\theta^A$. It follows from the linearity of $\widehat{\mathcal{H}}$ that $\widehat{g}_\theta = g_\theta + \widehat{g}_\theta^A + \widehat{g}_\theta^B + \widehat{g}_\theta^C + \widehat{g}_\theta^D$.

Assumption A4. Suppose that for some sequence δ_T as in A1:

$$\sup_{(a,x,\theta)\in A\times X\times\Theta} \left|\left(\widehat{\mathcal{H}} - \mathcal{H}\right) m_\theta (x, a)\right| = o_p(\delta_T),$$

i.e., $\|(\widehat{\mathcal{H}} - \mathcal{H})m\| = o_p(\delta_T)$ for any $m \in C(\mathcal{X})$.

A4 assumes the desirable properties of the conditional density estimators (cf. A1 and A3).

Consequence C4. Under A1–A4:

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta} \left|\widehat{g}_\theta^B (a, x) - g_\theta^B (a, x)\right| = o_p\left(T^{-2/5}\right),$$

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta} \left|\widehat{g}_\theta^C (a, x) - g_\theta^C (a, x)\right| = o_p\left(T^{-2/5}\right),$$

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta} \left|\widehat{g}_\theta^D (a, x)\right| = o_p\left(T^{-2/5}\right).$$

This follows immediately from A5 and the properties of the elements that define $\overline{m}_\theta^B$.

Assumption A5. Suppose that

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta} \left|g_\theta^C (a, x)\right| = o_p\left(T^{-2/5}\right).$$

A5 follows since the operator $\mathcal{H}$ is a global smooth, hence it reduces the variance of $\overline{m}_\theta^C$.

As with $\widehat{m}_\theta^A$, we can approximate $\widehat{g}_\theta^A$ by simpler terms.

Assumption A6. For any $m \in C(\mathcal{X})$ and for each $(a, x) \in A \times \mathcal{X}$:

$$\widehat{g}_\theta^A (a, x) = \left(\widehat{\mathcal{H}} - \mathcal{H}\right) m_\theta (x, a) = \widehat{g}_\theta^E (a, x) + \widehat{g}_\theta^F (a, x) + \widehat{g}_\theta^G (a, x),$$

where $\widehat{g}_\theta^E, \widehat{g}_\theta^F$ and $\widehat{g}_\theta^G$ satisfy:

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta} \left|\widehat{g}_\theta^E (a, x)\right| = O_p\left(T^{-2/5}\right) \quad \text{with } \widehat{g}_\theta^E \quad \text{deterministic},$$

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta} \left|\widehat{g}_\theta^F (a, x)\right| = o_p\left(T^{-2/5+\xi}\right) \quad \text{for any } \xi > 0,$$

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta} \left|\widehat{g}_\theta^G (a, x)\right| = o_p\left(T^{-2/5}\right).$$

A6 follows from a standard decomposition of the kernel conditional density estimator (cf. A3).

Proposition 2. *Suppose that A1–A6 holds for some estimators $\widehat{r}_\theta$, $\widehat{\mathcal{L}}$ and $\widehat{\mathcal{H}}$. Define $\widehat{m}_\theta$ as any solution of $\widehat{m}_\theta = \widehat{r}_\theta + \widehat{\mathcal{L}}\widehat{m}_\theta$ and $\widehat{g}_\theta = \widehat{\mathcal{H}}\widehat{m}_\theta$. Then the following expansion holds for $\widehat{g}_\theta$*

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta} \left|\widehat{g}_\theta (a, x) - g_\theta (a, x)\right.$$
$$\left. - g_\theta^B (a, x) - \widehat{g}_\theta^E (a, x) - \widehat{g}_\theta^F (a, x)\right| = o_p\left(T^{-2/5}\right),$$

where all of the terms above have been defined previously, in particular g_θ^B and $\widehat{g}_\theta^E$ are non-stochastic and the leading variance terms is $\widehat{g}_\theta^F$. Similarly to, $\overline{m}_\theta^j (x)$ for $j = B, C, D$, we define: $\overline{g}_\theta^B (a, x) = g_\theta^B (a, x) + \widehat{g}_\theta^E (a, x)$, $\overline{g}_\theta^C (a, x) = \widehat{g}_\theta^F (a, x)$, and $\overline{g}_\theta^D (a, x) = \widehat{g}_\theta (a, x) - g_\theta (a, x) - \overline{g}_\theta^B (a, x) - \overline{g}_\theta^C (a, x)$.

A.2. Proofs of Theorems 1 and 2 and high level conditions A1–A6

We assume $B1'$, B2–B6 and B8 throughout this subsection. Set $\delta_T = T^{\xi-3/10}$, this rate is arbitrarily close to the rate of convergence of 1-dimensional nonparametric density estimates when h_T decays at the rate specified by B6. For notational simplicity we assume that X^D is empty. The presence of discrete states do not affect any of the results below, we can simply replace any formula involving the density (and analogously for the conditional density) $f_X(dx_t)$ by $f_X\left(dx_t, x_t^d\right)$. We shall denote generic finite and positive constants by C_0 that may take different values in different places. The uniform rate of convergence proof of various components utilize some exponential inequalities found in [B], as done in [LM], the details are deferred to Appendix B.

It is useful to begin by defining various components that make up the bias and variance terms in Theorems 1 and 2. Define $e_{a,t} = \mathbf{1}[a_t = a] - P(a|x_t)$, and let

$$\eta_{P_a}(x) = 2\frac{\frac{\partial P(a|x)}{\partial x}\frac{\partial f_X(x)}{\partial x}}{f_X(x)} + \frac{\partial P^2(a|x)}{\partial x^2} \text{and } \omega_{P_a}(x) = \kappa_2 \frac{\sigma_a^2(x)}{f_X(x)} \tag{41}$$

respectively denote the bias and variance contributions from the Nadaraya–Watson estimator $\widehat{P}(a|x)$, where $\sigma_a^2(x) = E[e_{a,t}^2 | x_t = x]$. Define a real value function $\zeta_{a,x,\theta}$, indexed by (a, x, θ), such that

$\zeta_{a,x,\theta}(t) = t(\pi_\theta(a, x) + \log t) + \gamma, \quad$ for $t > 0$,

where γ is the Euler constant. Then, from (14) and (15), we can write $r_\theta(x) = \sum_{a \in A} \zeta_{a,x,\theta}(P(a|x))$. Let $\zeta'_{a,x,\theta}$ denotes the derivative of $\zeta_{a,x,\theta}$, then:

$$\eta_{r,\theta}(x) = \sum_{a \in A} \eta_{P_a}(x)\, \zeta'_{a,x,\theta}(P(a|x)) \tag{42}$$

$$\omega_{r,\theta}(x) = \frac{\kappa_2}{f_X(x)} \times \begin{pmatrix} \sum_{a \in A} \left(\zeta'_{a,x,\theta}(P(a|x))\right)^2 P(a|x)(1 - P(a|x)) \\ -2\sum_{a \neq \widetilde{a}} \zeta'_{a,x,\theta}(P(a|x))\, \zeta'_{\widetilde{a},x,\theta}(P(\widetilde{a}|x))\, P(a|x)\, P(\widetilde{a}|x) \end{pmatrix}, \tag{43}$$

respectively denote the bias and variance contributions for the intercept in the integral equation (r_θ). Also, let

$$\eta_{\mathcal{L},\theta}(x) = \beta \begin{pmatrix} \frac{1}{f_X(x)} \int m_\theta(x') \left(\frac{\partial^2 f_{X',X}(x', x)}{\partial x'^2} + \frac{\partial^2 f_{X',X}(x', x)}{\partial x^2} \right) dx' \\ -\frac{\frac{\partial^2 f_X(x)}{\partial x^2}}{f_X(x)} \int m_\theta(x')\, f_{X'|X}(dx'|x) \end{pmatrix}$$

$$\eta_{\mathcal{H},\theta}(a, x) = \frac{\int m_\theta(x') \left(\frac{\partial^2 f_{X',X,A}(x',x,a)}{\partial^2 x'} + \frac{\partial^2 f_{X',X,A}(x',x,a)}{\partial^2 x} \right) dx'}{f_{X,A}(x, a)} - \frac{\frac{\partial^2 f_{X,A}(x,a)}{\partial^2 x}}{f_{X,A}(x, a)} \int m_\theta(x')\, f_{X'|X,A}(dx'|x, a)$$

denote the bias contributions that arise from estimating the integral operators $\mathcal{L}$ and $\mathcal{H}$ (operating on m_θ) respectively.

Proof of Theorem 1. We proceed by providing the pointwise distribution theory for $\widehat{P}(a|x)$, for any $(a, x) \in A \times \mathcal{X}$, and the functionals thereof. By the definition of $\widehat{P}(a|x)$:

$$\widehat{P}(a|x) - P(a|x) = \frac{1}{T}\sum_{t=1}^{T} (\mathbf{1}[a_t = a] - P(a|x)) \times K_h(x_t - x) / \widehat{f}_X(x),$$

focusing on the numerator

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} (\mathbf{1}[a_t = a] - P(a|x)) K_h(x_t - x) &= \frac{1}{T}\sum_{t=1}^{T} (P(a|x_t) - P(a|x)) K_h(x_t - x) \\ &\quad + \frac{1}{T}\sum_{t=1}^{T} e_{a,t} K_h(x_t - x) \\ &= A_{1,T}(a, x) + A_{2,T}(a, x), \end{aligned}$$

where $e_{a,t} = \mathbf{1}[a_t = a] - P(a|x_t)$. The term $A_{1,T}(a, x)$ is dominated by the bias, by the usual change of variables and Taylor's expansion, more specifically

$$\begin{aligned} E\left[A_{1,T}(a, x)\right] &= E\left[(P(a|x_t) - P(a|x)) K_h(x_t - x)\right] \\ &= \frac{1}{2}\mu_2 h_T^2 \left(2\frac{\partial P(a|x)}{\partial x}\frac{\partial f_X(x)}{\partial x} + \frac{\partial^2 P(a|x)}{\partial x^2} f_X(x) \right) + o\left(h_T^2\right). \end{aligned}$$

By construction $E\left[e_{a,t}|x_t\right] = 0$ for all a and t. The variance of the summands $A_{2,T}(a, x)$ is dominated by the variances as covariance terms are of smaller order, e.g. see [M],

$$\begin{aligned} \mathrm{var}\left(A_{2,T}(a, x)\right) &= \mathrm{var}\left(\frac{1}{T}\sum_{t=1}^{T} e_{a,t} K_h(x_t - x) \right) \\ &= \frac{1}{T}\mathrm{var}\left(e_{a,t} K_h(x_t - x)\right) + o\left(\frac{1}{Th_T}\right) \\ &= \frac{1}{T} E\left[\sigma_a^2(x_t) K_h(x_t - x)\right] + o\left(\frac{1}{Th_T}\right) \\ &= \frac{\kappa_2}{Th_T}\sigma_a^2(x) f_X(x) + o\left(\frac{1}{Th_T}\right), \end{aligned}$$

where it is easy to see that $\sigma_a^2(x) = \mathrm{var}(\mathbf{1}[a_t = a] | x_t = x) = P(a|x)(1 - P(a|x))$. For the CLT, Lemma 7.1 of [R] can be used repeated throughout this section. To obtain the asymptotic distribution for $\widehat{r}_\theta(x)$, we next provide the joint distribution of $\{\widehat{P}(a|x)\}$. It follows from [R] that for any $a \in A$

$$\sqrt{Th_T}\left(\widehat{P}(a|x) - P(a|x) - \frac{1}{2}\mu_2 h_T^2 \eta_{P_a}(x) \right) \Longrightarrow \mathcal{N}\left(0, \omega_{P_a}(x)\right),$$

and, from the Cramér–Wold device,

$$\sqrt{Th_T} \begin{pmatrix} \widehat{P}(1|x) - P(1|x) - \frac{1}{2}\mu_2 h_T^2 \eta_{P_1}(x) \\ \vdots \\ \widehat{P}(K|x) - P(K|x) - \frac{1}{2}\mu_2 h_T^2 \eta_{P_K}(x) \end{pmatrix} \Longrightarrow \mathcal{N}\left(0, \begin{pmatrix} \omega_{P_1}(x) & \omega_{P_{2,1}}(x) & \cdots & \omega_{P_{K,1}}(x) \\ \omega_{P_{1,2}}(x) & \ddots & & \vdots \\ \vdots & & \ddots & \omega_{P_{K,K-1}}(x) \\ \omega_{P_{1,K}}(x) & \cdots & \omega_{P_{K-1,K}}(x) & \omega_{P_K}(x) \end{pmatrix}\right)$$

where η_{P_a} and ω_{P_1} are both defined in (41), and $\omega_{P_{a,\widetilde{a}}}(x) = \kappa_2 \frac{-P(a|x)P(\widetilde{a}|x)}{f_X(x)}$ for $a, \widetilde{a} \in A$. Note that the covariance matrix in the above display is rank deficient due to the constraint that $\sum_{a \in A} \widehat{P}(a|x) = 1$ for all $x \in \mathcal{X}$. Recall that $\widehat{r}_\theta(x) = \sum_{a \in A} \zeta_{a,x,\theta}\left(\widehat{P}(a|x)\right)$, and by the mean value theorem (MVT)

$$\begin{aligned} &\zeta_{a,x,\theta}\left(\widehat{P}(a|x)\right) - \zeta_{x,a,\theta}\left(P(a|x) + \frac{1}{2}\mu_2 h_T^2 \eta_{P_a}(x)\right) \\ &\quad = \zeta'_{a,x,\theta}(P(a|x)) \left(\widehat{P}(a|x) - P(a|x) - \frac{1}{2}\mu_2 h_T^2 \eta_{P_a}(x) \right) + o_p(1), \end{aligned}$$

where $\zeta'_{a,x,\theta}(t) = \pi_\theta(a, x) + \log t + 1$ for $t > 0$. By using MVT again,

$$\zeta_{a,x,\theta}\left(P(a|x)+\frac{1}{2}\mu_2 h_T^2\eta_{P_a}(x)\right)$$
$$=\zeta_{a,x,\theta}(P(a|x))+\frac{1}{2}\mu_2 h_T^2\eta_{P_a}(x)\,\zeta'_{a,x,\theta}(P(a|x))+o_p\left(h_T^2\right),$$

and by the continuous mapping theorem

$$\sqrt{Th_T}\left(\widehat{r}_\theta(x)-r_\theta(x)-\frac{1}{2}\mu_2 h_T^2\eta_{r,\theta}\right)\Longrightarrow\mathcal{N}\left(0,\omega_{r,\theta}(x)\right),$$

where $\eta_{r,\theta}$ and $\omega_{r,\theta}$ are defined in (42) and (43) respectively. Note we can relate the components of the expansion of $\widehat{r}_\theta(x)$, in (29), to the terms above as follows

$$\widehat{r}_\theta^B(x)=\frac{1}{2}\mu_2 h_T^2\eta_{r,\theta}(x),\tag{44}$$

$$\widehat{r}_\theta^C(x)=\sum_{a\in A}\frac{\zeta'_{a,x,\theta}(P(a|x))}{f_X(x)}\times\left(\frac{1}{T}\sum_{t=1}^{T}e_{a,t}K_h(x_t-x)\right).\tag{45}$$

We next provide the statistical properties for $\widehat{m}_\theta^A(x)$. First consider $\left(\widehat{\mathcal{L}}-\mathcal{L}\right)m_\theta(x)$:

$$\left(\widehat{\mathcal{L}}-\mathcal{L}\right)m_\theta(x)=\beta\int m_\theta\left(x'\right)\left(\widehat{f}_{X'|X}\left(dx'|x\right)-f_{X'|X}\left(dx'|x\right)\right)$$
$$=\frac{\beta}{f_X(x)}\int m_\theta\left(x'\right)\left(\widehat{f}_{X,X}\left(dx',x\right)-f_{X,X}\left(dx',x\right)\right)$$
$$-\frac{\beta}{f_X(x)}\left(\widehat{f}_X(x)-f_X(x)\right)$$
$$\times\int m_\theta\left(x'\right)f_{X'|X}\left(dx'|x\right)+o_p\left(T^{-2/5}\right)$$
$$=B_{1,\theta,T}(x)+B_{2,\theta,T}(x)+o_p\left(T^{-2/5}\right).$$

To analyze $B_{1,\theta,T}(x)$, proceed with the usual decomposition of $\widehat{f}_{X',X}\left(x',x\right)-f_{X',X}\left(x',x\right)$ then integrating out x'. We have:

$$B_{1,\theta,T}(x)=B_{1,\theta,T}^B(x)+B_{1,\theta,T}^C(x)+o_p\left(T^{-2/5}\right),$$

$$B_{1,\theta,T}^B(x)=\frac{1}{2}\mu_2 h_T^2\beta\int\left(\frac{m_\theta\left(x'\right)}{f_X(x)}\left(\frac{\partial^2 f_{X',X}\left(x',x\right)}{\partial x'^2}+\frac{\partial^2 f_{X',X}\left(x',x\right)}{\partial x^2}\right)\right)dx',\tag{46}$$

$$B_{1,\theta,T}^C(x)=\frac{\beta}{f_X(x)}\left(\frac{1}{T-1}\sum_{t=1}^{T-1}\int m_\theta\left(x'\right)\times\left(\begin{array}{c}K_h\left(x_{t+1}-x'\right)K_h(x_t-x)\\-E\left[K_h\left(x_{t+1}-x'\right)K_h(x_t-x)\right]\end{array}\right)\right)dx',\tag{47}$$

and it is easy to show that

$$\sqrt{Th_T}B_{1,\theta,T}^C(x)\Longrightarrow\mathcal{N}\left(0,\frac{\beta^2}{f_X(x)}\kappa_2\int\left(m_\theta\left(x'\right)\right)^2 f_{X'|X}\left(dx'|x\right)\right).$$

For $B_{2,\theta,T}(x)$, this is just the kernel density estimator of $f_X(x)$ multiplied by a non-stochastic term,

$$B_{2,\theta,T}(x)=B_{2,\theta,T}^B(x)+B_{2,\theta,T}^C(x)+o_p\left(T^{-2/5}\right),$$

$$B_{2,\theta,T}^B(x)=-\frac{1}{2}\mu_2 h_T^2\frac{\partial^2 f_X(x)}{\partial x^2}\times\left(\frac{\beta}{f_X(x)}\int m_\theta\left(x'\right)f_{X'|X}\left(dx'|x\right)\right),\tag{48}$$

$$B_{2,\theta,T}^C(x)=-\left(\frac{\beta}{f_X(x)}\int m_\theta\left(x'\right)f_{X'|X}\left(dx'|x\right)\right)\frac{1}{T}\times\sum_{t=1}^{T}\left(K_h(x_t-x)-E\left[K_h(x_t-x)\right]\right),\tag{49}$$

and it follows that

$$\sqrt{Th_T}B_{2,\theta,T}^C(x)\Longrightarrow\mathcal{N}\left(0,\kappa_2 f_X(x)\left(\frac{\beta}{f_X(x)}\int m_\theta\left(x'\right)f_{X'|X}\left(dx'|x\right)\right)^2\right).$$

Combining these we have

$$\widehat{m}_\theta(x)=m_\theta(x)+\overline{m}_\theta^B(x)+\overline{m}_\theta^C(x)+o_p\left(T^{-2/5}\right),$$

where $\overline{m}_\theta^B(x)=(I-\mathcal{L})^{-1}(B_{1,\theta,T}^B+B_{2,\theta,T}^B+\widehat{r}_\theta^B)(x)$ and $\overline{m}_\theta^C(x)=B_{1,\theta,T}^C(x)+B_{2,\theta,T}^C(x)+\widehat{r}_\theta^C(x)$. Note also that as $T\to\infty$:

$$\sqrt{Th_T}\left(B_{1,\theta,T}^C(x)+B_{2,\theta,T}^C(x)\right)\Longrightarrow\mathcal{N}\left(0,\frac{\kappa_2\beta^2}{f_X(x)}\mathrm{var}\left(m_\theta(x_{t+1})|x_t=x\right)\right),$$

$$\mathrm{Cov}\left(\sqrt{Th_T}\left(B_{1,\theta,T}^C(x)+B_{2,\theta,T}^C(x)\right),\sqrt{Th_T}\widehat{r}_\theta^C(x)\right)\to 0.$$

This provides us with the pointwise theory for $\widehat{m}_\theta$ for any $x\in\mathcal{X}$ and $\theta\in\Theta$.

$$\sqrt{Th_T}\left(\widehat{m}_\theta(x)-m_\theta(x)-\frac{1}{2}\mu_2 h_T^2\eta_{m,\theta}(x)\right)\Longrightarrow\mathcal{N}\left(0,\omega_{m,\theta}(x)\right),$$

where $\eta_{m,\theta}$ and $\omega_{m,\theta}$ are defined in (25) and (26) respectively. The proof of pairwise asymptotic independence across distinct x is obvious. □

Proof of Theorem 2. Similarly to the decomposition of $(\widehat{\mathcal{L}}-\mathcal{L})m_\theta(x)$, we have

$$\left(\widehat{\mathcal{H}}-\mathcal{H}\right)m_\theta(x,a)$$
$$=\int m_\theta\left(x'\right)\left(\widehat{f}_{X'|X,A}\left(dx'|x,a\right)-f_{X'|X,A}\left(dx'|x,a\right)\right)$$
$$=C_{1,\theta,T}(a,x)+C_{2,\theta,T}(a,x)+o_p\left(T^{-2/5}\right).$$

The properties for $C_{1,\theta,T}$ and $C_{2,\theta,T}$ are closely related to that of $B_{1,\theta,T}$ and $B_{2,\theta,T}$, specifically:

$$C_{1,\theta,T}(a,x)=C_{1,\theta,T}^B(a,x)+C_{1,\theta,T}^C(a,x)+o_p\left(T^{-2/5}\right),$$

$$C_{1,\theta,T}^B(a,x)=\frac{1}{2}\mu_2 h_T^2\int\frac{m_\theta\left(x'\right)}{f_{X,A}(x,a)}\times\left(\frac{\partial^2 f_{X',X,A}\left(x',x,a\right)}{\partial x'^2}+\frac{\partial^2 f_{X',X,A}\left(x',x,a\right)}{\partial x^2}\right)dx',$$

$$C_{1,\theta,T}^C(a,x)=\frac{1}{f_{X,A}(x,a)}\int\left(\frac{1}{T-1}\sum_{t=1}^{T-1}m_\theta\left(x'\right)\times\left(\begin{array}{c}K_h\left(x_{t+1}-x'\right)K_h(x_t-x)\mathbf{1}[a_t=a]\\-E\left[K_h\left(x_{t+1}-x'\right)K_h(x_t-x)\mathbf{1}[a_t=a]\right]\end{array}\right)\right)dx',$$

and, as in the case of $B_{1,\theta,T}^C$,

$$\sqrt{Th_T}C_{1,\theta,T}^C(a,x)\Longrightarrow\mathcal{N}\left(0,\frac{\kappa_2}{f_{X,A}(x,a)}\times\int\left(m_\theta\left(x'\right)\right)^2 f_{X'|X,A}\left(dx'|x,a\right)\right).$$

Similarly for $C_{2,\theta,T}$:

$$C^B_{2,\theta,T}(a,x) = -\frac{1}{2}\mu_2 h_T^2 \frac{\partial^2 f_{X,A}(x,a)}{\partial x^2} \times \left(\frac{1}{f_{X,A}(x,a)}\int m_\theta(x') f_{X'|X,A}(dx'|x,a)\right),$$

$$C^C_{2,\theta,T}(a,x) = -\left(\frac{1}{f_{X,A}(x,a)}\int m_\theta(x') f_{X'|X,A}(dx'|x,a)\right) \times \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} K_h(x_t-x)\mathbf{1}[a_t=a] \\ -E[K_h(x_t-x)\mathbf{1}[a_t=a]]\end{pmatrix},$$

$$\sqrt{Th_T}C^C_{2,\theta,T}(a,x) \Longrightarrow \mathcal{N}\left(0, \frac{\kappa_2}{f_{X,A}(x,a)} \times \int (m_\theta(x'))^2 f_{X'|X,A}(dx'|x,a)\right).$$

Combining these we have

$$\widehat{g}_\theta(a,x) = g_\theta(a,x) + \overline{g}^B_\theta(a,x) + \overline{g}^C_\theta(a,x) + o_p\left(T^{-2/5}\right),$$

where $\overline{g}^B_\theta(a,x) = C^B_{1,\theta,T}(a,x) + C^B_{2,\theta,T}(a,x) + \mathcal{H}\widehat{r}^B_\theta(a,x)$ and $\overline{g}^C_\theta(a,x) = C^C_{1,\theta,T}(a,x) + C^C_{2,\theta,T}(a,x)$. This provides us with the pointwise distribution theory for $\widehat{g}$, for any $x \in \mathcal{X}$, $a \in A$ and $\theta \in \Theta$

$$\sqrt{Th_T}\left(\widehat{g}(a,x) - g(a,x) - \frac{1}{2}\mu_2 h_T^2 \eta_{g,\theta}(a,x)\right) \Longrightarrow \mathcal{N}\left(0, \omega_{g,\theta}(a,x)\right),$$

where $\eta_{g,\theta}$ and $\omega_{g,\theta}$ as defined in (27) and (28) respectively. The pairwise asymptotic independence across distinct x completes the proof. □

Proof of A1. It suffices to show that

$$\sup_{(x',x)\in\mathcal{X}\times\mathcal{X}}\left|\widehat{f}_{X',X}(x',x) - f_{X',X}(x',x)\right| = o_p(\delta_T),$$

$$\sup_{x\in\mathcal{X}}\left|\widehat{f}_X(x) - f_X(x)\right| = o_p(\delta_T).$$

These uniform rates are bounded by the rates for the bias squared and the rates of the centered process. The former is standard, and holds uniformly over $\mathcal{X}\times\mathcal{X}$ (and $\mathcal{X}$). See Appendix B, where proof of A1 falls under Case 1. □

Proof of A2. The components for the decomposition have been provided by (44)–(45). By uniform boundedness of η_{P_a} and $\zeta_{a,x,\theta}$ over $A\times\mathcal{X}\times\Theta$ and the triangle inequality, the order of the leading bias and remainder terms are as stated in (30) and (33) respectively. For the stochastic term, we can utilize the exponential inequality, see Case 2 of Appendix B. We next check (32). [LM] make use of an eigen-expansion to construct the kernel of the new integral operator and showed that it had nice properties in their problem. In contrast, we use the Neumann's series to construct the kernel of the transform $\mathcal{L}(I-\mathcal{L})^{-1}$ directly. For any $\phi \in C(\mathcal{X})$

$$\mathcal{L}(I-\mathcal{L})^{-1}\phi = \sum_{\tau=1}^{\infty}\mathcal{L}^\tau\phi,$$

where $\mathcal{L}^\tau$ represents the linear operator of a τ-step ahead predictor with discounting, this follows from Chapman–Kolmogorov equation for homogeneous Markov chains

$$\mathcal{L}^\tau\phi(x) = \beta^\tau\int\phi(x')f_{(\tau)}(dx'|x),$$

$$f_{(\tau)}(x_{t+\tau}|x_t) = \int f_{X'|X}(x_{t+\tau}|x_{t+\tau-1})\prod_{k=1}^{\tau-1}f_{X'|X}(dx_{t+\tau-k}|x_{t+\tau-k-1}),$$

where $f_{(\tau)}(dx_{t+\tau}|x_t)$ denotes the conditional density of τ-steps ahead. First note that $\mathcal{L}(I-\mathcal{L})^{-1}\phi \in C(\mathcal{X})$ since for any $\phi \in C(\mathcal{X})$ and $x\in\mathcal{X}$

$$\begin{aligned}\left|\mathcal{L}(I-\mathcal{L})^{-1}\phi(x)\right| &= \left|\sum_{\tau=1}^{\infty}\beta^\tau\int\phi(x')f_{(\tau)}(dx'|x)\right| \\ &\le \sum_{\tau=1}^{\infty}\beta^\tau\int f_{(\tau)}(dx'|x)\|\phi\| \\ &\le \frac{\beta}{1-\beta}\|\phi\| \\ &< \infty.\end{aligned}$$

We denote the kernel of the integral transform $\mathcal{L}(I-\mathcal{L})^{-1}$ by the limit of the partial sum $\lambda_{\mathcal{T}}$, λ, where

$$\lambda_{\mathcal{T}}(x',x) = \sum_{\tau=1}^{\mathcal{T}}\beta^\tau f_{(\tau)}(x'|x). \tag{50}$$

Note that λ is continuous on $\mathcal{X}\times\mathcal{X}$ since $f_{(\tau)}$ is continuous and is uniformly bounded for all τ by $\sup_{(x',x)\in\mathcal{X}\times\mathcal{X}}\left|f_{X'|X}(x'|x)\right|$; by completeness $\lambda_{\mathcal{T}}$ converges to a continuous function (with Lipschitz constant no larger than $\frac{\beta}{1-\beta}\sup_{(x',x)\in\mathcal{X}\times\mathcal{X}}\left|f_{X'|X}(x'|x)\right|$). Then for the proof of (32), from (45), it is sufficient to show

$$\sup_{(a,x)\in A\times\mathcal{X}}\left|\frac{1}{T}\sum_{t=1}^{T}e_{a,t}\nu_{a,\theta}(x_t,x)\right| = o_p\left(T^{-2/5}\right),$$

where $\nu_{a,\theta}$ is defined as

$$\nu_{a,\theta}(x_t,x) = \int\frac{\zeta'_{a,x',\theta}(P(a|x'))}{f_X(x')}K_h(x_t-x')\lambda(dx',x).$$

The uniform bound can be obtained by applying exponential inequality, see Case 3 of Appendix B for details. □

Proof of A3. Following the decomposition of $\widehat{f}_{X'|X}$, the leading bias term is the sum of (46) and (48), and the variance term is the sums of (47) and (49). The results regarding the rates of convergence follow similarly to the proof of A2. □

Proof of A4. This is essentially the same as the proof of A1. □

Proof of A5. Since $\overline{m}^C_\theta$ consists of $\widehat{r}^C_\theta$ and $\widehat{r}^F_\theta$. We need to show,

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta}\left|\mathcal{H}\widehat{r}^C_\theta(a,x)\right| = o_p\left(T^{-2/5}\right)$$

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta}\left|\mathcal{H}\widehat{r}^F_\theta(a,x)\right| = o_p\left(T^{-2/5}\right).$$

The proof follows from exponential inequalities, see Appendix B. □

Proof of A6. This is essentially the same as proof of A3. □

A.3. Proofs of Theorems 3–5

We begin with two lemmas for the uniform expansion of the partial derivatives of $\widehat{m}_\theta$ and $\widehat{g}_\theta$ w.r.t. θ.

Lemma 1. *Under conditions* B1′, B2–B6 *and* B8 *hold. Then the following expansion holds for $k = 0, 1, 2$ and $j = 1, \ldots, J$*

$$\max_{1\le j\le J}\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\frac{\partial^k\widehat{m}_\theta(x)}{\partial\theta_j^k} - \frac{\partial^k m_\theta(x)}{\partial\theta_j^k} - \frac{\partial^k\overline{m}^B_\theta(x)}{\partial\theta_j^k} - \frac{\partial^k\overline{m}^C_\theta(x)}{\partial\theta_j^k}\right| = o_p\left(T^{-2/5}\right),$$

where $\frac{\partial^k m_\theta}{\partial \theta^k}$ *is defined as the solution to*

$$\frac{\partial^k m_\theta}{\partial \theta_j^k} = \frac{\partial^k r_\theta}{\partial \theta_j^k} + \mathcal{L}\frac{\partial^k m_\theta}{\partial \theta_j^k}, \tag{51}$$

and $\frac{\partial^k \widehat{m}_\theta}{\partial \theta_j^k}$ *defined as the solution to the analogous empirical integral equation. Standard definition for partial derivative applies for* $\frac{\partial^k \overline{m}_\theta^b}{\partial \theta_j^k}$ *with* $b = B, C$. *When* $k = 0$ *the expansion above coincides with the terms previously defined in Proposition 1. Further,*

$$\max_{1\leq j\leq J}\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\frac{\partial^k \overline{m}_\theta^B(x)}{\partial \theta_j^k}\right| = O_p\left(T^{-2/5}\right) \quad \text{with} \frac{\partial^k \overline{m}_\theta^B(x)}{\partial \theta^k} \quad \text{deterministic},$$

$$\max_{1\leq j\leq J}\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\frac{\partial^k \overline{m}_\theta^C(x)}{\partial \theta_j^k}\right| = o_p\left(T^{\xi-2/5}\right) \quad \text{for any } \xi > 0.$$

Proof of Lemma 1. Comparing the integral equations in (10) and (51), these involve the same integral operator but different intercepts. Since $\zeta_{a,x,\theta}$ and m_θ are twice continuously differentiable in θ, on Θ over $A \times X$, Dominated Convergence Theorem (DCT) can be utilized throughout. Hence all the arguments used to verify the definition of $\frac{\partial^k \overline{m}_\theta}{\partial \theta_j^k}$ and their uniformity results, analogous to A2–A3, follow immediately. □

Lemma 2. *Under conditions* B1′, B2–B6 *hold. Then the following expansion holds for* $k = 0, 1, 2$ *and* $j = 1, \ldots, J$

$$\max_{1\leq j\leq J}\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta}\left|\frac{\partial^k \widehat{g}_\theta(a,x)}{\partial \theta_j^k} - \frac{\partial^k g_\theta(a,x)}{\partial \theta_j^k} - \frac{\partial^k \overline{g}_\theta^B(a,x)}{\partial \theta_j^k} - \frac{\partial^k \overline{g}_\theta^C(a,x)}{\partial \theta_j^k}\right| = o_p\left(T^{-2/5}\right),$$

where all of the terms above are defined analogously to those found in Lemma 1, and for $k = 1, 2$

$$\max_{1\leq j\leq J}\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta}\left|\frac{\partial^k \overline{g}_\theta^B(a,x)}{\partial \theta_j^k}\right| = O_p\left(T^{-2/5}\right) \quad \text{with} \frac{\partial^k \overline{g}_\theta^B(a,x)}{\partial \theta_j^k} \quad \text{deterministic},$$

$$\max_{1\leq j\leq J}\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta}\left|\frac{\partial^k \overline{g}_\theta^C(a,x)}{\partial \theta_j^k}\right| = o_p\left(T^{\xi-2/5}\right) \quad \text{for any } \xi > 0.$$

Proof of Lemma 2. Same as the proof of Lemma 1. □

Proof of Theorem 3. We first proceed to show the consistency result of the estimator.

Consistency. Consider any estimator θ_T of θ_0 that asymptotically maximizes $\widehat{Q}_T(\theta)$:

$$Q_T(\theta_T) \geq \sup_{\theta\in\Theta} Q_T(\theta) - o_p(1).$$

Under B1 and B9, by a standard argument, for example see Newey and McFadden (1994), consistency of such extremum estimator will follow if we can show

$$\sup_{\theta\in\Theta}\left|\widehat{Q}_T(\theta) - Q(\theta)\right| = o_p(1).$$

By the triangle inequality, this is implied by

$$\sup_{\theta\in\Theta}|Q_T(\theta) - Q(\theta)| = o_p(1) \tag{52}$$

$$\sup_{\theta\in\Theta}\left|\widehat{Q}_T(\theta) - Q_T(\theta)\right| = o_p(1). \tag{53}$$

For (52), since $\ell : A \times X \times \Theta \to \mathbb{R}$ is continuous on the compact set $X \times \Theta$, for any $a \in A$, hence by Weierstrass Theorem

$$\sup_{(a,x,\theta)\in A\times X\times\Theta}|\ell(a,x;\theta,g_\theta)| < \infty.$$

This ensures that $E|\ell(a_t,x_t;\theta,v_\theta)| < \infty$, and by the LLN for ergodic and stationary processes we have

$$Q_T(\theta) \xrightarrow{p} Q(\theta) \quad \text{for each } \theta\in\Theta.$$

The convergence above can be made uniform since Q_T is stochastic equicontinuous and Q is uniformly continuous by DCT. To proof (53) we partition $\widehat{Q}_T(\theta) - Q_T(\theta)$ into two components

$$\widehat{Q}_T(\theta) - Q_T(\theta) = \frac{1}{T}\sum_{t=1}^T \mathfrak{c}_{t,T}\left(\ell(a_t,x_t;\theta,\widehat{g}_\theta) - \ell(a_t,x_t;\theta,g_\theta)\right) + \frac{1}{T}\sum_{t=1}^T\left(1-\mathfrak{c}_{t,T}\right)\ell(a_t,x_t;\theta,\widehat{g}_\theta).$$

The second term is $o_p(1)$. To see this, denote $1-\mathfrak{c}_{t,T}$ by $\mathfrak{d}_{t,T}$, then

$$\left|\frac{1}{T}\sum_{t=1}^T \mathfrak{d}_{t,T}\ell(a_t,x_t;\theta,\widehat{g}_\theta)\right| \leq \sup_{(a,x,\theta)\in A\times X\times\Theta}|\ell(a,x;\theta,g_\theta)|\frac{1}{T}\sum_{t=1}^T \mathfrak{d}_{t,T} = o_p(1).$$

The inequality holds w.p.a. 1 and the equality is the result of $\mathfrak{d}_{t,T} = o_p(\vartheta_T)$ for any sequence $\vartheta_T = o(1)$. To proof (53), it suffices to show

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta}|\ell(a,x;\theta,\widehat{g}_\theta) - \ell(a,x;\theta,g_\theta)| = o_p(1).$$

Recall that

$$\ell(a,x;\theta,\widehat{g}_\theta) - \ell(a,x;\theta,g_\theta) = \widehat{v}_\theta(a,x) - v_\theta(a,x) + \log\left(\frac{\sum_{\widetilde{a}\in A}\exp(v_\theta(\widetilde{a},x))}{\sum_{\widetilde{a}\in A}\exp(\widehat{v}_\theta(\widetilde{a},x))}\right),$$

where $v_\theta = \pi_\theta + g_\theta$ and $\widehat{v}_\theta$ differs from v_θ by replacing g_θ with $\widehat{g}_\theta$. We have shown earlier that, for some $\delta_T = o(1)$,

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta}|\widehat{g}_\theta(a,x) - g_\theta(a,x)| = o_p(\delta_T).$$

So we have uniform convergence of $\widehat{v}$ to v at the same rate. We know that, for any continuously differentiable function ϕ (in this case, $\exp(\cdot)$ and $\log(\cdot)$), MVT implies

$$\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta}|\phi(\widehat{v}_\theta(a,x)) - \phi(v_\theta(a,x))| = o_p(\delta_T).$$

Therefore

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\sum_{a\in A}\exp(\widehat{v}_\theta(a,x)) - \sum_{a\in A}\exp(v_\theta(a,x))\right| = o_p(1).$$

Since we have, at least w.p.a. 1, $\exp(\widehat{v}_\theta(a,x))$ and $\exp(v_\theta(a,x))$ are positive a.s.

$$\left|\frac{\sum_{a\in A}\exp\left(v_\theta\left(a,x\right)\right)}{\sum_{a\in A}\exp\left(\widehat{v}_\theta\left(a,x\right)\right)}-1\right|=\left|\frac{1}{\sum_{a\in A}\exp\left(\widehat{v}_\theta\left(a,x\right)\right)}\right|$$
$$\times\left|\sum_{a\in A}\exp\left(v_\theta\left(a,x\right)\right)-\sum_{a\in A}\exp\left(\widehat{v}_\theta\left(a,x\right)\right)\right|,$$

and by the Weierstrass Theorem $\inf_{(a,x,\theta)\in A\times X\times\Theta}\exp\left(\widehat{v}_\theta\left(a,x\right)\right)>0$, hence we have

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\frac{\sum_{a\in A}\exp\left(v_\theta\left(a,x\right)\right)}{\sum_{a\in A}\exp\left(\widehat{v}_\theta\left(a,x\right)\right)}-1\right|=o_p\left(1\right).$$

The proof of (53) is completed once we apply another mean value expansion to obtain

$$\sup_{(x,\theta)\in\mathcal{X}\times\Theta}\left|\log\left(\frac{\sum_{a\in A}\exp\left(v_\theta\left(a,x\right)\right)}{\sum_{a\in A}\exp\left(\widehat{v}_\theta\left(a,x\right)\right)}\right)\right|=o_p\left(1\right).\quad\square$$

Asymptotic normality. Our estimator asymptotically satisfies the first order condition

$$\frac{\partial\widehat{Q}_T\left(\widehat{\theta}\right)}{\partial\theta}=o_p\left(1/\sqrt{T}\right).$$

Taking the mean value expansion around θ_0,

$$\sqrt{T}\left(\widehat{\theta}-\theta_0\right)=\left(-\frac{\partial^2\widehat{Q}_T\left(\overline{\theta}\right)}{\partial\theta^2}\right)^{-1}\sqrt{T}\frac{\partial\widehat{Q}_T\left(\theta_0\right)}{\partial\theta}+o_p\left(1\right),$$

for some intermediate value $\overline{\theta}$ in the $\|\widehat{\theta}-\theta_0\|$ neighborhood of θ_0. To complete the proof of Theorem 3, we will show:

AN1 Asymptotic positive definiteness of $\frac{\partial^2\widehat{Q}_T(\theta)}{\partial\theta^2}$ near θ_0,

$$\inf_{\|\theta-\theta_0\|<\epsilon_T}\lambda_{\min}\left(-\frac{\partial^2\widehat{Q}_T\left(\theta\right)}{\partial\theta\partial\theta^\top}\right)>C_0+o_p\left(1\right)$$

for any $\epsilon_T=o\left(1\right)$ and some $C_0>0$,

$$\frac{\partial^2\widehat{Q}_T\left(\overline{\theta}\right)}{\partial\theta\partial\theta^\top}\xrightarrow{p}\mathcal{J}=E\left[\frac{\partial^2\ell\left(a_t,x_t;\theta_0,g_{\theta_0}\right)}{\partial\theta\partial\theta^\top}\right].$$

AN2 Asymptotic normality of $\sqrt{T}\frac{\partial\widehat{Q}_T(\theta_0)}{\partial\theta}=D_{1,T}+D_{2,T}+o_p\left(1\right)\Longrightarrow\mathcal{N}\left(0,\mathcal{I}\right)$, where

$$D_{1,T}=\frac{1}{\sqrt{T}}\sum_{t=1}^T\frac{\partial\ell\left(a_t,x_t;\theta_0,g_{\theta_0}\right)}{\partial\theta},$$
$$D_{2,T}=\frac{1}{\sqrt{T}}\sum_{t=1}^T\mathfrak{c}_{t,T}\times\left(\frac{\partial\ell\left(a_t,x_t;\theta_0,\widehat{g}_{\theta_0}\right)}{\partial\theta}-\frac{\partial\ell\left(a_t,x_t;\theta_0,g_{\theta_0}\right)}{\partial\theta}\right).$$

Proof of AN1: By B10, it is sufficient to show

$$\sup_{\|\theta-\theta_0\|<\epsilon_T}\left\|\frac{\partial^2\widehat{Q}_T\left(\theta\right)}{\partial\theta\partial\theta^\top}-E\left[\frac{\partial^2\ell\left(a_t,x_t;\theta,g_\theta\right)}{\partial\theta\partial\theta^\top}\right]\right\|=o_p\left(1\right).\tag{54}$$

Since the second derivative of ℓ is continuous on the compact set $A\times X\times\Theta$, standard arguments for uniform convergence implies that

$$\sup_{\|\theta-\theta_0\|<\epsilon_T}\left\|\frac{\partial^2Q_T\left(\theta\right)}{\partial\theta\partial\theta^\top}-E\left[\frac{\partial^2\ell\left(a_t,x_t;\theta,g_\theta\right)}{\partial\theta\partial\theta^\top}\right]\right\|=o_p\left(1\right).$$

By the triangle inequality, (54) will hold if we can show

$$\sup_{\|\theta-\theta_0\|<\epsilon_T}\left\|\frac{\partial^2\widehat{Q}_T\left(\theta\right)}{\partial\theta\partial\theta^\top}-\frac{\partial^2Q_T\left(\theta\right)}{\partial\theta\partial\theta^\top}\right\|=o_p\left(1\right).$$

This condition is implied by

$$\sup_{(a,x)\in A\times\mathcal{X},\|\theta-\theta_0\|<\epsilon_T}\left\|\frac{\partial^2\ell\left(a,x;\theta,\widehat{g}_\theta\right)}{\partial\theta\partial\theta^\top}-\frac{\partial^2\ell\left(a,x;\theta,g_\theta\right)}{\partial\theta\partial\theta^\top}\right\|=o_p\left(1\right).$$

We begin by writing the score in terms of v_θ,

$$\frac{\partial\ell\left(a_t,x_t;\theta,g_\theta\right)}{\partial\theta}=\frac{\partial v_\theta\left(a_t,x_t\right)}{\partial\theta}-\frac{\sum_{a\in A}\left(\frac{\partial v_\theta(a,x_t)}{\partial\theta}\right)\exp\left(v_\theta\left(a,x_t\right)\right)}{\sum_{a\in A}\exp\left(v_\theta\left(a,x_t\right)\right)},$$

and similarly for the Hessian

$$\frac{\partial^2\ell\left(a_t,x_t;\theta,g_\theta\right)}{\partial\theta\partial\theta^\top}=\frac{\partial^2v_\theta\left(a_t,x_t\right)}{\partial\theta\partial\theta^\top}-\frac{\sum_{a\in A}\left(\frac{\partial^2v_\theta(a,x_t)}{\partial\theta\partial\theta^\top}+\frac{\partial v_\theta(a,x_t)}{\partial\theta}\frac{\partial v_\theta(a,x_t)}{\partial\theta^\top}\right)\exp\left(v_\theta\left(a,x_t\right)\right)}{\sum_{a\in A}\exp\left(v_\theta\left(a,x_t\right)\right)}$$
$$+\frac{\sum_{a\in A}\sum_{\widetilde{a}\in A}\frac{\partial v_\theta(a,x_t)}{\partial\theta}\frac{\partial v_\theta(\widetilde{a},x_t)}{\partial\theta^\top}\exp\left(v_\theta\left(a,x_t\right)+v_\theta\left(\widetilde{a},x_t\right)\right)}{\left(\sum_{a\in A}\exp\left(v_\theta\left(a,x_t\right)\right)\right)^2}.$$

We show (54) holds by a tedious but straightforward calculation, using similar arguments for proving (53); repeatedly making use of MVT and uniform convergence of the following partial derivatives

$$\max_{1\leq j\leq J}\sup_{(a,x,\theta)\in A\times\mathcal{X}\times\Theta}\left|\frac{\partial^k\widehat{v}_\theta\left(a,x\right)}{\partial\theta_j^k}-\frac{\partial^kv_\theta\left(a,x\right)}{\partial\theta_j^k}\right|=o_p\left(1\right)\quad\text{for }k=0,1,2,$$

which follows from Lemmas 1 and 2. $\square$

Proof of AN2: From the definition of $\widehat{Q}_T$, we write

$$\sqrt{T}\frac{\partial\widehat{Q}_T\left(\theta_0\right)}{\partial\theta}=\frac{1}{\sqrt{T}}\sum_{t=1}^T\mathfrak{c}_{t,T}\frac{\partial\ell\left(a_t,x_t;\theta_0,\widehat{g}_{\theta_0}\right)}{\partial\theta}$$
$$=\frac{1}{\sqrt{T}}\sum_{t=1}^T\frac{\partial\ell\left(a_t,x_t;\theta_0,g_{\theta_0}\right)}{\partial\theta}+\frac{1}{\sqrt{T}}\sum_{t=1}^T\mathfrak{c}_{t,T}\left(\frac{\partial\ell\left(a_t,x_t;\theta_0,\widehat{g}_{\theta_0}\right)}{\partial\theta}-\frac{\partial\ell\left(a_t,x_t;\theta_0,g_{\theta_0}\right)}{\partial\theta}\right)$$
$$+\frac{1}{\sqrt{T}}\sum_{t=1}^T\mathfrak{d}_{t,T}\frac{\partial\ell\left(a_t,x_t;\theta_0,\widehat{g}_{\theta_0}\right)}{\partial\theta}$$
$$=D_{1,T}+D_{2,T}+D_{3,T}.$$

Note that $D_{1,T}$ is asymptotically normal with mean zero and variance Λ_1, of the infeasible MLE, where

$$\Lambda_1 = E\left[\frac{\partial \ell\left(a_t, x_t; \theta_0, g_{\theta_0}\right)}{\partial \theta} \frac{\partial \ell\left(a_t, x_t; \theta_0, g_{\theta_0}\right)}{\partial \theta}^{\top}\right] + \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}(T-t) \times \left(\begin{array}{c} E\left[\frac{\partial \ell\left(a_t, x_t; \theta_0, g_{\theta_0}\right)}{\partial \theta} \frac{\partial \ell\left(x_0, a_0; \theta_0, g_{\theta_0}\right)}{\partial \theta}^{\top}\right] \\ + E\left[\frac{\partial \ell\left(a_t, x_t; \theta_0, g_{\theta_0}\right)}{\partial \theta} \frac{\partial \ell\left(x_0, a_0; \theta_0, g_{\theta_0}\right)}{\partial \theta}^{\top}\right]^{\top} \end{array}\right).$$

This follows immediately from the CLT for stationary and geometric mixing process. Also, $D_{3,T}$ is $o_p(1)$ since $\frac{\partial \ell\left(a_t, x_t; \theta_0, g_{\theta_0}\right)}{\partial \theta}$ is uniformly bounded and $\mathfrak{d}_{t,T} = o_p\left(\sqrt{T}\right)$ for all t. So we focus on $D_{2,T}$. We proceed by writing $D_{2,T}$ as a finite linear combination of U-statistics and show their leading terms have a normal limiting distribution. Consider the j-th element of $D_{2,T}$, by linearizing the score function

$$\begin{aligned}
\left(D_{2,T}\right)_j &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \mathfrak{c}_{t,T}\left(\frac{\partial \widehat{v}_{\theta_0}(a_t, x_t)}{\partial \theta_j} - \frac{\partial v_{\theta_0}(a_t, x_t)}{\partial \theta_j}\right) \\
&\quad - \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\sum_{a\in A} \mathfrak{c}_{t,T}\psi_1(a, x_t) \times \left(\frac{\partial \widehat{v}_{\theta_0}(a, x_t)}{\partial \theta_j} - \frac{\partial v_{\theta_0}(a, x_t)}{\partial \theta_j}\right) \\
&\quad - \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\sum_{a\in A} \mathfrak{c}_{t,T}\psi_{2,j}(a, x_t) \times \left(\widehat{v}_{\theta_0}(a, x_t) - v_{\theta_0}(a, x_t)\right) \\
&\quad + \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\sum_{a\in A} \mathfrak{c}_{t,T}\psi_{2,j}(a, x_t) \times \left(\sum_{\widetilde{a}\in A} P(\widetilde{a}|x_t)\left(\widehat{v}_{\theta_0}(\widetilde{a}, x_t) - v_{\theta_0}(\widetilde{a}, x_t)\right)\right) + o_p(1) \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(E_{1,t,T}\right)_j + \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(E_{2,t,T}\right)_j + \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(E_{3,t,T}\right)_j + \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(E_{4,t,T}\right)_j + o_p(1),
\end{aligned}$$

where

$$\psi_1(a, x_t) = P(a|x_t), \tag{55}$$

$$\psi_{2,j}(a, x_t) = P(a|x_t)\frac{\partial v_{\theta_0}(a, x_t)}{\partial \theta_j}, \tag{56}$$

and the remainder terms are of smaller order since our nonparametric estimates converge uniformly to the true at the rate faster than $T^{-1/4}$ on the trimming set, as proven in Theorems 1 and 2.

The asymptotic properties of these terms are tedious but simple to obtain by repeatedly utilizing the projection results and law of large numbers for U-statistics, see Lee (1990). We also note that all of the relevant kernels for our statistics are uniformly bounded, along with the assumption [B1], this ensures the residuals from the projections can be ignored. Now we give some details for deriving the distribution of $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(E_{1,t,T}\right)_j$. We consider the normalized sum of $\frac{\partial^k \widehat{g}_{\theta_0}}{\partial \theta_j^k} - \frac{\partial^k g_{\theta_0}}{\partial \theta_j^k}$ for $k = 0, 1$, whose leading terms are

$$\left(\widehat{\mathcal{H}} - \mathcal{H}\right)\frac{\partial^k m_{\theta_0}}{\partial \theta_j^k} + \mathcal{H}(I-\mathcal{L})^{-1}\left(\widehat{\mathcal{L}} - \mathcal{L}\right)\frac{\partial^k m_{\theta_0}}{\partial \theta_j^k} + \mathcal{H}(I-\mathcal{L})^{-1} \times \left(\frac{\partial^k \widehat{r}_{\theta_0}}{\partial \theta_j^k} - \frac{\partial^k r_{\theta_0}}{\partial \theta_j^k}\right).$$

First consider the normalized sum of $(\widehat{\mathcal{H}} - \mathcal{H})\frac{\partial^k m_{\theta_0}}{\partial \theta_j^k}$, with further linearization, see the decomposition $\widehat{\mathcal{L}} - \mathcal{L}$ and $\widehat{\mathcal{H}} - \mathcal{H}$ in the proof of [A1], we obtain the (scaled) U-statistics $\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\mathfrak{c}_{t,T}\left[\left(\widehat{\mathcal{H}} - \mathcal{H}\right)\frac{\partial m_{\theta_0}}{\partial \theta_j}(x_t, a_t)\right]$ as given in Box I. Hoeffding (H-)decomposition provides the following as leading term, note that the bias is asymptotically negligible under assumptions B6 and B7, after disposing the trimming factors

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\left(\frac{\partial m_{\theta_0}(x_{t+1})}{\partial \theta_j} - E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial \theta_j}\right| x_t, a_t\right]\right). \tag{57}$$

To obtain the projection of the second term is more labor intensive. We first split it up into two parts,

$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\mathfrak{c}_{t,T}\left[\mathcal{H}(I-\mathcal{L})^{-1}\left(\widehat{\mathcal{L}} - \mathcal{L}\right)\frac{\partial^k m_{\theta_0}}{\partial \theta_j^k}(x_t, a_t)\right] \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\mathfrak{c}_{t,T}\left[\mathcal{H}\left(\widehat{\mathcal{L}} - \mathcal{L}\right)\frac{\partial^k m_{\theta_0}}{\partial \theta_j^k}(x_t, a_t)\right] \\
&\quad + \frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\mathfrak{c}_{t,T}\left[\mathcal{H}\mathcal{L}(I-\mathcal{L})^{-1}\left(\widehat{\mathcal{L}} - \mathcal{L}\right)\frac{\partial^k m_{\theta_0}}{\partial \theta_j^k}(x_t, a_t)\right].
\end{aligned}$$

The summands of the first term takes the form given in Box II with standard change of variable and usual symmetrization. This is a U-statistic with kernel (see Box III), whose leading term from the H-decomposition gives rise to the following centered process

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\beta\left(\frac{\partial m_{\theta_0}(x_{t+1})}{\partial \theta_j} - E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial \theta_j}\right| x_t\right]\right), \tag{58}$$

notice the conditional expectation term is a two-step ahead predictor, zero mean follows from stationarity assumption and the law of iterated expectation. As for the second part of the second term, using the Neumann series representation discussed in the proof of A2, the kernel of the relevant U-statistics is given in Box IV where λ is defined as the limit of discounted sum of the conditional densities, λ_T, defined in (50). The leading term of the projection of the U-statistic with the above kernel is

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\left(\frac{\beta^2}{1-\beta}\right)\left(\frac{\partial m_{\theta_0}(x_{t+1})}{\partial \theta_j} - E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial \theta_j}\right| x_t\right]\right). \tag{59}$$

The last term of $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(E_{1,t,T}\right)_j$ can be treated similarly, recall we have

$$\begin{aligned}
&\mathcal{H}(I-\mathcal{L})^{-1}\left(\frac{\partial^k \widehat{r}_{\theta_0}}{\partial \theta_j^k} - \frac{\partial^k r_{\theta_0}}{\partial \theta_j^k}\right) \\
&= \mathcal{H}\left(\frac{\partial^k \widehat{r}_{\theta_0}}{\partial \theta_j^k} - \frac{\partial^k r_{\theta_0}}{\partial \theta_j^k}\right) + \mathcal{H}\mathcal{L}(I-\mathcal{L})^{-1}\left(\frac{\partial^k \widehat{r}_{\theta_0}}{\partial \theta_j^k} - \frac{\partial^k r_{\theta_0}}{\partial \theta_j^k}\right),
\end{aligned}$$

then

$$\begin{aligned}
&\mathcal{H}\left(\frac{\partial^k \widehat{r}_{\theta_0}}{\partial \theta_j^k} - \frac{\partial^k r_{\theta_0}}{\partial \theta_j^k}\right)(x_t, a_t) \\
&= \frac{1}{T-1}\sum_{a\in A}\sum_{s\neq t}\int\left(\frac{\partial^k \zeta'_{a,x',\theta_0}\left(P\left(a|x'\right)\right)}{\partial \theta_j^k}\frac{e_{a,s}K_h\left(x_s - x'\right)}{f_X\left(x'\right)}\right) \times f'_{X|X,A}\left(dx'|x_t, a_t\right) + o_p\left(T^{-1/2}\right),
\end{aligned}$$

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\mathfrak{c}_{t,T}\left[\left(\widehat{\mathcal{H}}-\mathcal{H}\right)\frac{\partial m_{\theta_0}}{\partial\theta_j}(x_t,a_t)\right]$$
$$=\frac{1}{\sqrt{T}}\frac{1}{T-1}\sum_{t=1}^{T-1}\sum_{s\neq t}\mathfrak{c}_{t,T}\left(\begin{array}{c}\frac{\partial m_{\theta_0}(x_{s+1})}{\partial\theta_j}\frac{K_h(x_s-x_t)\mathbf{1}[a_s=a_t]}{f_{X,A}(x_t,a_t)}-E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\right|x_t,a_t\right]\\-\frac{K_h(x_s-x_t)\mathbf{1}[a_s=a_t]-f_{X,A}(x_t,a_t)}{f_{X,A}(x_t,a_t)}E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\right|x_t,a_t\right]\end{array}\right)+o_p(1)$$
$$=\sqrt{T}\binom{T-1}{2}^{-1}\sum_{t=1}^{T-1}\sum_{s>t}\frac{1}{2}\left(\begin{array}{c}\mathfrak{c}_{t,T}\frac{\partial m_{\theta_0}(x_{s+1})}{\partial\theta_j}\frac{K_h(x_s-x_t)\mathbf{1}[a_s=a_t]}{f_{X,A}(x_t,a_t)}-\mathfrak{c}_{t,T}E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\right|x_t,a_t\right]\\+\mathfrak{c}_{s,T}\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\frac{K_h(x_t-x_s)\mathbf{1}[a_t=a_s]}{f(x_s,a_s)}-\mathfrak{c}_{s,T}E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\right|x_s,a_s\right]\end{array}\right)$$
$$-\sqrt{T}\binom{T-1}{2}^{-1}\sum_{t=1}^{T-1}\sum_{s>t}\frac{1}{2}\left(\begin{array}{c}\mathfrak{c}_{t,T}\frac{K_h(x_s-x_t)\mathbf{1}[a_s=a_t]-f_{X,A}(x_t,a_t)}{f_{X,A}(x_t,a_t)}E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\right|x_t,a_t\right]\\+\mathfrak{c}_{s,T}\frac{K_h(x_t-x_s)\mathbf{1}[a_t=a_s]-f(x_s,a_s)}{f(x_s,a_s)}E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\right|x_s,a_s\right]\end{array}\right)+o_p(1).$$

Box I.

$$\mathfrak{c}_{t,T}\beta\int\left(\begin{array}{c}\int\frac{\partial m_{\theta_0}(x'')}{\partial\theta_j}\frac{1}{f_X(x')}\left(\widehat{f}_{X'X}(dx'',x')-f_{X'X}(dx'',x')\right)\\-\frac{\widehat{f}_X(x')-f_X(x')}{f_X(x')}E\left[\left.\frac{\partial m_{\theta_0}(x_{t+2})}{\partial\theta_j}\right|x_{t+1}=x'\right]\end{array}\right)f_{X'|X,A}(dx'|x_t,a_t),$$

Box II.

$$\frac{\beta}{2}\left(\begin{array}{c}\mathfrak{c}_{t,T}\frac{\partial m_{\theta_0}(x_{s+1})}{\partial\theta_j}\frac{f'_{X|X,A}(x_s|x_t,a_t)}{f_X(x_s)}-\mathfrak{c}_{t,T}E\left[E\left[\left.\frac{\partial m_{\theta_0}(x_{t+2})}{\partial\theta_j}\right|x_{t+1}\right]\middle|x_t,a_t\right]\\+\mathfrak{c}_{s,T}\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\frac{f'_{X|X,A}(x_t|x_s,a_s)}{f_X(x_t)}-\mathfrak{c}_{s,T}E\left[E\left[\left.\frac{\partial m_{\theta_0}(x_{s+2})}{\partial\theta_j}\right|x_{s+1}\right]\middle|x_s,a_s\right]\end{array}\right)$$
$$-\frac{\beta}{2}\left(\begin{array}{c}\mathfrak{c}_{t,T}E\left[\left.\frac{\partial m_{\theta_0}(x_{s+1})}{\partial\theta_j}\right|x_s\right]\frac{f'_{X|X,A}(x_s|x_t,a_t)}{f_X(x_s)}-\mathfrak{c}_{t,T}E\left[E\left[\left.\frac{\partial m_{\theta_0}(x_{t+2})}{\partial\theta_j}\right|x_{t+1}\right]\middle|x_t,a_t\right]\\+\mathfrak{c}_{s,T}E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\right|x_t\right]\frac{f'_{X|X,A}(x_t|x_s,a_s)}{f_X(x_t)}-\mathfrak{c}_{s,T}E\left[E\left[\left.\frac{\partial m_{\theta_0}(x_{s+2})}{\partial\theta_j}\right|x_{s+1}\right]\middle|x_s,a_s\right]\end{array}\right),$$

Box III.

$$\frac{\beta}{2}\left(\begin{array}{c}\mathfrak{c}_{t,T}\frac{\partial m_{\theta_0}(x_{s+1})}{\partial\theta_j}\int\frac{\lambda(x_s,x')}{f_X(x_s)}f'_{X|X,A}(dx'|x_t,a_t)-\mathfrak{c}_{t,T}\sum_{\tau=1}^{\infty}\beta^{\tau}E\left[E\left[\left.\frac{\partial m_{\theta_0}(x_{t+\tau+2})}{\partial\theta_j}\right|x_{t+1}\right]\middle|x_t,a_t\right]\\+\mathfrak{c}_{s,T}\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\int\frac{\lambda(x_t,x')}{f_X(x_t)}f'_{X|X,A}(dx'|x_s,a_s)-\mathfrak{c}_{s,T}\sum_{\tau=1}^{\infty}\beta^{\tau}E\left[E\left[\left.\frac{\partial m_{\theta_0}(x_{s+\tau+2})}{\partial\theta_j}\right|x_{s+1}\right]\middle|x_s,a_s\right]\end{array}\right)$$
$$-\frac{\beta}{2}\left(\begin{array}{c}\mathfrak{c}_{t,T}E\left[\left.\frac{\partial m_{\theta_0}(x_{s+1})}{\partial\theta_j}\right|x_s\right]\int\frac{\lambda(x_s,x')}{f_X(x_s)}f'_{X|X,A}(dx'|x_t,a_t)-\mathfrak{c}_{t,T}\sum_{\tau=1}^{\infty}\beta^{\tau}E\left[E\left[\left.\frac{\partial m_{\theta_0}(x_{t+\tau+2})}{\partial\theta_j}\right|x_{t+1}\right]\middle|x_t,a_t\right]\\+\mathfrak{c}_{s,T}E\left[\left.\frac{\partial m_{\theta_0}(x_{t+1})}{\partial\theta_j}\right|x_t\right]\int\frac{\lambda(x_t,x')}{f_X(x_t)}f'_{X|X,A}(dx'|x_s,a_s)-\mathfrak{c}_{s,T}\sum_{\tau=1}^{\infty}\beta^{\tau}E\left[E\left[\left.\frac{\partial m_{\theta_0}(x_{s+\tau+2})}{\partial\theta_j}\right|x_{s+1}\right]\middle|x_s,a_s\right]\end{array}\right),$$

Box IV.

$$=\frac{1}{T-1}\sum_{a\in A}\sum_{s\neq t}f'_{X|X,A}(x_s|x_t,a_t)$$
$$\times\frac{\partial^k\zeta'_{a,x_s,\theta_0}(P(a|x))}{\partial\theta_j^k}\frac{e_{a,s}}{f_X(x_s)}+o_p\left(T^{-1/2}\right).$$

Normalizing the projection of the corresponding U-statistics obtains

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathcal{H}\left(\frac{\partial^k\widehat{r}_{\theta_0}}{\partial\theta_j^k}-\frac{\partial^k r_{\theta_0}}{\partial\theta_j^k}\right)(x_t,a_t)$$

$$=\frac{1}{\sqrt{T}}\sum_{a\in A}\sum_{t=1}^{T}\frac{\partial^k\zeta'_{a,x_t,\theta_0}(P(a|x_t))}{\partial\theta_j^k}e_{a,t}+o_p(1). \quad (60)$$

We do the same for the remaining terms, in particular we show that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T-1}\mathcal{H}\mathcal{L}(I-\mathcal{L})^{-1}\left(\frac{\partial^k\widehat{r}_{\theta_0}}{\partial\theta_j^k}-\frac{\partial^k r_{\theta_0}}{\partial\theta_j^k}\right)(x_t,a_t)$$

$$= \frac{1}{\sqrt{T}} \sum_{a \in A} \sum_{t=1}^{T} \frac{\beta}{1-\beta} \frac{\partial^k \zeta'_{a,x_t,\theta_0}(P(a|x_t))}{\partial \theta_j^k} e_{a,t} + o_p(1). \quad (61)$$

Collecting (57)–(61), for $k = 1$, we obtain the leading terms of $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(E_{1,t,T}\right)_j$. For $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(E_{2,t,T}\right)_j$ and $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(E_{3,t,T}\right)_j$, we again use the projection technique for the U-statistics to obtain their leading terms. We have provided a detailed analysis for the former case as the remaining terms in $\left(D_{2,T}\right)$ can be treated in a similar fashion. In particular, it is simple to show that the projections of various relevant U-statistics, defined below with some elements $\varpi_k \in C(X)$, $\varsigma_k \in C(A \times X)$ and $a \in A$, have the following linear representations

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varsigma_k(x_t, a) \left[\left(\widehat{\mathcal{H}} - \mathcal{H} \right) \varpi_k(x_t, a) \right]$$
$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{T-1} \left(\frac{\varsigma_k(x_t, a) f_X(x_t) \mathbf{1}[a_t = a]}{f(x_t, a)} \right)$$
$$\times \left(\varpi_k(x_{t+1}) - E\left[\varpi_k(x_{t+1}) | x_t, a_t = a \right] \right) + o_p(1)$$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varsigma_k(x_t, a) \left[\mathcal{H} \left(\widehat{\mathcal{L}} - \mathcal{L} \right) \varpi_k(x_t, a) \right]$$
$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{T-1} \beta \left(\frac{\int \varsigma_k(v, a) f'_{X|X,A}(x_t|v, a) f_X(dv)}{f_X(x_t)} \right)$$
$$\times \left(\varpi_k(x_{t+1}) - E\left[\varpi_k(x_{t+1}) | x_t \right] \right) + o_p(1)$$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varsigma_k(x_t, a) \left[\mathcal{H}\mathcal{L}(I - \mathcal{L})^{-1} \left(\widehat{\mathcal{L}} - \mathcal{L} \right) \varpi_k(x_t, a) \right]$$
$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{T-1} \beta$$
$$\times \left(\frac{\iint \varsigma_k(v, a) \varphi(x_t|w) f'_{X|X,A}(dw|dv, a) f_X(dv)}{f_X(x_t)} \right)$$
$$\times \left(\varpi_k(x_{t+1}) - E\left[\varpi_k(x_{t+1}) | x_t \right] \right) + o_p(1).$$

In correspondence of $\left(E_{k+1,t,T}\right)_j$ for $k = 1, 2$, we have: $\varsigma_1(\cdot) = \psi_1(a, \cdot)$, $\varsigma_2(\cdot) = \psi_{2,j}(a, \cdot)$, $\varpi_1(\cdot) = \frac{\partial m_{\theta_0}(\cdot)}{\partial \theta_j}$, and $\varpi_2(\cdot) = m_{\theta_0}(\cdot)$, where ψ_1 and $\psi_{2,j}$ are defined in (55) and (56). Similarly, we also have

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varsigma_k(x_t, a) \mathcal{H} \left(\frac{\partial^k \widehat{r}_{\theta_0}}{\partial \theta_j^k} - \frac{\partial^k r_{\theta_0}}{\partial \theta_j^k} \right)(x_t, a)$$
$$= \frac{1}{\sqrt{T}} \sum_{\widetilde{a} \in A} \sum_{t=1}^{T-1} \left(\int \varsigma_k(v, a) f'_{X|X,A}(x_t|v, a) f_X(dv) \right)$$
$$\times \frac{\partial^k \zeta'_{x_t,a,\theta_0}(P(\widetilde{a}|x_t))}{\partial \theta_j^k} \frac{e_{\widetilde{a},t}}{f_X(x_t)} + o_p(1).$$

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varsigma_k(x_t, a) \left[\mathcal{H}\mathcal{L}(I - \mathcal{L})^{-1} \left(\frac{\partial^k \widehat{r}_{\theta_0}}{\partial \theta_j^k} - \frac{\partial^k r_{\theta_0}}{\partial \theta_j^k} \right)(x_t, a) \right]$$
$$= \frac{1}{\sqrt{T}} \sum_{\widetilde{a} \in A} \sum_{t=1}^{T} \left[\iint \varsigma_k(v, a) \varphi(x_t|w) f'_{X|X,A} \right.$$
$$\left. \times (dw|v, a) f_X(dv) \right] \frac{\partial^k \zeta'_{x_t,a,\theta_0}(P(\widetilde{a}|x_t))}{\partial \theta_j^k} \frac{e_{\widetilde{a},t}}{f_X(x_t)} + o_p(1).$$

From all the projections above, their leading terms form a finite linear combination of mean zero processes each satisfying the CLT. Therefore $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(E_{k,t,T}\right)_j = O_p(1)$ for $k = 1, 2, 3$ and $j = 1, \ldots, J$, and by Cramer–Wold device

$$\sqrt{T} D_{2,T} \Rightarrow N(0, \Lambda_2), \quad \text{where}$$
$$\Lambda_2 = \lim_{T \to \infty} \operatorname{var} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(E_{1,t,T} + E_{2,t,T} + E_{3,t,T} \right) \right).$$

In sum, we have

$$D_{1,T} + D_{2,T} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(\frac{\partial \ell\left(a_t, x_t; \theta_0, g_{\theta_0}\right)}{\partial \theta} \right.$$
$$\left. + E_{1,t,T} + E_{2,t,T} + E_{3,t,T} \right) + o_p(1) \Rightarrow N(0, \mathcal{I}). \quad \square$$

Proof of Theorems 4 and 5. Under the assumed smoothness assumptions, the results follow from MVT. □

Appendix B

We now show various centered processes in the previous section converge uniformly at desired rates on a compact set X. We outline the main steps below and proof the results for relevant cases. Our approach here is similar to the analysis in [LM], where they employ the exponential inequalities from [B] for various quantities similar to ours.

Consider some process $\mathrm{l}_T(x) = \frac{1}{T} \sum l(x_t, x)$, where $l(x_t, x)$ has mean zero. For some positive sequence, δ_T, converging monotonically to zero, we first show that $|\mathrm{l}_T(x)| = o_p(\delta_T)$ pointwise on X, then we use the continuity property of $l(x_t, x)$ to show that this rate of convergence is preserved uniformly over X.

To obtain the pointwise rates, specializing Theorem 1.3 of [B], we have the following inequality.

$$\Pr\left(|\mathrm{l}_T(x)| > \delta_T\right) \le 4 \exp\left(-\frac{\delta_T^2 T^\beta}{8 v^2\left(T^\beta\right)} \right)$$
$$+ 22 \left(1 + \frac{4 b_T}{\delta_T} \right)^{1/2} T^\beta \alpha \left(\left\lfloor \frac{T^{1-\beta}}{2} \right\rfloor \right)$$
$$\le \exp\left(-G_{1,T}\right) + G_{2,T},$$

for some $\beta \in (0, 1)$,

$$b_T = O\left(\sup_{(x',x) \in \mathcal{X} \times \mathcal{X}} l\left(x', x\right) \right),$$

$$v^2(\beta) = \operatorname{var}\left(\frac{1}{\left\lfloor \frac{T^\beta}{2} + 1 \right\rfloor} \sum_{t=1}^{\left\lfloor \frac{T^\beta}{2} + 1 \right\rfloor} l(x_t, x) \right) + \frac{b_T \delta_T}{2}. \quad (62)$$

We need $G_{1,T} \to \infty$ for the exponential term to converge to zero. The main calculation required here is the variance term in v^2. Following [M], we can generally show that the uniform order of such term comes from the variances and the covariance terms are of smaller order. We note that the bounds on these variances are independent on the trimming set. For our purposes, the natural choice of δ_T^2often reduces us to choosing β to satisfy $b\delta_T = o\left(\delta_T^2 T^\beta\right)$. The rate of $G_{2,T}$ is easy to control since all of the quantities involved increase (decrease) at a power rate, the mixing coefficient can be made to decay sufficiently fast so $G_{2,T} = O\left(T^{-\eta}\right)$ for some $\eta > 0$, hence $\Pr\left(|\mathrm{l}_T(x)| > \delta_T\right) = O\left(T^{-\eta}\right)$.

To obtain the uniform rates over $\mathcal{X}$, compactness implies there exist an increasing number, N_T, of shrinking hyper-cubes $\left\{I_{n,N_T}\right\}$ whose length of each side is $\{\epsilon_T\}$ with centers $\{x^{n,N_T}\}$. These cubes cover $\mathcal{X}$, namely for some C_0 and d, $\epsilon_T^d N_T \le C_0 < \infty$. In particular, we will have N_T grow at a power rate in our applications. Then we have

$$\Pr\left(\sup_{x\in\mathcal{X}}\left|\mathrm{l}_T(x)\right|>\delta_T\right)$$
$$\leq \Pr\left(\max_{1\leq n\leq N_T}\left|\mathrm{l}_T\left(x^{n,N_T}\right)\right|>\delta_T\right)$$
$$+\Pr\left(\max_{1\leq n\leq N_T}\sup_{x\in I_{n,N_T}}\left|\mathrm{l}_T(x)-\mathrm{l}_T\left(x^{n,N_T}\right)\right|>\delta_T\right)$$
$$=G_{3,T}+G_{4,T},$$

where $G_{3,T}=O\left(N_TT^{-\eta}\right)$ by Bonferroni Inequality. Provided the rate of decay of the mixing coefficient, i.e. η, is sufficiently large relative to the rate N_T grows we shall have $N_T=o\left(T^{\eta}\right)$. For the second term, since the opposing behavior of (ϵ_T,N_T) is independent of the mixing coefficient, $\max_{1\leq n\leq N_T}\sup_{x\in I_{n,N_T}}\left|\mathrm{l}_T(x)-\mathrm{l}_T\left(x^{n,N_T}\right)\right|=o\left(\delta_T\right)$ can be shown using Lipschitz continuity when the hyper cubes shrink sufficiently fast.

Before we proceed with the specific cases we validate our treatment of the trimming factor. The pointwise rates are clearly unaffected by bias at the boundary so long $x\in\mathcal{X}$. The technique used to obtain uniformity also accommodates expanding sets X_T, so long we use the sequence $\{c_T\}$ to satisfy condition stated in [B9]. The uniform rate of convergence is also unaffected, when replace $\mathcal{X}$ with X_T, since the covering of an expanding of a compact subsets of a compact set can still grow (and shrink) at the same rate in each of the cases below. Therefore we can simply replace $\mathcal{X}$ everywhere by X_T.

Combining the results of uniform convergence of the zero mean processes and their biases, the uniform rates to various quantities in the previous section can now be established. We note that the treatment to allow for additional discrete observable states only requires trivial extension. We provide illustrate this for the first case of kernel density estimation, and for brevity, thenceforth assume that we only have purely continuous observable state variables.

Case 1. Density estimators: $\widehat{f}_X(x)$, $\widehat{f}_{X',X}\left(x',x\right)$ and $\widehat{f}_{X',X,A}\left(x',x,a\right)$.

We first establish the pointwise rate of convergence of a demeaned kernel density estimator.

$$\mathrm{l}_T(x)=\widehat{f}_X(x)-E\widehat{f}_X(x),$$
$$l\left(x_t,x\right)=K_h\left(x_t-x\right)-EK_h\left(x_t-x\right)$$

when (x_t,x) are d-dimensional vectors we use a product kernel $K_h(x_t-x)=\prod_{l=1}^{d}K_h(x_t^{(l)}-x^{(l)})$. The main elements for studying the rate of $G_{1,T}$ are: $\varpi=\frac{1}{\sqrt{Th_T^d}}$, $\delta_T=T^{\xi}\varpi$ for some $\xi>0$, $b_T=O\left(h_T^{-d}\right)$, and $\upsilon^2\left(T^{\beta}\right)=O\left(\varpi^2T^{1-\beta}\vee T^{\xi}\varpi h_T^{-d}\right)$. We obtain from simple algebra

$$G_{1,T}=O\left(\frac{T^{2\xi}T^{\beta}}{T^{1-\beta}+T^{\xi}T^{1/2}h_T^{-d/2}}\right).$$

As mentioned in the previous section, we have $d=2$ and $h_T=O\left(T^{-1/5}\right)$. This means $\delta_T=T^{\xi-3/10}$, and if $\beta\in(7/10,1)$ then we have $G_{1,T}\to\infty$. Clearly, the same choice of β will suffice for $d=1$ as well.

To make this uniform on X_T, with product kernels and the Lipschitz continuity of K, for any $\left(x,x^{n,N_T}\right)$ we have

$$\left|K_h\left(x_t-x\right)-K_h\left(x_t-x^{n,N_T}\right)\right|\leq\frac{C_0}{h_T^3}\epsilon_T.$$

So it follows that

$$\delta_T^{-1}\max_{1\leq n\leq N_T}\sup_{x\in I_{n,N_T}}\left|\mathrm{l}_T(x)-\mathrm{l}_T\left(x^{n,N_T}\right)\right|$$
$$=O\left(\frac{\epsilon_T}{\delta_Th_T^3}\right)=O\left(\frac{T^{-\zeta/2}}{T^{\xi-9/10}}\right).$$

Define $N_T=T^{\zeta}$, for some $\zeta>0$, this requires $9/5<\zeta<\eta$.

We can allow for additional discrete control variable and/or observable state variables. As an illustration, consider the density estimator of one continuous random variable and some discrete random variable, we have

$$\mathrm{l}_T(x)=\widehat{f}_{X^C,X^D}\left(x_c,x_d\right)-E\widehat{f}_{X^C,X^D}\left(x_c,x_d\right),$$
$$l\left(x_t,x\right)=K_h\left(x_{c,t}-x_c\right)\mathbf{1}\left(x_{d,t}=x_d\right)$$
$$-EK_h\left(x_{c,t}-x_c\right)\mathbf{1}\left(x_{d,t}=x_d\right).$$

Same rates as the purely continuous case apply. For the pointwise part, the variance is clearly of the same order. For the bounds on the uniform rates observe that,

$$\left|K_h\left(x_{c,t}-x_c\right)\mathbf{1}\left(x_{d,t}=x_d\right)-K_h\left(x_{c,t}-x_c^{n,N_T}\right)\mathbf{1}\left(x_{d,t}=x_d\right)\right|$$
$$\leq\left|K_h\left(x_{c,t}-x_c\right)-K_h\left(x_{c,t}-x_c^{n,N_T}\right)\right|.$$

Same reasoning also applies for the kernel estimator of the density of the control and observable state variables.

Case 2. $\widehat{r}^C(x)$. For any $a\in A$

$$l\left(x_t,x\right)=\frac{e_{a,t}K_h\left(x_t-x\right)}{f_X(x)}.$$

Since $\left\{e_{a,T}\right\}$ is uniformly bounded (a.s.) it follows, as shown in Case 1, the choice $\beta\in(3/5,1)$ will suffice to have $G_{1,T}\to\infty$.

To make this uniform on X_T, by boundedness of $\left\{e_{a,T}\right\}$, Lipschitz continuity of K,f and their appropriate bounds, we have for any $\left(x,x^{n,N_T}\right)\in I_{n,N_T}$,

$$\left|K_h\left(x_t-x\right)-K_h\left(x_t-x^{n,N_T}\right)\right|\leq\frac{C_2}{h_T^2}\epsilon_T.$$

So it follows that

$$\delta_T^{-1}\max_{1\leq n\leq N_T}\sup_{x\in I_{n,N_T}}\left|\mathrm{l}_T(x)-\mathrm{l}_T\left(x_n\right)\right|=O\left(\frac{T^{-\zeta}}{T^{\xi-7/10}}\right)=o(1),$$

for some $\zeta>0$, this requires $7/10<\zeta<\eta$.

Case 3. $\mathcal{L}\left(I-\mathcal{L}\right)^{-1}\widehat{r}_{\theta}^C(x)$. For any $a\in A$, we have

$$l\left(x_t,x\right)=e_{a,t}\nu_{a,\theta}\left(x_t,x\right),$$

where the definition of $\nu_{a,\theta}$ is provided in the proof of (A2). Using Billingsley's Inequality, it is straightforward to show that, with the additional smoothing, the variance of $\mathrm{l}_T(x)$ converges at the parametric rate on X_T. Selecting $\beta\in(1/2,1)$ will yield $G_{1,T}\to\infty$ for $\Pr\left(\left|\mathrm{l}_T(x)\right|>T^{-2/5}\right)=o(1)$, for any $x\in X_T$.

To make this uniform on X_T, by boundedness of $\left\{e_{a,T}\right\}$ and Lipschitz continuity ofλ, we have for any $\left(x,x^{n,N_T}\right)\in I_{n,N_T}$,

$$\left|e_{a,t}\nu\left(x_t,x\right)-e_{a,t}\nu\left(x_t,x^{n,N_T}\right)\right|\leq C_3\epsilon_T.$$

So it follows that

$$\delta_T^{-1}\max_{1\leq n\leq N_T}\sup_{x\in I_{n,N_T}}\left|\mathrm{l}_T(x)-\mathrm{l}_T\left(x_n\right)\right|=O\left(\frac{T^{-\zeta}}{T^{-2/5}}\right),$$

for some $\zeta>0$, this requires $2/5<\zeta<\eta$.

Case 4. $m_{1,\theta}^C(x)$. Here we have

$$\mathrm{l}_T(x)=\frac{\beta}{f_X(x)}\int\left(\widehat{f}_{X',X}\left(x',x\right)-E\widehat{f}_{X',X}\left(x',x\right)\right)m_{\theta}\left(x'\right)dx'.$$

As mentioned in the previous section, under our smoothness assumptions, we have uniformly on X_T that

$$\int \widehat{f}_{X',X}\left(x',x\right) m_\theta\left(x'\right) dx'$$
$$= \frac{1}{T-1}\sum_{t=1}^{T-1} K_h\left(x_t - x\right) m_\theta\left(x_{t+1}\right) + O\left(h_T^2\right).$$

The same choices of bounding parameters used in Case 2 directly apply. □

References

Aguirregabiria, V., Mira, P., 2002. Swapping nested fixed point algorithm: a class of estimators for discrete Markov decision models. Econometrica 70, 1519–1543.

Aguirregabiria, V., Mira, P., 2007. Sequential estimation of dynamic discrete games. Econometrica 75, 1–53.

Aguirregabiria, V., Mira, P., 2010. Dynamic discrete choice structural models: a survey. Journal of Econometrics 156, 38–67.

Altug, S., Miller, R.A., 1998. The effect of work experience on female wages and labour supply. Review of Economic Studies 65, 45–85.

Andrews, D.W.K., 2005. Higher-order improvements of the parametric bootstrap for Markov processes. In: J.H. Stock and D.W.K. Andrews, (Eds.), Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, Cambridge, pp. 171–215.

Arcidiacono, P., Miller, R.A., 2010. CCP estimation of dynamic discrete choice models with unobserved heterogeneity. Working Paper. Duke University.

Bajari, P., Benkard, C.L., Levin, J., 2007. Estimating dynamic models of imperfect competition. Econometrica 75, 1331–1370.

Bajari, P., Chernozhukov, V., Hong, H., Nekipelov, D., 2008. Semiparametric estimation of a dynamic game of incomplete information. Working Paper. University of Minnesota.

Bellman, R.E., 1957. Dynamic Programming. Princeton University Press.

Bertsekas, D.P., Shreve, S., 1978. Stochastic Optimal Control: The Discrete-Time Case. Academic Press.

Bosq, D., 1998. Nonparametric Statistics for Stochastic Processes. Springer.

Carrasco, M., Florens, J.P., Renault, E., 2007. Linear inverse problems and structural econometrics estimation based on spectral decomposition and regularization. In: Heckman, J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 6. North-Holland, Amsterdam, pp. 5633–5751.

Chen, X., Linton, O.B., van Keilegom, I., 2003. Estimation of semiparametric models when the criterion function is not smooth. Econometrica 71, 1591–1608.

Chen, X., Pouzo, D., 2009. Efficient estimation of semiparametric conditional moment models with possibly nonsmooth residuals. Journal of Econometrics 152, 46–60.

Hajivassiliou, V., Ruud, P., 1994. Classical estimation methods for LDV models using simulation. In: Engle, R.F., McFadden, D.L. (Eds.), Handbook of Econometrics, vol. 4. North-Holland, Amsterdam, pp. 2383–2441.

Horowitz, J.L., 2003. Bootstrap methods for Markov processes. Econometrica 71, 1049–1082.

Hotz, V., Miller, R.A., 1993. Conditional choice probabilities and the estimation of dynamic models. Review of Economic Studies 60, 497–531.

Hotz, V., Miller, R.A., Smith, S., Smith, J., 1994. A simulation estimator for dynamic models of discrete choice. Review of Economic Studies 61, 265–289.

Judd, K., 1998. Numerical Methods in Economics. MIT Press.

Kasahara, H., Shimotsu, K., 2008. Pseudo-likelihood estimation and bootstrap inference for structural discrete Markov decision models. Journal of Econometrics 146, 92–106.

Kress, R., 1999. Linear Integral Equations. Springer.

Li, Q., Racine, J.S., 2006. Nonparametric Econometrics. Princeton University Press.

Linton, O.B., Mammen, E., 2003. Estimating semiparametric ARCH(∞) models by kernel smoothing methods. STICERD Working Paper EM/2003/453.

Linton, O.B., Mammen, E., 2005. Estimating semiparametric ARCH(∞) models by kernel smoothing methods. Econometrica 73, 771–836.

Masry, E., 1996. Multivariate local polynomial regression for time series: uniform strong consistency and rates. Journal of Time Series Analysis 17, 571–599.

McFadden, D.L., 1974. Conditional logit analysis of qualitative choice behavior. Frontier in Econometrics. New York.

Miller, R.A., 1997. Estimating models of dynamic optimization with microeconomic data. In: Pesaran, M.H., Schmidt, P. (Eds.), Handbook of Applied Econometrics, vol. 2. Basil Blackwell, pp. 246–299.

Newey, W.K., McFadden, D.L., 1994. Large sample estimation and hypothesis testing. In: Engle, R.F., McFadden, D.L. (Eds.), Handbook of Econometrics, vol. 4. North-Holland, Amsterdam, pp. 2111–2245.

Ouyang, D., Li, Q., Racine, J.S., 2006. Cross-validation and the estimation of probability distributions with categorical data. Journal of Nonparametric Statistics 18, 69–100.

Pakes, A., Olley, S., 1995. A limit theorem for a smooth class of semiparametric estimators. Journal of Econometrics 65, 295–332.

Pakes, A., Ostrovsky, M., Berry, S., 2004. Simple estimators for the parameters of discrete dynamic games (with entry/exit example). Working Paper. Harvard University.

Pakes, A., Pollard, D., 1985. Simulation and the asymptotics of optimization estimators. Econometrica 57, 1027–1057.

Pesendorfer, M., Schmidt-Dengler, P., 2003. Identification and estimation of dynamic games. NBER Working Paper.

Pesendorfer, M., Schmidt-Dengler, P., 2008. Asymptotic least squares estimator for dynamic games. The Review of Economic Studies 75, 901–928.

Powell, J.L., Stock, J.H., Stoker, T.M., 1989. Semiparametric estimation of index coefficients. Econometrica 57, 1403–1430.

Robinson, P.M., 1983. Nonparametric estimators for time series. Journal of Time Series Analysis 4, 185–207.

Robinson, P.M., 1988. Root-n consistent semiparametric regression. Econometrica 56, 931–954.

Rosenblatt, M., 1956. Remarks on some nonparametric estimates of a density function. Annals of Mathematical Statistics 27, 832–837.

Roussas, G., 1967. Nonparametric estimation in Markov processes. Annals of the Institute of Statistical Mathematics 21, 73–87.

Roussas, G., 1969. Nonparametric estimation of the transition distribution function of a Markov process. Annals of Mathematical Statistics 40, 1386–1400.

Rust, J., 1987. Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher. Econometrica 55, 999–1033.

Rust, J., 1994. Estimation of dynamic structural models: problems and prospects part I: discrete decision processes. In: Sims, C.A. (Ed.), Proceedings of the 6th World Congress of the Econometric Society. Cambridge University Press, pp. 119–171.

Rust, J., 1996. Numerical dynamic programming in economics. In: Amman, H., Kendrick, D., Rust, J. (Eds.), Handbook of Computational Economics, vol. 1. North-Holland, Amsterdam, pp. 619–729.

Stokey, N.L., Lucas, R.E., 1989. Recursive Methods in Economics Dynamics. Harvard University Press, Cambridge, MA.

Wand, M.P., Jones, M.C., 1994. Kernel Smoothing. In: Chapman & Hall/CRC Monographs on Statistics & Applied Probability.