Three-Step CCP Estimation of Dynamic Programming Discrete Choice Models with Large State Space

Cheng Chou
University of Leicester

Geert Ridder
University of Southern California

March 30, 2017

Abstract

The existing estimation methods of structural dynamic discrete choice models mostly require knowing or estimating the state transition distributions. When some or all state variables are continuous and/or the dimension of the vector of state variables is moderately large, the estimation of state transition distributions becomes difficult and has to rely on tight distributional assumptions. We show that state transition distributions are indeed not necessary to identify and estimate the flow utility functions in the presence of panel data and excluded variables. A state variable is called excluded variable if it does not affect the flow utility but affects the choice probability by affecting the future payoff of current choice. We propose a new estimator of flow utility functions without estimating or specifying the state transition law or solving agents' dynamic programming problems. The estimator can be applied to both infinite horizon stationary model or general dynamic discrete choice models with time varying flow utility functions and state transition law.

JEL Classification: C14, C23, C35, J24

Keywords: Semiparametric estimation, dynamic discrete choice, continuous state variables, large state space

1 Introduction

We extend the two-step Conditional Choice Probability (CCP) estimator of Hotz and Miller (1993, HM93 hereafter) to a three-step CCP estimator. We first explain HM93's two-step method to make our motivation clear. The CCP is a function of the difference between expected payoffs from different alternatives. In dynamic programming discrete choice model, the expected payoff of one alternative equals the sum of the expected flow utility function and conditional valuation function of that alternative. Both are functions of the current observed state s_t and choice a_t. The expected flow utility functions are usually parameterized. The difficulty is the *conditional valuation function*, which is the expected remaining lifetime utility from period t onwards given the current state s_t and choice a_t. The key observation in HM93 is that under certain conditions the conditional valuation functions are explicit functions of all *future* CCP, *future* expected flow utility functions and *future* state transition distributions.[1] So the first point is that the CCP in period t can be expressed as a function of all CCP, flow utility functions and state transition distributions beyond period t. Moreover, the CCP in period t of choosing alternative a is also the conditional mean of the dummy variable that equals 1 when alternative a was chosen in period t given the observed state variables. So the second point is that after parameterizing the flow utility functions and knowing all CCP and state transition distributions, we can set up moment conditions about the unknown parameters in flow utility functions using the conditional mean interpretation of the CCP.Based on these two points, the first step of HM93's two-step estimator is to estimate the CCP and state transition distributions, nonparametrically. Plugging these estimates into the moment conditions defined by the CCP formula, the second step is to estimate the unknown parameters in the flow utility functions via the generalized methods of moments (GMM) estimator.

The motivation for this paper is that when the dimension of the vector of observable state variables s_t is even moderately large, the nonparametric estimation of the state transition distributions $F(s_{t+1} \,|\, s_t)$ in the first step of HM93 becomes difficult. The presence of continuous state variables will only make the task harder. The same difficulty appears in the other estimators following HM93, including Hotz, Miller, Sanders, and Smith (1994); Aguirregabiria and Mira (2002); Pesendorfer and Schmidt-Dengler (2008); Arcidiacono and Miller (2011); Srisuma and Linton (2012); Arcidiacono and Miller (2016), whose implementation requires the state transition distributions. For large state space, we have to grid the

[1] Two main conditions for this conclusion are the flow utility functions are additive in the utility shocks, and the utility shocks are serially uncorrelated. Arcidiacono and Miller (2011) address the serially correlated unobserved heterogeneity by including finite types in the model.

state space coarsely, or impose various conditional independence restrictions and parametric specification of state transition distributions. The aim is to obtain the state transition distributions somehow.

The main point of this paper is that it is not necessary to estimate the state transition distributions to estimate the expected flow utility functions by using HM93's original result with *panel* data and *excluded variables*. A state variable is called excluded variable if it does not affect flow utility but affect future payoffs. Let $p_t(s_t)$ be the vector of CCP in period t and $u(a_t, s_t; \theta_t)$ be the expected flow utility functions known up to a finitely dimensional θ_t. HM93 shows that given observed state s_t, the expected *optimal* flow utility in period t, denoted by U_t^o, is an explicit function of CCP and the expected flow utility functions at time t. Denote such a function by

$$U_t^o(p_t(s_t), u(a_t, s_t; \theta_t)).$$

For example, if the utility shocks are independent and follow type-1 extreme value distribution, it follows from HM93 that

$$U_t^o = \sum_{a \in A} \Pr(a_t = a \mid s_t)[u(a_t = a, s_t; \theta_t) - \ln \Pr(a_t = a \mid s_t)] + 0.5776, \tag{1}$$

where A is the choice set. Under the assumptions of HM93, the conditional valuation function of choosing a_t at time t equals

$$\sum_{r=t+1}^{T_*} \beta^{r-t} \operatorname{E}\Big(U_r^o(p_r(s_r), u(a_r, s_r; \theta_r)) \Big| s_t, a_t\Big), \tag{2}$$

where β is the discount factor, and $T_* \leq \infty$ is the last decision period. HM93 needs the state transition distributions to evaluate the conditional expectations from period $t+1$ to T_*. Hotz, Miller, Sanders, and Smith (1994) needs the state transition distributions to simulate future states and choices to evaluate these conditional expectations. Our idea is that since U_t^o can be explicitly written a function of CCP and flow utility functions, e.g. eq. (1), we might directly estimate the conditional expectation terms in eq. (2) by nonparametric regressions after estimating CCP and parameterizing flow utility functions if we have *panel* data. More concretely, take eq. (1) as an example and assume s_t is a scalar. Letting $u(a_t = a, s_t; \theta_t) = s_t \theta_t(a)$, to estimate the conditional valuation function of eq. (2), it is only necessary to estimate the following terms

$$\operatorname{E}\Big(\Pr(a_r = a \mid s_r) s_r \Big| s_t, a_t\Big) \qquad \text{and} \qquad \operatorname{E}\Big(\Pr(a_r = a \mid s_r) \ln \Pr(a_r = a \mid s_r) \Big| s_t, a_t\Big) \tag{3}$$

for *each* period $r = t+1, \ldots, T_*$. Why would it help? It helps because estimating $F(s_{t+1}|s_t)$ involves the dimension of $2 \cdot \dim s_t$, while the nonparametric regressions, like eq. (3), involves the dimension of $\dim s_t + 1$. In addition, we have abundant resource to deal with the curse-of-dimensionality in nonparametric regressions.

Careful readers may now raise two questions. First, for the infinite horizon stationary dynamic programming problem, in which $T_* = \infty$, do we have to estimate infinite number of nonparametric regressions, which is infeasible given that any practical panel data have finite sampling periods? Second, for finite horizon problem ($T_* < \infty$), do we need panel data covering agents' entire decision horizon in order to estimate eq. (3) for each period $r = t+1, \dots, T_*$? The answer is "no" to both questions. For infinite horizon stationary problem, we require the panel data to cover two consecutive decision periods. For finite horizon problem, we require the panel data to cover four consecutive decision periods. The two-period requirement for infinite horizon stationary problem is easier to understand. For an infinite horizon stationary problem, both CCP and state transition distributions are not time varying. The time invariant CCP can be estimated from single period data, and two periods data are enough for estimating the state transition distributions. Once time invariant CCP and state transition distributions are known, in principle it is possible to calculate the conditional expectations like eq. (3). The four-period requirement for finite horizon problem is more subtle. In this paper, the identification of the model will be achieved by using the Exclusion Restriction proposed by Chou (2016). The Exclusion Restriction says that there are excluded variables that do not affect the immediate payoff of the discrete choice but affect future payoffs. One useful conclusion of Chou (2016) is that the value function in the last *sampling* period of panel data is nonparametrically identified when panel data cover at least four consecutive decision periods. It is useful because we can write the conditional valuation function in eq. (2) for finite horizon problem as follows,

$$\sum_{r=t+1}^{T-1} \beta^{r-t} \, \mathrm{E}\Big(U_r^o(p_r(s_r), u(a_r, s_r; \theta_r)) \,\Big|\, s_t, a_t\Big) + \beta^{T-t} \, \mathrm{E}\Big(\bar{V}_T(s_T) \,\Big|\, s_t, a_t\Big),$$

where T is the last decision period *in sample*, and $\bar{V}_T(s_T)$ is the value function in period T after integrating out the unobserved utility shocks. The nonparametric regressions involved in the first summation term have no problem. The second term is not a problem either, because we can express $\bar{V}_T(s_T)$ as a series expansion and the series coefficients are identifiable.

Our three-step extension of HM93 proceeds in the following steps. The first step is to estimate the CCP nonparametrically. The second step is to estimate the conditional valuation function by nonparametric regressions of generated dependent variables, which are the terms in the expression of $U_t^o(\widehat{p}_t(s_t), u(a_t, s_t; \theta_t))$, on the current state and choice (s_t, a_t). The third step is GMM estimation of θ_t of the flow utility functions. Although the exact moment conditions that we used for the third step is different from HM93, the essential difference between our three-step estimator and the original HM93's two-step procedure is that we replace the nonparametric estimation of state transition distributions in HM93 with the nonparametric regression with generated dependent variables. This can be complementary to the literature when the dimension of the state variables is large. The additional cost is that for finite horizon problem, we need excluded variables in order to identify the value function in the last sampling period. The excluded variable will also be needed to identify the infinite horizon stationary dynamic programming discrete choice model.[2]

Since HM93, a series of CCP estimation papers (Altuğ and Miller, 1998; Arcidiacono and Miller, 2011, 2016) have been using the concept of "finite dependence" to simplify the CCP estimator. Given "finite dependence" property, such as terminal or renewal action, the current choice will not alter the distribution of future states after a certain number of periods. Therefore, the conditional valuation function will depend only on the CCP and flow utility functions in a small number of periods ahead. This can substantially simplify the CCP estimation. Such "finite dependence" property can also be used to simplify our three-step estimation, because we are essentially also estimating the conditional valuation functions. We are not going to pursue this direction in the present paper.

To derive the asymptotic variance of the proposed estimator, we derive the general formulas for the asymptotic variance of the three-step semiparametric M-estimators with generated dependent variables for the nonparametric regressions of the second step. It turns out that unlike the three-step semiparametric estimators with generated regressors (Hahn and Ridder, 2013), the sampling error in the first-step estimation will always affect the influence function of the three-step semiparametric estimators with generated dependent variables.

The rest is organized as follows. The dynamic discrete choice model is described in section 2. Section 3 and section 4 show the three-step estimation of infinite horizon stationary model and general nonstationary model, resepctively. Section 5 is a numerical study of the proposed three-step CCP estimator. Section 6 concludes the paper. Some technical details are included in Appendix A, B, C, and D.

[2]One way identify the infinite horizon stationary model without using the excluded variable is to impose the normalization assumption that lets the flow utility function of a reference alternative being zero for all state values. This normalization will be problematical for doing counterfactual analysis as shown in our discussions after assumption 4 of our model. Parametric specification of flow utility function might be able to identify the model in some cases. For example, in the numerical example of Srisuma and Linton (2012), they estimate an infinite horizon stationary model without using the excluded variables or the normalization. They did not address the identification issue, though their Monte Carlo experiments suggest the parameters are identified.

2 Dynamic Programming Discrete Choice Model

First, we set up the model. Second, we decribe the data required for our estimator. Third, we briefly discuss the identification and conclude with the moment conditions that are the basis of our three step estimator.

2.1 The Model

We focus on the binary choice case. For period t, let $\mathit{\Omega}_t$ be the vector of state variables that could be relevant to the current and future choices or utilities apart from time itself. In each period t, an agent makes a binary choice $a_t \in A \equiv \{0, 1\}$ based on $\mathit{\Omega}_t$. The choice a_t affects both the agent's flow utility in period t and the distribution of the next period state variables $\mathit{\Omega}_{t+1}$. The vector $\mathit{\Omega}_t$ is completely observable to the agent in period t but partly observable to econometricians. Let $\mathit{\Omega}_t \equiv (x_t, z_t, \varepsilon_t)$. Econometricians only observe x_t and z_t, and let the vector ε_t denote the unobservable utility shocks. Also denote $s_t \equiv (x_t, z_t)$ the observable state variables. Assumption 1 assumes that $\mathit{\Omega}_t$ is a controlled first-order Markov process, which is standard in the literature. Assumption 2 has two points: (i) the flow utility is additive in ε_t; (ii) given x_t, z_t does not affect flow utility (the *Exclusion Restriction*). The additivity in utility shocks is usually maintained in the literature with notable exception of Kristensen, Nesheim, and de Paula (2014). We will use the Exclusion Restriction to identify the model. The exclusion restriction exists in the applied literature, e.g. Fang and Wang (2015) and Blundell, Costa Dias, Meghir, and Shaw (2016). We will use the notation by Aguirregabiria and Mira (2010) in their survey paper as much as possible. The first difference is that due to the presence of the excluded state variables z_t, we decided to use $\mathit{\Omega}_t \equiv (s_t, \varepsilon_t)$ and $s_t \equiv (x_t, z_t)$ to denote the vector of *all* state variables and the vector of *observable* state variables, respectively.[3]

Assumption 1. $\Pr(\mathit{\Omega}_{t+1} \mid \mathit{\Omega}_t, a_t, \mathit{\Omega}_{t-1}, a_{t-1}, ...) = \Pr(\mathit{\Omega}_{t+1} \mid \mathit{\Omega}_t, a_t)$.

Assumption 2. *The agent receives flow utility* $U_t(a_t, \mathit{\Omega}_t)$ *in period* t. *Letting* $\varepsilon_t \equiv (\varepsilon_t(0), \varepsilon_t(1))'$, *assume*

$$U_t(a_t, \mathit{\Omega}_t) = u_t(a_t, x_t) + \varepsilon_t(a_t).$$

The flow utility $u_t(a_t, x_t)$ in assumption 2 has a subscript "t". This is a bit uncommon in the literature. By adding "t" as a subscript, we allow the flow utility to depend on the decision period, which is "age" in many applications of labor economics, *per se*.

[3] Aguirregabiria and Mira (2010) used s_t and x_t to denote the vector of all state variables and the vector of observable state variables, respectively.

Let $T_* \leq \infty$ be the last decision period. In each period t, the agent makes a sequence of choices $\{a_t, \dots, a_{T_*}\}$ to maximize the expected discounted remaining lifetime utility,

$$\sum_{r=t}^{T_*} \beta^{r-t} \,\mathrm{E}\Big(U_r(a_r, \Omega_r) \,\Big|\, a_t, \Omega_t\Big)$$

where $0 \leq \beta < 1$ is the discount factor.[4] The agent's problem is a Markov decision process, which can be solved by dynamic programming. Let $V_t(\Omega_t)$ be the optimal value function, which solves Bellman's equation,

$$V_t(\Omega_t) = \max_{a \in A} u_t(a_t = a, x_t) + \varepsilon_t(a) + \beta \,\mathrm{E}\Big(V_{t+1}(\Omega_{t+1}) \,\Big|\, s_t, \varepsilon_t, a_t = a\Big). \tag{4}$$

Accordingly, if the agent follows the optimal decision rule, the observed choice a_t satisfies

$$a_t = \arg\max_{a \in A} u_t(a_t = a, x_t) + \beta \,\mathrm{E}\Big(V_{t+1}(\Omega_{t+1}) \,\Big|\, s_t, \varepsilon_t, a_t = a\Big) + \varepsilon_t(a). \tag{5}$$

Equation (5) is similar to the data generating process of a static binary choice model with the exception of the additional term $\beta \,\mathrm{E}(V_{t+1}(\Omega_{t+1}) | s_t, \varepsilon_t, a_t = a)$ (the conditional valuation function associated with choosing a in period t in HM93's terminology) in the "expected payoff" of alternative a. Without further restriction, the conditional valuation function is non-separable from the unobserved ε_t. To avoid dealing with non-separable models, we make the following assumption, which ensures

$$\mathrm{E}\Big(V_{t+1}(\Omega_{t+1}) \,\Big|\, s_t, \varepsilon_t, a_t\Big) = \mathrm{E}\Big(\bar{V}_{t+1}(s_{t+1}) \,\Big|\, s_t, a_t\Big), \tag{6}$$

where

$$\bar{V}_{t+1}(s_{t+1}) \equiv \mathrm{E}\Big(V_{t+1}(s_{t+1}, \varepsilon_{t+1}) \,\Big|\, s_{t+1}\Big). \tag{7}$$

Assumption 3. (i) $s_{t+1}, \varepsilon_{t+1} \perp\!\!\!\perp \varepsilon_t \,|\, s_t, a_t$, (ii) $s_{t+1} \perp\!\!\!\perp \varepsilon_{t+1} \,|\, s_t, a_t$, *and* (iii) $\varepsilon_{t+1} \perp\!\!\!\perp s_t, a_t$.

This assumption is also common in the literature. Equation (6) implies that the agent's expectation of her optimal future value depends only on the observable states variables s_t and choice a_t. This could be restrictive in some applications, e.g. it excludes fixed effect that affects flow utility or state transition or both. In Magnac and Thesmar's (2002) research about the identification of the dynamic programming discrete choice model, they show the identification of flow utility allowing for fixed effect that could affect both flow utility and state transition. One of their conditions is that the flow utility and the conditional valuation function of one reference alternative equal to zero. In our notation, taking "0" as

[4] In general, we can allow for time varying discount factor, β_t, but for the simplicity of exposition, we do not consider such an extension here.

the reference alternative, they assume $u_t(a_t = 0, x_t) = 0$ and $\mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t, a_t = 0) = 0$. Such a condition turns out to be too restrictive in some applications, as we will discuss more after assumption 4.

Using the simplification of eq. (6), the observed discrete choice a_t satisfies

$$a_t = 1\Big(\varepsilon_t(0) - \varepsilon_t(1) < v_t(a_t = 1, s_t) - v_t(a_t = 0, s_t)\Big),$$

where

$$v_t(a_t, s_t) = u_t(a_t, x_t) + \beta\,\mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,\big|\, s_t, a_t) \tag{8}$$

is the choice-specific value function minus the flow utility shock in Aguirregabiria and Mira's (2010) terminology. Let $\tilde{\varepsilon}_t \equiv \varepsilon_t(0) - \varepsilon_t(1)$ and $G(\cdot|s_t)$ be the cumulative distribution function (CDF) of $\tilde{\varepsilon}_t$ given s_t. In terms of $G(\cdot \mid s_t)$, the CCP $p_t(s_t) = \Pr(a_t = 1 \,|\, s_t)$ is

$$\begin{aligned} p_t(s_t) &= G\Big(v_t(a_t = 1, s_t) - v_t(a_t = 0, s_t) \,\Big|\, s_t\Big) \\ &= G\Big(u_t(a_t = 1, x_t) - u_t(a_t = 0, x_t) + \beta\,\mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,\big|\, s_t, a_t = 1) - \beta\,\mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,\big|\, s_t, a_t = 0) \,\Big|\, s_t\Big). \end{aligned} \tag{9}$$

Our analysis will heavily use the notation of flow utility *difference* and the conditional expectation *difference*. Without simple notation, the discussion would be cumbersome as shown by the above displayed equation. So we define the following new notation

$$\begin{aligned} \tilde{u}_t(x_t) &\equiv u_t(a_t = 1, x_t) - u_t(a_t = 0, x_t) \\ \tilde{\mathrm{E}}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t) &\equiv \mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t, a_t = 1) - \mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t, a_t = 0). \end{aligned}$$

The estimation of $u_t(a_t, x_t)$ is equivalent to the estimation of $\tilde{u}_t(x_t)$ and $u_t(a_t = 0, x_t)$.

When the CDF $G(\cdot \mid s_t)$ is unknown, even the difference $v_t(a_t = 1, s_t) - v_t(a_t = 0, s_t)$ cannot be identified in general, let alone the flow utility functions $u_t(a_t, x_t)$.[5] Suppose that the CDF $G(\cdot \mid s_t)$ is known, the absolute level of $u_t(a_t, x_t)$ cannot be identified. Take $\beta = 0$ for example, for any constant $c \in \mathbb{R}$,

$$\begin{aligned} p_t(s_t) &= G(u_t(a_t = 1, x_t) - u_t(a_t = 0, x_t) \,|\, s_t) \\ &= G\Big([u_t(a_t = 1, x_t) + c] - [u_t(a_t = 0, x_t) + c] \,\Big|\, s_t\Big). \end{aligned}$$

The following assumption is to address these concerns.

[5]With additional assumptions, the CDF $G(\cdot \mid s_t)$ is identifiable as shown by Aguirregabiria (2010, page 205). His arguments are based on the following observation. Besides the presence of the conditional valuation function $\beta\,\mathrm{E}(\bar{V}_{t+1}(s_{t+1})|s_t, a_t)$, the CCP formula eq. (9) is similar to the CCP in the binary static discrete choice model studied by Matzkin (1992), in which the CDF $G(\cdot \mid s_t)$ can be nonparametrically identified with the "special regressors" and the zero-median assumption. In this paper, we focus on the estimation when $G(\cdot \mid s_t)$ is known.

Assumption 4. (i) *Assume* $s_t \perp\!\!\!\perp \tilde{\varepsilon}_t$*, and* $\tilde{\varepsilon}_t$ *is a continuous random variable with real line support. Letting* $G(\cdot)$ *be the CDF of* $\tilde{\varepsilon}_t$*, assume* $G(\cdot)$ *is a known strictly increasing function, and* $\mathrm{E}(\varepsilon_t(0)) = 0$*.*

(ii) *For every period* t*, let* $u_t(a_t = 0, x_t = x_*) = 0$ *for some* x_**.*

The independence assumption $s_t \perp\!\!\!\perp \tilde{\varepsilon}_t$ is not particularly strong and is commonly maintained in the literature given that we have to assume the conditional CDF of $\tilde{\varepsilon}_t$ given s_t is known anyway. The normalization in assumption 4.(ii) differs from the commonly used normalization by letting one flow utility function be zero,

$$u_t(a_t = 0, x_t = x) = 0, \qquad \text{for all } t \text{ and all } x. \tag{10}$$

The normalization of eq. (10) implies that the flow utility of alternative 0 in *each* period does not vary with respect to the values of the state variable x_t, and more importantly with respect to the *distribution* of x_t. This has serious consequence. Because the future value of a current choice is the discounted expected sum of all future optimal flow utilities by following the optimal strategy, prohibiting $u_t(a_t = 0, x_t)$ from changing with respect to x_t for each period t is going to greatly restrict how the future value of the current choice changes with respect to the distribution of $(x_{t+1}, x_{t+2}, \dots)$ given x_t and a_t. When the counterfactual policy is to change the state transition distributions, it is known that the normalization of eq. (10) is not innocuous, but the normalization of assumption 4.(ii) is mostly harmless for making the counterfactual policy predictions (Norets and Tang, 2014; Kalouptsidi, Scott, and Souza-Rodrigues, 2015; Chou, 2016). With the Exclusion Restriction, Chou (2016) shows that the flow utility functions are nonparametrically identifiable without imposing the "strong" normalization of eq. (10).

2.2 Data and Structural Parameters

There are n agents in data. For each agent $i = 1, \dots, n$, we observe her decisions a_{it} and observable state variables s_{it} from decision period 1 to $T \leq T_*$. Taking the female labor force participation model as an example, the decision period is "age", and a_{it} is of course whether or not woman i is on the labor market at the age of t. In the female labor supply study (e.g. Eckstein and Wolpin, 1989; Altuğ and Miller, 1998; Blundell, Costa Dias, Meghir, and Shaw, 2016), some common state variables in s_{it} include woman i's education, accumulated assets, working experience, her family background, her partner's wage (which would zero if she is single), education and working experience, and the number of dependent children in the household. The point is s_{it} usually has a large dimension and includes both continuous (partner's wage) and discrete variables (education). These features make the direct estimation of the state transition distribution hard. It should be remarked that the first decision period 1 in sample does not need to be the initial period of her dynamic programming problem, nor does the last period T in sample correspond to the terminal decision period T_*.

The structural parameters of this model include the flow utility functions $u_t(a_t, x_t)$, discount factor β and state transition distributions $F_{t+1}(s_{t+1}|s_t, a_t)$ in each period t. The state transition distributions are identified from data. By adding time subscript in the state transition distributions, we allow the distribution vary across decision periods. This can be useful in applications because for example data show that mean earnings are hump shaped over the working lifetime (see e.g. Heckman, Lochner, and Todd, 2003; Huggett, Ventura, and Yaron, 2011; Blundell, Costa Dias, Meghir, and Shaw, 2016).

2.3 Identification from Linear Moment Conditions

The identification of the structural parameters (excepting for the state transition distributions, which are directly identified from data) is based on the following equations,

$$\varphi(p_t(s_t)) = v_t(a_t = 1, s_t) - v_t(a_t = 0, s_t) = \tilde{u}_t(x_t) + \beta\, \tilde{\mathrm{E}}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t), \tag{11a}$$

$$\psi(p_t(s_t)) = \bar{V}_t(s_t) - u_t(a_t = 0, x_t) - \beta\, \mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t, a_t = 0), \tag{11b}$$

for every period t in data. Our estimator will be based on an alternative expression of the above equations, which is eq. (14) at the end of this section. Here $\varphi(p)$ and $\psi(p)$ are constructed from $G(\cdot)$, the known strictly increasing CDF of $\tilde{\varepsilon}_t = \varepsilon_t(0) - \varepsilon_t(1)$:

$$\begin{aligned} \varphi(p) &\equiv G^{-1}(p) \\ \psi(p) &\equiv \int \max[0, G^{-1}(p) - t]\, \mathrm{d}\, G(t), \end{aligned} \tag{12}$$

for $p \in [0, 1]$. Clearly, $\varphi(p)$ is the quantile function. If we substitute p in the definition of $\psi(p)$ with the CCP $p_t(s_t)$, $\psi(p_t(s_t))$ viewed as a function of s_t is the so called McFadden's social surplus function, that is the expected life-time utility following the optimal policy minus the expected life-time utility of choosing alternative zero in period t and following the optimal policy thereafter. The fact that the social surplus function depends only on the CCP is known in the literature (see e.g. page 501–502 of HM93 and Proposition 2 of Aguirregabiria, 2010).

Equation (11a) follows from inverting the CDF $G(\cdot)$ in eq. (9) and the definition of $\varphi(p)$ in eq. (12). It is a special case of the Hotz and Miller's inversion result (HM93, Proposition

1), which establishes the invertibility for multinomial discrete choice. Though not explicitly presented in HM93, eq. (11b) can be derived using the similar arguments of HM93, page 501. It follows from integrating out ε_t from the both sides of Bellman's eq. (4). First, in terms of $v_t(a_t, s_t)$, eq. (4) is rewritten as follows,

$$V_t(s_t, \varepsilon_t) = \max_{a \in A} v_t(a_t = a, s_t) + \varepsilon_t(a).$$

The integration of ε_t for both sides of the above display gives

$$\begin{aligned}\bar{V}_t(s_t) &= \int \max\Big(v_t(a_t = 0, s_t) + \varepsilon_t(0),\ v_t(a_t = 1, s_t) + \varepsilon_t(1)\Big)\ \mathrm{d}\, F(\varepsilon_t(0), \varepsilon_t(1)) \\ &= v_t(a_t = 0, s_t) + \int \max(0,\ v_t(a_t = 1, s_t) - v_t(a_t = 0, s_t) - \tilde{\varepsilon}_t)\ \mathrm{d}\, G(\tilde{\varepsilon}_t).\end{aligned}$$

Here we used eq. (7) and $\mathrm{E}(\varepsilon_t(0)) = 0$ in assumption 4. It then follows from

$$v_t(a_t = 1, s_t) - v_t(a_t = 0, s_t) = G^{-1}(p_t(s_t))$$

that

$$\begin{aligned}\bar{V}_t(s_t) &= v_t(a_t = 0, s_t) + \int \max(0,\ G^{-1}(p_t(s_t)) - \tilde{\varepsilon})\ \mathrm{d}\, G(\tilde{\varepsilon}) \\ &= v_t(a_t = 0, s_t) + \psi(p_t(s_t)).\end{aligned}$$

The above display gives eq. (11b) by substituting $v_t(a_t = 0, s_t)$ with its definition in eq. (8). In addition, the two functional of CCP, $\varphi(p_t(s_t))$ and $\psi(p_t(s_t))$, are related. Letting a_t^o be the optimal choice at time t given s_t, we have

$$\mathrm{E}(\varepsilon_t(a_t^o) \,|\, s_t) = \psi(p_t(s_t)) - p_t(s_t)\varphi(p_t(s_t)). \tag{13}$$

This can be verified by calculating $\bar{V}_t(s_t)$ in the following way,

$$\begin{aligned}\bar{V}_t(s_t) &= \mathrm{E}(v_t(a_t^o, s_t) + \varepsilon_t(a_t^o) \,\big|\, s_t) \\ &= p_t(s_t)v_t(a_t = 1, s_t) + (1 - p_t(s_t))v_t(a_t = 0, s_t) + \mathrm{E}(\varepsilon_t(a_t^o) \,|\, s_t) \\ &= v_t(a_t = 0, s_t) + p_t(s_t)[v_t(a_t = 1, s_t) - v_t(a_t = 0, s_t)] + \mathrm{E}(\varepsilon_t(a_t^o) \,|\, s_t) \\ &= v_t(a_t = 0, s_t) + p_t(s_t)\varphi(p_t(s_t)) + \mathrm{E}(\varepsilon_t(a_t^o) \,|\, s_t).\end{aligned}$$

Comparing the above display with eq. (11b), we have eq. (13). In the rest of the paper, we work with the term $\psi(p_t(s_t)) - p_t(s_t)\varphi(p_t(s_t))$ as a whole. So we define

$$\eta(p(s)) \equiv \psi(p(s)) - p(s)\varphi(p(s)).$$

Equation (11) can be viewed as linear moment conditions about the structural parameters. We observe CCP, hence $\varphi(p_t(s_t))$ and $\psi(p_t(s_t))$, from data. Given the discount factor β, the observable $\varphi(p_t(s_t))$ and $\psi(p_t(s_t))$ are linear in the unknown flow utility and value functions. Given the flow utility and value functions, $\varphi(p_t(s_t))$ and $\psi(p_t(s_t))$ are linear in β. Chou (2016) shows that with the Exclusion Restriction and certain rank conditions, we can nonparametrically identify (i) the value function $\bar{V}_t(s_t)$; (ii) the difference between the flow utility functions $\tilde{u}_t(x_t)$; (iii) the flow utility function of alternative zero $u_t(a_t = 0, x_t)$. Below, we assume that the model is identified and focus on the estimation issue, though during the presentation of our estimator, you will see at least hints about the identification in remark 1 and remark 2.

In the rest, we transform eq. (11) to eq. (14), which turns out to be more useful for the estimation job. Lemma 1 is to get rid of the event $a_t = 0$ in the conditional expectation $\mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t, a_t = 0)$ in eq. (11b), so that the law of iterated expectation arguments, like $\mathrm{E}(\mathrm{E}(v_{t+k} \,|\, s_{t+k-1}) \,|\, s_t) = \mathrm{E}(v_{t+k} \,|\, s_t)$, can be used.

Lemma 1. *We have*

$$\mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t, a_t = 0) = \mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t) - \beta^{-1} p_t(s_t)[\varphi(p_t(s_t)) - \tilde{u}_t(x_t)]$$

Proof. It follows from the law of total probability that (we suppress s_{t+1} in $\bar{V}_{t+1}(s_{t+1})$ and s_t in $p_[(s_[)t])$

$$\begin{aligned}\mathrm{E}(\bar{V}_{t+1} \mid s_t) &= (1 - p_t)\,\mathrm{E}(\bar{V}_{t+1} \mid s_t, a_t = 0) + p_t\,\mathrm{E}(\bar{V}_{t+1} \mid s_t, a_t = 1) \\ &= \mathrm{E}(\bar{V}_{t+1} \mid s_t, a_t = 0) + p_t\,\tilde{\mathrm{E}}(\bar{V}_{t+1} \mid s_t).\end{aligned}$$

So we have

$$\begin{aligned}\mathrm{E}(\bar{V}_{t+1} \mid s_t, a_t = 0) &= \mathrm{E}(\bar{V}_{t+1} \mid s_t) - p_t\,\tilde{\mathrm{E}}(\bar{V}_{t+1} \mid s_t) \\ &= \mathrm{E}(\bar{V}_{t+1} \mid s_t) - \beta^{-1} p_t[\varphi(p_t) - \tilde{u}_t(x_t)],\end{aligned}$$

where the second line follows from eq. (11b). ■

Using lemma 1, we can rewrite eq. (11b) as follows,

$$\psi(p_t(s_t)) = \bar{V}_t(s_t) - u_t(a_t = 0, x_t) - \beta\,\mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t) + p_t(s_t)[\varphi(p_t(s_t)) - \tilde{u}_t(x_t)],$$

hence

$$\bar{V}_t(s_t) = \Big[u_t(a_t = 0, x_t) + p_t(s_t)\tilde{u}_t(x_t) + \eta(p_t(s_t))\Big] + \beta\,\mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t).$$

The term within the bracket is indeed the expected optimal flow utility in period t,

$$U_t^o(s_t) = \mathrm{E}(u_t(a_t^o, x_t) + \varepsilon_t(a_t^o) \mid s_t),$$

because

$$\begin{aligned} \mathrm{E}(u_t(a_t^o, x_t) \mid s_t) &= p_t(s_t) u_t(a_t = 1, x_t) + (1 - p_t(s_t)) u_t(a_t = 0, x_t) \\ &= u_t(a_t = 0, x_t) + p_t(s_t) \tilde{u}_t(x_t), \end{aligned}$$

and eq. (13).

In summary, we have

$$\varphi(p_t(s_t)) = \tilde{u}_t(x_t) + \beta \, \tilde{\mathrm{E}}(\bar{V}_{t+1}(s_{t+1}) \bigm| s_t) \tag{14a}$$

$$\bar{V}_t(s_t) = U_t^o(s_t) + \beta \, \mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \bigm| s_t) \tag{14b}$$

$$U_t^o(s_t) \equiv u_t(a_t = 0, x_t) + p_t(s_t) \tilde{u}_t(x_t) + \eta(p_t(s_t)), \tag{14c}$$

for every decision period t. Below, we show how to estimate $u_t(a_t, x_t)$ using these equations.

3 Estimation of Infinite Horizon Stationary Markov Decision Processes

We first show our three-step estimator for the infinite horizon stationary dynamic programming discrete choice model, because of its simple structure and importance in the literature. Since the structural parameters of an infinite horizon stationary model are all time invariant, we omit t from CCP, flow utility and value functions below.

The plan for this section is the following. First, we derive a simpler moment condition about the flow utility functions from eq. (14). This simpler moment condition, eq. (23), resembles linear regression, but it has infinite number of unknown conditional expectations in both dependent and independent variables. Second, we describe our three-step semiparametric estimation method assuming that we can estimate the infinite number of unknown conditional expectations with finite-period panel data. Third, we show how to estimate the infinite number of conditional expectations with two-period panel data by providing an estimable approximation formula of all conditional expectations and the order of approximation error. Fourth, we derive the influence function for our three-step semiparametric estimator.

3.1 Linear moment equation

Starting with eq. (14b), we have

$$\bar{V}(s_t) - \beta\,\mathrm{E}(\bar{V}(s_{t+1}) \,|\, s_t) = U^o(s_t). \tag{15}$$

By the recursive structure of eq. (15), we expect to express $\bar{V}(s_t)$ in terms of $\mathrm{E}(U^o_r(s_r) \,|\, s_t)$ with $r = t, t+1, \ldots$. Lemma 2 verifies our expectation.

Lemma 2. *If assumption 1-4 hold, we have* $\bar{V}(s_t) = \sum_{j=0}^{\infty} \beta^j\,\mathrm{E}(U^o_{t+j}(s_{t+j}) \,|\, s_t)$.

Proof. Let $F(s' \,|\, s)$ be the time invariant conditional CDF of s_{t+1} given s_t. In the proof, we consider the case that the conditional CDF $F(s' \,|\, s)$ of $s_{t+1} \,|\, s_t$ is absolutely continuous for every s, and let $f(s' \,|\, s)$ be the conditional probability density function (PDF). This does not lose generality, since one can redefine PDF with respect to other measures depending on the type of s_t.

Let $\mathcal{V}$ be a Banach space of $\bar{V}(s)$ with norm $\|\bar{V}(s)\| = \sup_{s\in\mathcal{S}} |\bar{V}(s)|$, where $\mathcal{S}$ is the domain of the value function. We can define a linear operator L such that

$$[L(\bar{V})](s) = \mathrm{E}(\bar{V}(s') \,|\, s) = \int \bar{V}(s') F(\mathrm{d}\, s' \,|\, s).$$

The left-hand-side of eq. (15) equals $[(I - \beta L)(\bar{V})](s_t)$, where I denotes identity operator.

First, we show $I - \beta L$ is invertible. Because L is a linear integral operator with the kernel function being the conditional PDF $f(s' \,|\, s)$, we know that

$$\begin{aligned} \|L\| = \sup_{\bar{V}\in\mathcal{V}, \|\bar{V}\|=1} \big\|[L(\bar{V})](s)\big\| &= \sup_{\bar{V}\in\mathcal{V}, \|\bar{V}\|=1} \left\| \int \bar{V}(s') f(s' \,|\, s)\ \mathrm{d}\, s' \right\| \\ &= \sup_{s\in\mathcal{S}} \int |f(s' \,|\, s)|\ \mathrm{d}\, s' = \sup_{s\in\mathcal{S}} \int f(s' \,|\, s)\ \mathrm{d}\, s' = 1. \end{aligned}$$

We then know that $[\beta L(\bar{V})](s_t) = \beta\,\mathrm{E}(\bar{V}(s_{t+1}) \,|\, s_t)$, as a linear operator, has norm $\beta < 1$. It then follows from the geometric series theorem for linear operators (e.g. see Helmberg, 2008, Theorem 4.23.3) that $I - \beta L$ is invertible. Hence, we conclude that $\bar{V}(s_t) = (I - \beta L)^{-1} U^o_t(s_t)$.

Second, we have $(I - \beta L)^{-1} = \sum_{j=0}^{\infty} (\beta L)^j$ from the geometric series theorem, hence $\bar{V}(s_t) = \sum_{j=0}^{\infty} \beta^j [L^j(U^o(s))](s_t)$.

Third, we show that

$$[L^j(U^o(s))](s_t) = \mathrm{E}(U^o_{t+j}(s_{t+j}) \,|\, s_t), \qquad \text{for } j = 0, 1, 2, \ldots \tag{16}$$

by induction. For $j = 0$, because $L^0 = I$ and $\mathrm{E}(U^o_t(s_t) \,|\, s_t) = U^o_t(s_t)$, the above display holds. Suppose the above display holds for $j = J$, we need to show that it holds for $j = J+1$.

To see this, we have

$$
\begin{aligned}
[L^{J+1}(U^o(s))](s_t) &= \Big[L\Big(L^J(U^o(s))\Big)\Big](s_t) \\
&= \int L^J(U^o(s))(s_{t+1}) f(s_{t+1} \,|\, s_t) \,\mathrm{d}\, s_t \\
&= \int \int U^o_{t+1+J}(s_{t+1+J}) f(s_{t+1+J} \,|\, s_{t+1}) \,\mathrm{d}\, s_{t+1+J} f(s_{t+1} \,|\, s_t) \,\mathrm{d}\, s_t. \qquad (17)
\end{aligned}
$$

The last line used $[L^J(U^o(s))](s_{t+1}) = \mathrm{E}(U^o_{t+1+J}(s_{t+1+J}) \,|\, s_{t+1})$. By the Markov property, we have $f(s_{t+1+J} \,|\, s_{t+1}) = f(s_{t+1+J} \,|\, s_{t+1}, s_t)$, hence

$$
f(s_{t+1+J} \,|\, s_{t+1}) f(s_{t+1} \,|\, s_t) = f(s_{t+1+J}, s_{t+1} \,|\, s_t).
$$

As a result,

$$
\begin{aligned}
\text{eq. (17)} &= \int \int U^o_{t+1+J}(s_{t+1+J}) f(s_{t+1+J}, s_{t+1} \,|\, s_t) \,\mathrm{d}\, s_{t+1+J} \,\mathrm{d}\, s_t \\
&= \mathrm{E}(U^o_{t+1+J}(s_{t+1+J}) \,|\, s_t).
\end{aligned}
$$

By induction, we conclude that eq. (16) is true.

Last, given eq. (16) and $\bar{V}(s_t) = \sum_{j=0}^{\infty} \beta^j [L^j(U^o(s))](s_t)$, the statement is proved. ■

Expressing the integrated value function $\bar{V}(s_t)$ as a sum of discounted expected future optimal flow utilities given the current state s_t is not an innovation relative to the literature. For discrete state space, HM93, Hotz, Miller, Sanders, and Smith (1994) and Miller (1997) all write the un-discounted conditional valuation function $\mathrm{E}(\bar{V}(s_{t+1}) \,\big|\, s_t, a_t = a)$ as

$$
\sum_{j=1}^{\infty} \beta^j \,\mathrm{E}\big(U^o(s_{t+j}) \,\big|\, s_t, a_t = a\big), \tag{18}
$$

which follows from the expression of $\bar{V}(s_{t+1})$ in our lemma 2. The existing CCP estimators in discrete state space proceeds by expressing the conditional expectations of eq. (18) as the product the state transition probability matrices and the optimal flow utilities in each state. More concretely, for discrete state space, let F denote the state transition probability matrix from s_t to s_{t+1}, which is constant across time for stationary problem. Let F_a be the conditional state transition probability matrix from s_t to s_{t+1} given $a_t = a$. And finally let $\vec{U}^o$ denote the vector the optimal flow utility in each state of the state space. HM93 and other CCP estimators estimate eq. (18) by

$$
\beta F_a \sum_{j=0}^{\infty} \beta^j F^j \vec{U}^o = \beta F_a (I - \beta F)^{-1} \vec{U}^o.
$$

For small state space, the estimation of the transition probability matrices is straightforward. For large state space, the estimation requires various restrictions, which are not always easy to form and justify. Of course, the above formula breaks when s_t contains continuous state variables. Excepting for griding and parametric specification, another remedy for continuous state variables is to use Srisuma and Linton (2012). Their estimator essentially views eq. (15) as a Fredholm integral equation of type 2 with unknown kernel $F(s' \mid s)$, which was nonparametrically estimated in their paper. Still, their method has to estimate the state transition probabilities.

The departure from the literature and the innovation is that we will estimate the conditional expectation terms in eq. (18) by nonparametric regressions. Of course, we do not need to restrict the state space to be discrete. To see why and how we estimate eq. (18) by nonparametric regressions, we first derive a linear moment condition of the flow utility functions.

We now derive the linear moment equation eq. (23). Equation (14a) for infinite horizon stationary problem reads

$$\varphi(p(s_t)) = \tilde{u}(x_t) + \beta\, \tilde{\mathrm{E}}(\bar{V}(s_{t+1}) \,|\, s_t).$$

Replacing $\bar{V}(s_{t+1})$ in the above display with its series expression in lemma 2 and applying

$$\mathrm{E}\Big(\mathrm{E}(U^o_{t+j}(s_{t+j}) \,|\, s_{t+1}) \,\Big|\, s_t\Big) = \mathrm{E}\Big(U^o_{t+j}(s_{t+j}) \,\Big|\, s_t\Big),$$

which follows from the Markov property of s_t, we have the moment condition

$$\varphi(p(s_t)) = \tilde{u}(x_t) + \sum_{j=1}^{\infty} \beta^j\, \tilde{\mathrm{E}}\Big(U^o_{t+j}(s_{t+j}) \,\Big|\, s_t\Big), \tag{19}$$

where

$$U^o_t(s_t) = u(a_t = 0, x_t) + p(s_t)\tilde{u}(x_t) + \eta(p(s_t)).$$

For notational simplicity, let

$$\tilde{u}(x_t) = x_t'\delta \qquad \text{and} \qquad u(a_t = 0, x_t) = x_t'\alpha. \tag{20}$$

The flow utility functions can also be nonparametrically estimated if we write $\tilde{u}(x_t)$ and $u(a_t = 0, x_t)$ as series expansion. We are interested in estimating $\theta \equiv (\delta', \alpha')'$.

It follows from eq. (20) that eq. (19) becomes

$$\varphi(p(s_t)) = x_t'\delta - \sum_{j=1}^{\infty} \beta^j h_{1j}(s_t) + \sum_{j=1}^{\infty} \beta^j h_{2j}(s_t)'\alpha + \sum_{j=1}^{\infty} \beta^j h_{3j}(s_t)'\delta, \tag{21}$$

where

$$h_{1j}(s_t) \equiv \tilde{\mathrm{E}}(\eta(p(s_{t+j})) \,|\, s_t), \quad h_{2j}(s_t) \equiv \tilde{\mathrm{E}}(x_{t+j} \,|\, s_t), \quad h_{3j}(s_t) \equiv \tilde{\mathrm{E}}(p(s_{t+j})x_{t+j} \,|\, s_t).$$

Suppose we have data $(a_{it}, x'_{it}, z'_{it})$ for $i = 1, \dots, n$, $t = 1, \dots, T \geq 2$ and a known discount factor β. For $i = 1, \dots, n$, letting

$$y_{it} \equiv \varphi(p(s_{it})) + \sum_{j=1}^{\infty} \beta^j h_{1j}(s_{it}), \tag{22a}$$

$$r'_{it} \equiv \Big(x'_{it} + \sum_{j=1}^{\infty} \beta^j h_{3j}(s_{it})', \sum_{j=1}^{\infty} \beta^j h_{2j}(s_{it})'\Big), \tag{22b}$$

eq. (21) is written as the following linear equation of θ,

$$y_{it} = r'_{it}\theta. \tag{23}$$

Letting $Y_t = (y_{1t}, \dots, y_{nt})'$, $Y = (Y'_1, \dots, Y'_T)'$, $R_t = (r_{1t}, \dots, r_{nt})'$, and $R = (R'_1, \dots, R'_T)'$, we have $Y = R\theta$, or

$$\theta = (R'R)^{-1}(R'Y).$$

3.2 Three-step estimator

The estimation of θ is to obtain the sample analog of R and Y. First, we approximate y_{it} and r_{it} by truncating the infinite series in eq. (21). We will justify the truncation latter, but given the presence of the discount factor $\beta < 1$, it should not be a surprise that we can truncate the series for estimation. Define

$$y_{it,J} \equiv \varphi(p(s_{it})) + \sum_{j=1}^{J} \beta^j h_{1j}(s_{it}), \tag{24a}$$

$$r'_{it,J} \equiv \Big(x'_{it} + \sum_{j=1}^{J} \beta^j h_{3j}(s_{it})', \sum_{j=1}^{J} \beta^j h_{2j}(s_{it})'\Big). \tag{24b}$$

The notation $Y_{t,J}$, Y_J, $R_{t,J}$, and R_J are then defined similarly from $y_{it,J}$ and $r_{it,J}$. It will be helpful to denote $h_k^J(s_t) \equiv (h_{k1}(s_t), \dots, h_{kJ}(s_t))'$ for $k = 1, 2, 3$ and $h^J(s_t)' \equiv (h_1^J(s_t)', h_2^J(s_t)', h_3^J(s_t)')$ for calculating the influence functions latter. Our estimator proceeds in three steps.

Step 1: Estimate the CCP $\mathrm{E}(a_{it} \,|\, s_{it})$. Let $\hat{p}(s)$ be the CCP estimator.

Step 2: Estimate $\varphi(p(s))$ and $h_{kj}(s)$ for $k = 1, 2, 3$ and $j = 1, \dots, J$. Let $\varphi(\hat{p}(s))$ be the estimators of $\varphi(p(s))$. Taking the estimation of $h_{3j}(s) = \tilde{\mathrm{E}}(p(s_{t+j})x_{t+j} \,|\, s_t)$ for example, we need to estimate

$$\mathrm{E}(p(s_{t+j})x_{t+j}\,|\,s_t, a_t = a),$$

for $a = 0, 1$. When $t + j \leq T$, these conditional expectations can be estimated by nonparametric regression of $\hat{p}(s_{i,t+j})x_{i,t+j}$ on s_{it} and a_{it}. If $t + j > T$, we can still estimate these conditional means by using the stationarity property. We will explain the estimation of h_{kj} when $t + j > T$ in detail after completing the procedure. For $k = 1, 2, 3$, let $\hat{h}_{kj}(s)$ be the estimator of $h_{kj}(s)$.

Step 3: Then $\hat{y}_{it,J}$ and $\hat{r}_{it,J}$, hence $\hat{Y}_J$ and $\hat{R}_J$, are constructed by replacing the unknown φ and h_{kj} with their respective estimates. We have the estimator

$$\hat{\theta}_J = (\hat{R}'_J \hat{R}_J)^{-1}(\hat{R}'_J \hat{Y}_J).$$

We want to emphasize that the implementation of the entire estimation procedure involves only certain number of nonparametric regressions and one ordinary linear regression. The number of nonparametric regressions depends on the dimension of x_t linearly. So we claim that the numerical implementation is not difficult, and the computation burden grows linearly in the dimension of the state variables.

Remark 1 (Role of the excluded variable z_t with linear flow utility functions). We claimed that the excluded variable z_t was required to identify the model. This point can be now better understood from the linear regression perspective using eq. (23). To identify θ, the "regressors" in r_t cannot be perfectly collinear. For simplicity, let x_t be a scalar. So there are two regressors in r_t:

$$x_t + \sum_{j=1}^{\infty} \beta^j \,\tilde{\mathrm{E}}(p(s_{t+j})x_{t+j}\,|\,x_t, z_t) \qquad \text{and} \qquad \sum_{j=1}^{\infty} \beta^j \,\tilde{\mathrm{E}}(x_{t+j}\,|\,x_t, z_t).$$

Without excluded variable z_t, these two regressors become

$$x_t + \sum_{j=1}^{\infty} \beta^j \,\tilde{\mathrm{E}}(p(s_{t+j})x_{t+j}\,|\,x_t) \qquad \text{and} \qquad \sum_{j=1}^{\infty} \beta^j \,\tilde{\mathrm{E}}(x_{t+j}\,|\,x_t).$$

The term x_t is likely to be collinear with $\tilde{\mathrm{E}}(x_{t+j}|x_t)$ depending on the conditional distribution of x_{t+j} given x_t and a_t. Due to the presence of CCP, we cannot tell how much is the correlation between $\tilde{\mathrm{E}}(p(s_{t+j})x_{t+j}\,|\,x_t)$ and $\tilde{\mathrm{E}}(x_{t+j}\,|\,x_t)$. However, because CCP is smaller than one and

$$\tilde{\mathrm{E}}(p(s_{t+j})x_{t+j}\,|\,x_t) = \mathrm{E}(p(s_{t+j})x_{t+j}\,|\,x_t, a_t = 1) - \mathrm{E}(p(s_{t+j})x_{t+j}\,|\,x_t, a_t = 0),$$

we expect that $\sum_{j=1}^{\infty} \beta^j \tilde{\mathrm{E}}(p(s_{t+j})x_{t+j} \mid x_t)$ is dominated by x_t. With linear flow utility specification and without the excluded variable, the conclusion is the two regressors are going to be highly correlated in some cases making the variance of $\hat{\theta}_J$ large. ■

Remark 2 (Role of the excluded variable z_t with nonparametric flow utility functions). If we leave the flow utility functions nonparametrically unknown, we can write the original moment equation eq. (19) as follows,

$$\begin{aligned}\varphi(p(s_t)) + \sum_{j=1}^{\infty} \tilde{\mathrm{E}}(\eta(p(s_{t+j})) \mid x_t) = {} & \Big[\tilde{u}(x_t) + \sum_{j=1}^{\infty} \beta^j \tilde{\mathrm{E}}(p(s_{t+j})\tilde{u}(x_{t+j}) \mid x_t)\Big] \\ & + \sum_{j=1}^{\infty} \beta^j \tilde{\mathrm{E}}(u(a_{t+j} = 0, x_{t+j}) \mid x_t).\end{aligned}$$

The left-hand-side of the above display is known from data. However, two terms on the right-hand-side are both unknown functions of x_t, hence neither $\tilde{u}(x_t)$ nor $u(a_t = 0, x_t)$ is identifiable. ■

3.3 Estimation of conditional expectations with two-period data

We now explain the estimation of $h_{kj}(s)$ when $t + j > T$ in more detail. In particular, we consider the most "difficult" case in which $T = 2$. Let $g(s)$ denote a generic function of s. We only need to show the approximation of $\mathrm{E}(g(s_{t+J}) \mid s_t, a_t = a)$ for all $J \geq 1$ and $a \in A$. As a summary for the rest, first, we prove that

$$\mathrm{E}(g(s_{t+J}) \mid s_t) = q^K(s_t)' \mathit{\Gamma}^J \rho + \text{error term;}$$

second, we derive the order of the error term; third, we show,

$$\mathrm{E}(g(s_{t+J}) \mid s_t, a_t) = \mathrm{E}(\mathrm{E}(g(s_{t+J}) \mid s_{t+1}) \mid s_t, a_t) = q^K(s_t)' \mathit{\Gamma}(a) \mathit{\Gamma}_{J-1} \rho + \text{error term,}$$

and derive the order of the approximation error. The notation used in the approximation formulas will be clear soon. The bottom line is the formula can be estimated from *two*-period panel data. The proofs of the results are not interesting by themselves, so we left them into the appendix.

For simplicity, let $s \in [0, 1]$ and let $g \in L_2(0, 1)$. Here $L_2(0, 1)$ denotes the set of functions $g : [0, 1] \mapsto \mathbb{R}$ such that $\int_0^1 g^2(s) \, \mathrm{d}\, s < \infty$. In this paper, we let $q_1(s), q_2(s), \ldots$ denote a sequence of generic approximating functions. Here, we let $q_1, q_2, \ldots$ form an orthonormal basis of $L_2(0, 1)$. We can always write

$$g(s) = \sum_{j=1}^{\infty} \rho_j q_j(s)$$

where $\rho_j = \int_0^1 g(s) q_j(s) \,\mathrm{d}\, s$. Denote

$$\bar{q}_k(s) \equiv \mathrm{E}(q_k(s_{t+1}) \,|\, s_t = s) \qquad \text{and} \qquad \bar{q}_k(s,a) \equiv \mathrm{E}(q_k(s_{t+1}) \,|\, s_t = s, a_t = a), \tag{25}$$

for $k = 1, 2, \dots$. Provided that $\bar{q}_k(s), \bar{q}_k(s,a) \in L_2(0,1)$ (e.g. assumption 5.(i) and 5.(iii) give two sufficient conditions), we can also write

$$\bar{q}_k(s) = \sum_{j=1}^{\infty} \gamma_{k,j} q_j(s) \qquad \text{and} \qquad \bar{q}_k(s,a) = \sum_{j=1}^{\infty} \gamma_{k,j}(a) q_j(s),$$

where $\gamma_{k,j} = \int_0^1 \bar{q}_k(s) q_j(s) \,\mathrm{d}\, s$, and $\gamma_{k,j}(a)$ is defined similarly. Let's truncate all series at K, and write

$$g(s) = q^K(s)'\rho + \nu(s), \qquad \bar{q}_k(s) = q^K(s)'\gamma_k + \omega_k(s), \qquad \bar{q}_k(s,a) = q^K(s)'\gamma_k(a) + \omega_k(s,a).$$

Here, $q^K = (q_1, \dots, q_K)'$, $\rho = (\rho_1, \dots, \rho_K)'$, $\nu(s) = \sum_{j=K+1}^{\infty} \rho_j q_j(s)$, γ_k, $\omega_k(s)$, $\gamma_k(a)$ and $\omega_k(s,a)$ are defined similarly. The following notation will be handy latter: let $\omega^K = (\omega_1, \dots, \omega_K)'$, $\omega^K(s,a) = (\omega_1(s,a), \dots, \omega_K(s,a))'$, $\mathit{\Gamma} = (\gamma_1, \dots, \gamma_K)$, and $\mathit{\Gamma}(a) = (\gamma_1(a), \dots, \gamma_K(a))$. Note both $\mathit{\Gamma}$ and $\mathit{\Gamma}(a)$ are estimable from two-period panel data using nonparametric regression. Depending on how smooth $g(s)$, $\bar{q}_k(s)$ and $\bar{q}_k(s,a)$ (for each a) are, we can bound the norm of $\nu(s)$, $\omega_k(s)$ and $\omega_k(s,a)$.

Assumption 5. *(i) Assume $\bar{q}_k(s)$ defined in eq.* (25) *is monotone is s, or the state transition density $f(s_{t+1} \,|\, s_t)$ is continuous in s_t. Either condition ensures $\bar{q}_k(s) \in L_2(0,1)$.*

(ii) A Sobolev ellipsoid is a set $B(m,c) = \left\{ b \middle| \sum_{j=1}^{\infty} a_j b_j^2 \le c^2 \right\}$, where $a^j \sim \pi j^{2m}$ as $j \to \infty$. Assume that the coefficients $(\rho_1, \rho_2, \dots)$ and $(\gamma_{k,1}, \gamma_{k,2}, \dots)$, for each $k = 1, 2, \dots$, belong to the Sobolev ellipsoid $B(m,c)$. So we have a bound for $\|\nu(s)\|_2$ and $\|\omega_k(s)\|_2$:

$$\|\nu(s)\|_2 = O(K^{-m}) \qquad \textit{and} \qquad \|\omega_k(s)\|_2 = O(K^{-m}).$$

The bound follows from the definition of $\nu(s)$ and $\omega_k(s)$ and Sobolev ellipsoid. Lemma 8.4 of Wasserman (2006) establishes this bound. Loosely speaking, the value of m depends on the order of the differentiability of the function.

(iii) For each $a \in A$, assume $\bar{q}_k(s,a)$ is monotone is s, or the state transition density $f(s_{t+1} \,|\, s_t, a_t = a)$ is continuous in s_t.

(iv) Assume that the coefficients $(\gamma_{k,1}(a), \gamma_{k,2}(a), \dots)$, for each $k = 1, 2, \dots$, belong to the Sobolev ellipsoid $B(m,c)$. So that $\|\omega_k(s,a)\|_2 = O(K^{-m})$ for each $a \in A$.

Lemma 3. *Suppose s_t is a first-order Markov process, and assumption 5.(i) holds. For any $J \geq 1$, we have*

$$\mathrm{E}(g(s_{t+J}) \mid s_t) = q^K(s_t)' \mathit{\Gamma}^J \rho + \sum_{j=1}^{J} \mathrm{E}(\omega^K(s_{t+J-j})' \mathit{\Gamma}^{j-1} \rho \mid s_t) + \mathrm{E}(\nu(s_{t+J}) \mid s_t),$$

for any t.

Proposition 1. *Suppose s_t is a first-order Markov process, and assumption 5.(i)-(ii) hold. For any $J \geq 1$ (including $J \to \infty$),*

$$\left\| \mathrm{E}(g(s_{t+J}) \mid s_t) - q^K(s_t)' \mathit{\Gamma}^J \rho \right\|_2 = O(K^{1/2-m}).$$

Proposition 2. *Suppose s_t is a first-order Markov process, and assumption 5 holds. For any $J \geq 1$ (including $J \to \infty$),*

$$\left\| \mathrm{E}(g(s_{t+J}) \mid s_t, a_t = a) - q^K(s_t)' \mathit{\Gamma}(a) \mathit{\Gamma}^{J-1} \rho \right\|_2 = O(K^{1/2-m}).$$

Given proposition 2, a reasonable estimator of $\mathrm{E}(g(s_{t+J}) | s_t, a_t = a)$ is $q^K(s_t)' \hat{\Gamma}(a) \hat{\Gamma}^{J-1} \rho$. The two matrices $\hat{\Gamma}(a)$ and $\hat{\Gamma}$ are easily obtained from the nonparametric regressions of $q_k(s_{i,t+1})$ on $s_{i,t}$ and $(s_{i,t}, a_{i,t})$, respectively, for all $k = 1, \dots, K$. The vector $\hat{\rho}$ is obtained from its definition $\rho_j = \int_0^1 g(s) q_j(s) \, \mathrm{d}\, s$.

3.4 Influence function for stationary decision process

Letting $N = nT$, our goal is to characterize the first-order asymptotic properties of $\hat{\theta}_J = (\hat{R}_J' \hat{R}_J)^{-1} (\hat{R}_J' \hat{Y}_J)$ when $n \to \infty$ and T is fixed by deriving its influence function. The derivation will mostly use the pathwise derivative approach by Newey (1994a) and Hahn and Ridder (2013).

Letting $p_*(s_t)$ be the true CCP, denote y_{it*} and r_{it*} the y_{it} and r_{it} of eq. (22) evaluated at the true CCP.The other objects with subscript "$*$" are defined in the same fashion. Define

$$\begin{aligned} \theta_* &= (R_*' R_*)^{-1} R_*' Y_* = [\mathrm{E}(r_{it*} r_{it*}')]^{-1} \mathrm{E}(r_{it*} y_{it*}), \qquad (26) \\ \theta_J &= (R_{J*}' R_{J*})^{-1} R_{J*}' Y_{J*}, \\ \theta_{J*} &= [\mathrm{E}(r_{it,J*} r_{it,J*}')]^{-1} \mathrm{E}(r_{it,J*} y_{it,J*}). \end{aligned}$$

The identity in eq. (26) follows from $y_{it*} = r_{it*}' \theta_*$. We decompose

$$\sqrt{N}(\hat{\theta}_J - \theta_*) = \sqrt{N}(\hat{\theta}_J - \theta_{J*}) + \sqrt{N}(\theta_{J*} - \theta_*),$$

and make the following assumption.

Assumption 6. (i) *There exists a constant* ζ, *such that* $\sup_s |h_{kj}(s)| < \zeta$ *for all* $k = 1, 2, 3$ *and* $j = 1, 2, \dots$.

(ii) *The smallest eigenvalue of* $\mathrm{E}(r_{it,J} r'_{it,J})$ *is bounded away from zero uniformly in* $J = 1, 2, \dots$.

In the appendix, we show that the bias term $\sqrt{N}(\theta_{J*} - \theta_*) = O_p(\beta^{J+1})$ (proposition B.2), so we can focus on the analysis of the variance term $\sqrt{N}(\hat{\theta}_J - \theta_{J*})$, which can be understood as a three-step M-estimator with generated dependent variables. Proposition B.2 justifies the truncation we made earlier. The moment equation is

$$m(x_{it}, \theta, h^J(s_{it}; p), p) = r_{it,J}(y_{it,J} - r'_{it,J}\theta).$$

Equation (24) says that $r'_{it,J} = \left(x'_{it} + \sum_{j=1}^{J} \beta^j h_{3j}(s_{it})', \sum_{j=1}^{J} \beta^j h_{2j}(s_{it})'\right)$ is a function of x_{it}, $h_2^J(s_{it})$ and $h_3^J(s_{it})$, and $y_{it,J} = \varphi(p(s_{it})) + \sum_{j=1}^{J} \beta^j h_{1j}(s_{it})$ is a function of $p(s_{it})$ and $h_1^J(s_{it})$. The vector $h^J(s_t)' \equiv (h_1^J(s_t)', h_2^J(s_t)', h_3^J(s_t)')$ itself is a function of p. So we write $h^J(s_{it}; p)$. The estimator $\hat{\theta}_J$ solves the moment equation,

$$\sum_{i,t=1}^{n,T} m(x_{it}, \theta, \hat{h}^J(s_{it}; \hat{p}), \hat{p})/N = 0,$$

and we have $\mathrm{E}\big(m(x_{it}, \theta_{J*}, h_*^J(s_{it}; p_*), p_*)\big) = 0$.

It is similar to Hahn and Ridder (2013) that using the pathwise derivative approach in Newey (1994a), the influence function associated with $\sqrt{N}(\hat{\theta}_J - \theta_{J*})$ can be expressed as a sum of the following terms: (i) the leading term $\sqrt{N}(\theta_J - \theta_{J*})$, (ii) a term that adjusts for the sampling variation in the nonparametric regressions for estimating $h_{kj}(s_t)$, when the CCP is known,

$$\frac{1}{\sqrt{N}} \sum_{i,t=1}^{n,T} \left(\sum_{i,t=1}^{n,T} r_{it,J}\Big(x_{it}, \hat{h}^J(s_{it}; p_*)\Big) r_{it,J}\Big(x_{it}, \hat{h}^J(s_{it}; p_*)\Big)' \right)^{-1} r_{it,J}\Big(x_{it}, \hat{h}^J(s_{it}; p_*)\Big) y_{it,J}\Big(p_*, \hat{h}^J(s_{it}; p_*)\Big) - \sqrt{N}\theta_J,$$

and (iii) an adjustment for the CCP estimation $\hat{p}$ in the first step, i.e.,

$$\frac{1}{\sqrt{N}} \sum_{i,t=1}^{n,T} \left(\sum_{i,t=1}^{n,T} r_{it,J}\Big(x_{it}, h_*^J(s_{it}; \hat{p})\Big) r_{it,J}\Big(x_{it}, h_*^J(s_{it}; \hat{p})\Big)' \right)^{-1} r_{it,J}\Big(x_{it}, h_*^J(s_{it}; \hat{p})\Big) y_{it,J}\Big(\hat{p}, h_*^J(s_{it}; \hat{p})\Big) - \sqrt{N}\theta_J,$$

Proposition B.1 of the Appendix shows that the leading term $\sqrt{N}(\theta_J - \theta_{J*})$ is $O_p(\beta^{J+1})$. We then just derive the other two terms.

In appendix D, we show the general results of the asymptotic variance of semiparametric M-estimators with generated dependent variables, which cover the present problem as a special case. The results here are derived using the general formulas in appendix D.

We consider the second adjustment term first, which can be analyzed by the pathwise derivative approach in Newey (1994a, page 1360-1361) or Hahn and Ridder (2013, page 327). The unknown functions in $h^J(s_t;\hat{p})$ in the second step nonparametric regressions are the difference between partial means in Newey (1994b)'s terminology. Letting

$$\kappa_{it} \equiv \frac{1(a_{it}=1)}{p_*(s_{it})} - \frac{1-1(a_{it}=1)}{1-p_*(s_{it})},$$

it can be shown that the second adjustment term equals

$$\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i,t,j=1}^{n,T,J}\frac{\partial m_*}{\partial y_{it}}\beta^j\Big[\kappa_{it}\eta(p_*(s_{i,t+j}))-h_{1j*}(s_{it})\Big] && (27)\\
&+\frac{1}{\sqrt{N}}\sum_{i,t,j=1}^{n,T,J}\frac{\partial m_*}{\partial r_{it}}\beta^j\begin{pmatrix}0\\1\end{pmatrix}\otimes[\kappa_{it}x_{i,t+j}-h_{2j*}(s_{it})] && (28)\\
&+\frac{1}{\sqrt{N}}\sum_{i,t,j=1}^{n,T,J}\frac{\partial m_*}{\partial r_{it}}\beta^j\begin{pmatrix}1\\0\end{pmatrix}\otimes[\kappa_{it}p_*(s_{i,t+j})x_{i,t+j}-h_{3j*}(s_{it})], && (29)\\
&+o_p(1)
\end{aligned}$$

where $\partial m_*/\partial y_{it}$ and $\partial m_*/\partial r_{it}$ denote the derivatives of the moment equation $m(x_{it},\theta_{J*},h_*^J(s_{it};p_*),p_*)$ with respect to $y_{it,J*}$ and $r_{it,J*}$, respectively. We have

$$\frac{\partial m_*}{\partial y_{it}} = r_{it,J*} \qquad \text{and} \qquad \frac{\partial m_*}{\partial r_{it}} = (y_{it,J*}-r'_{it,J*}\theta_{J*})I - r_{it,J*}\theta_{J*}.$$

Here I is a $2d_x \times 2d_x$ identity matrix, and d_x is the dimension of x_t. Equation (27), (28) and (29) are the adjustment terms for sampling error in $\hat{h}_1^J(s_{it};p_*)$, $\hat{h}_2^J(s_{it};p_*)$ and $\hat{h}_3^J(s_{it};p_*)$, respectively.

The third adjustment term accounts for the sampling variation in estimating the CCP. Unlike the adjustment for the generated regressors studied by Hahn and Ridder (2013), the adjustment for the generated dependent variables can be obtained by simply linearizing the moment equation with respect to the first step estimator. It can be shown using the formula in appendix D that the third adjustment term equals

$$\frac{1}{\sqrt{N}}\sum_{i,t=1}^{n,T}\frac{\partial m_*}{\partial y_{it}}\varphi'(p_*(s_{it}))(a_{it}-p_*(s_{it})) \tag{30}$$

$$+\frac{1}{\sqrt{N}}\sum_{i,t,j=1}^{n,T,J}\frac{\partial m_*}{\partial y_{it}}\beta^j\kappa_{it}\frac{\partial\eta(p_*(s_{i,t+j}))}{\partial p}(a_{it}-p_*(s_{it})) \tag{31}$$

$$+\frac{1}{\sqrt{N}}\sum_{i,t,j=1}^{n,T,J}\frac{\partial m_*}{\partial r_{it}}\beta^j\begin{pmatrix}1\\0\end{pmatrix}\otimes\kappa_{it}x_{i,t+j}(a_{it}-p_*(s_{it})) \tag{32}$$

$$+o_p(1), \tag{33}$$

where

$$\frac{\partial\eta(p_*(s_{i,t+j}))}{\partial p}=\varphi(p_*(s_{i,t+j}))+p_*(s_{i,t+j})\varphi'(p_*(s_{i,t+j}))-\psi'(p_*(s_{i,t+j})).$$

Equation (30), (31) and (32) are the adjustment terms for sampling error in $\varphi(\hat{p}(s_{it})$, $h_1^J(s_t;\hat{p})$ and $h_3^J(s_t;\hat{p})$.

Remark 3. From the influence functions, it easy to see that the discount factor affects the asymptotic variance of the three-step estimator. The closer to 1 is the discount factor β, the larger variance is the three-step estimator.

4 Estimation of General Markov Decision Processes

By "general" Markov decision processes, we allow for time varying flow utility functions and state transition distributions, and finite decision horizon.

The plan for this section is the following. First, we derive a simpler moment condition about the flow utility functions from eq. (14). This simpler moment condition, eq. (35), resembles a partially linear model, in which the unknown function is the integrated value function in the last sampling period $\bar{V}_T(s_T)$. Using the series approximation of $\bar{V}_T(s_T)$ and the partitioned linear regression arguments, we derive an estimable formula for the vector of parameters specifying the flow utility functions. Such a formula is then estimated by a three-step semiparametric estimator, which is our second part. Third, we derive the influence function for our three-step semiparametric estimator. The arguments there have two parts. We first show that the series approximation error of $\bar{V}_T(s_T)$ does not affect the influence function of our three-step estimator under certain conditions. Then, the rest arguments are similar to the ones we used in deriving the influence function for the infinite horizon stationary model.

4.1 Linear moment condition

Starting from eq. (14b), we have

$$\bar{V}_t(s_t) = U_t^o(s_t) + \beta\, \mathrm{E}(\bar{V}_{t+1}(s_{t+1}) \,|\, s_t).$$

Iteratively substituting $\bar{V}_{t+1}(s_{t+1})$ with the above expression and using the Markov property, we have

$$\bar{V}_t(s_t) = \sum_{j=0}^{T-t-1} \beta^j\, \mathrm{E}(U_{t+j}^o(s_{t+j}) \,|\, s_t) + \beta^{T-t}\, \mathrm{E}(\bar{V}_T(s_T) \,|\, s_t).$$

Applying this formula to $\bar{V}_{t+1}(s_{t+1})$, eq. (14a) becomes

$$\varphi(p_t(s_t)) = \tilde{u}_t(x_t) + \sum_{j=1}^{T-t-1} \beta^j\, \tilde{\mathrm{E}}(U_{t+j}^o(s_{t+j}) \,|\, s_t) + \beta^{T-t}\, \mathrm{E}(\bar{V}_T(s_T) \,|\, s_t),$$

where

$$U_t^o(s_t) = u_t(a_t = 0, x_t) + p_t(s_t)\tilde{u}_t(x_t) + \eta(p_t(s_t)).$$

Similar to the infinite horizon stationary model, let

$$u_t(a_t = 0, x_t) = x_t'\alpha_t \qquad \text{and} \qquad \tilde{u}_t(x_t) = x_t'\delta_t.$$

We have the following linear equations,

$$y_{T-1} = x_{T-1}'\delta_{T-1} + \beta\, \tilde{\mathrm{E}}(\bar{V}_T(s_T) \,|\, s_{T-1}) \tag{34a}$$

$$y_t = x_t'\delta_t + \sum_{j=1}^{T-t-1} \beta^j h_{2,t,j}(s_t)'\alpha_{t+j} + \sum_{j=1}^{T-t-1} \beta^j h_{3,t,j}(s_t)'\delta_{t+j} + \beta^{T-t}\, \tilde{\mathrm{E}}(\bar{V}_T(s_T) \,|\, s_t), \tag{34b}$$

for $t = 1, \ldots, T-2$, where

$$y_t \equiv \varphi(p_t(s_t)) + \sum_{j=1}^{T-t-1} \beta^j h_{1,t,j}(s_t) \qquad \text{and} \qquad y_{T-1} \equiv \varphi(p_{T-1}(s_{T-1})),$$

with

$$h_{1,t,j}(s_t) \equiv \tilde{\mathrm{E}}(\eta(p_{t+j}(s_{t+j}))\big|s_t), \quad h_{2,t,j}(s_t) \equiv \tilde{\mathrm{E}}(x_{t+j}|s_t), \quad h_{3,t,j}(s_t) \equiv \tilde{\mathrm{E}}(p_{t+j}(s_{t+j})x_{t+j}|s_t).$$

Equation (34) provides a set of linear moment conditions about $\theta = (\delta_1', \ldots, \delta_{T-1}', \alpha_2', \ldots, \alpha_{T-1}')'$. With some definition of matrices, we have a concise version of eq. (34). Suppose we have panel data (a_{it}, s_{it}') with $i = 1, \ldots, n$ and $t = 1, \ldots, T$. For each i, define a $(T-1) \times \dim(\theta)$ matrix

$$R_i = (R_{i,\delta}, R_{i,\alpha}),$$

where

$$R_{i,\delta} \equiv \begin{bmatrix} x'_{i,1} & \beta h_{3,1,1}(s_{i,1})' & \beta^2 h_{3,1,2}(s_{i,1})' & \cdots & \beta^{T-2} h_{3,1,T-2}(s_{i,1})' \\ & x'_{i,2} & \beta h_{3,2,1}(s_{i,2})' & \cdots & \beta^{T-3} h_{3,2,T-3}(s_{i,2})' \\ & & x'_{i,3} & \cdots & \beta^{T-4} h_{3,3,T-4}(s_{i,3})' \\ & & & \ddots & \vdots \\ 0 & & & & x'_{i,T-1} \end{bmatrix},$$

$$R_{i,\alpha} \equiv \left[\begin{array}{ccccc} \beta h_{2,1,1}(s_{i,1})' & \beta^2 h_{2,1,2}(s_{i,1})' & \beta^3 h_{2,1,3}(s_{i,1})' & \cdots & \beta^{T-2} h_{2,1,T-2}(s_{i,1})' \\ & \beta h_{2,2,1}(s_{i,2})' & \beta^2 h_{2,2,2}(s_{i,2})' & \cdots & \beta^{T-3} h_{2,2,T-3}(s_{i,2})' \\ & & \beta h_{2,3,1}(s_{i,3})' & \cdots & \beta^{T-4} h_{2,3,T-4}(s_{i,3})' \\ & & & \ddots & \vdots \\ 0 & & & & \beta h_{2,T-2,1}(s_{i,T-2})' \\ \hline & & 0 & & \end{array}\right].$$

Let

$$H_i = \left(\beta^{T-1}\,\tilde{\mathrm{E}}(\bar{V}_T(s_T)\,|\,s_{i1}), \dots, \beta\,\tilde{\mathrm{E}}(\bar{V}_T(s_T)\,|\,s_{i,T-1})\right)'$$

For each i, let $Y_i = (y_{i1}, \dots, y_{i,T-1})'$ and eq. (34) becomes

$$Y_i = R_i\theta + H_i.$$

Stacking Y_1, ..., Y_n, we have $Y = (Y_1', \dots, Y_n')$, which is a $N = n \cdot (T-1)$ dimensional vector. Define R and H similarly. We have

$$Y = R\theta + H. \tag{35}$$

Because H involves the unknown $\bar{V}_T(s_T)$, eq. (35) is similar to the partially linear model (Donald and Newey, 1994). It should be remarked that in discrete state space, Chou (2016) shows that given the exclusion restriction the integrated value function $\bar{V}_T(s_T)$ is identifiable up to a constant when $T \geq 4$.

4.2 Three-step estimator

Before listing our three-step recipe, we need to deal with the unknown function $\bar{V}_T(s_T)$ in H by series approximation. Let

$$q^K(s) = (q_1(s), \dots, q_K(s))'$$

be a vector of approximating functions, such as power series or splines. Let

$$Q_{i,K} = \begin{bmatrix} \beta^{T-1}\,\tilde{\mathrm{E}}(q^K(s_T)' \,|\, s_{i,1}) \\ \vdots \\ \beta\,\tilde{\mathrm{E}}(q^K(s_T)' \,|\, s_{i,T-1}) \end{bmatrix} \qquad \text{and} \qquad Q'_K = (Q'_{1K}, \dots, Q'_{nK}).$$

The estimator of θ are coefficients of R_i in the linear regression of Y_i on R_i and $Q_{i,K}$. By the usual partioned linear regression arguments, the estimator may be written as follows,

$$\theta_K = (R'M_K R)^{-1} R'M_K Y,$$

with $M_K = I - Q_K(Q'_K Q_K)^{-1}Q'_K$. The invertibility of $R'M_K R$ will be discussed latter. In the next subsection, we show the difference between θ_K and the true value of θ is asymptotically negligible. So our estimator of θ is the sample analog of θ_K formed in the following three steps.

Step 1: Estimate the CCP $\mathrm{E}(a_{it} \,|\, s_{it})$ and $\beta^{T-t}\,\tilde{\mathrm{E}}(q^K(s_T) \,|\, s_t)$ for $t = 1, \dots, T-1$. Let $\hat{p}_t(s_t)$ be the CCP estimator. Let $\hat{Q}_K$ be the estimate of Q, and let $\hat{M}_K$ be the estimate of M_K from replacing Q_K with $\hat{Q}_K$.

Step 2: Let $\varphi(\hat{p}_t(s_t))$ be the estimator of $\varphi(p_t(s_t))$. Let $\hat{h}$ be the estimator of h from nonparametric regressions with generated dependent variables.

Step 3: Then $\hat{Y}$ and $\hat{R}$ are constructed by replacing the unknown φ and h with their respective estimates. We have the estimator

$$\hat{\theta}_K = (\hat{R}'\hat{M}_K\hat{R})^{-1}\hat{R}'\hat{M}_K\hat{Y}.$$

4.3 Influence function for general decision process

Let $N = n \cdot (T-1)$. Letting θ_* be the true vector of parameters specifying the flow utility function, we can decompose

$$\sqrt{N}(\hat{\theta}_K - \theta_*) = \sqrt{N}(\hat{\theta}_K - \theta_K) + \sqrt{N}(\theta_K - \theta_*).$$

We can focus on $\sqrt{N}(\hat{\theta}_K - \theta_K)$ (variance term) if $\|\sqrt{N}(\theta_K - \theta_*)\|$ (bias term) is $o_p(1)$. We start by showing the conditions under which $\|\sqrt{N}(\theta_K - \theta_*)\| = o_p(1)$. Then the properties of $\sqrt{N}(\hat{\theta}_K - \theta_K)$ can be similarly analyzed using the influence function formula for the three-step semiparametric estimators with generated dependent variables. So the details will be omitted.

Proposition 3. *Suppose (i) The smallest eigenvalue of* $\mathrm{E}(N^{-1}R'M_KR)^{-1}$ *is bounded away from zero uniformly in* K*; (ii) The series approximation error of* $\bar{V}_T(s_T)$ *is of the order* $O(K^{-\xi})$*. More explicitly, there exists* $\xi > 0$*, such that* $\sup_{s\in\mathcal{S}}\left|\bar{V}_T(s_T) - \sum_{j=1}^{K}\rho_j q_j(s)\right| \le O(K^{-\xi})$*. Then we have* $\|\theta_K - \theta_*\| = O_p(K^{-\xi})$*. So if we impose undersmoothing conditions such that* $K^{-\xi} = o(1/\sqrt{n})$*, we can let the bias term* $\|\sqrt{N}(\theta_K - \theta_*)\| = o_p(1)$*.*

Proof. See appendix C. ■

We derive the influence function of $\hat{\theta}_K$ here. The derivation is similar to the stationary Markov decision process case. First, the three-step estimator is viewed as a M-estimator with the following moment equation,

$$m(x_i, \theta, h(s_i; p), p) = \tilde{R}_i'(\tilde{Y}_i - \tilde{R}_i\theta),$$

where $x_i = (x_{i1}', \ldots, x_{i,T-1}')'$, $s_i = (s_{i1}', \ldots, s_{iT}')'$, and $p(s_i) = (p_1(s_1), \ldots, p_{T-1}(s_{T-1}))'$ is the vector of CCP from period 1 to $T-1$. Here $\tilde{R}_i = M_{i,K}R$ and $\tilde{Y}_i = M_{i,K}Y$, with $M_K' = (M_{1,K}', \ldots, M_{n,K}')$. Each $M_{i,K}$ is an $(T-1)\times N$ matrix. We write $h(s_i; p)$ to emphasize the dependence of h on the CCP.The three-step estimator solves

$$\sum_{i=1}^{n} m(x_i, \theta, \hat{h}(s_i; \hat{p}), \hat{p})/n = 0.$$

We have $\mathrm{E}(m(x_i, \theta_K, h_*(s_i; p_*), p_*)) = 0$.

Using the pathwise derivative approach in Newey (1994a), we again can write $\sqrt{n}(\hat{\theta}_K - \theta_K)$ as the sum of the three components:

$$\frac{1}{\sqrt{n}}\Big(\sum_{i=1}^{n}\tilde{R}_i(x_i, h_*(s_i;p_*))'\tilde{R}_i(x_i, h_*(s_i;p_*))\Big)^{-1}\tilde{R}_i(x_i, h_*(s_i;p_*))'\tilde{Y}_i(h_*(s_i;p_*)) - \sqrt{n}\theta_K \quad \text{(i)}$$

$$\frac{1}{\sqrt{n}}\Big(\sum_{i=1}^{n}\tilde{R}_i(x_i, \hat{h}(s_i;p_*))'\tilde{R}_i(x_i, \hat{h}(s_i;p_*))\Big)^{-1}\tilde{R}_i(x_i, \hat{h}(s_i;p_*))'\tilde{Y}_i(\hat{h}(s_i;p_*)) - \sqrt{n}\theta_K \quad \text{(ii)}$$

$$\frac{1}{\sqrt{n}}\Big(\sum_{i=1}^{n}\tilde{R}_i(x_i, h_*(s_i;\hat{p}))'\tilde{R}_i(x_i, h_*(s_i;\hat{p}))\Big)^{-1}\tilde{R}_i(x_i, h_*(s_i;\hat{p}))'\tilde{Y}_i(h_*(s_i;\hat{p})) - \sqrt{n}\theta_K \quad \text{(iii)}$$

By the definition of θ_K, term (i) equals zero. We just need to calculate the other two terms.

Using the general results of the asymptotic variance of the three-step estimators with generated dependent variables in appendix D, we have the following results. Let

$$\begin{aligned}
\kappa_{it} &= \frac{1(a_{it}=1)}{p_{t*}(s_{it})} - \frac{1-1(a_{it}=1)}{1-p_{t*}(s_{it})}, \\
a_{i,t,j} &= \beta^j\left\{\kappa_{it}\eta(p_{t+j,*}(s_{i,t+j})) - h_{1,t,j*}(s_{it})\right\}, \\
b_{i,t,j} &= \beta^j\left[\kappa_{it}x_{i,t+j} - h_{2,t,j*}(s_{i,t})\right], \\
c_{i,t,j} &= \beta^j\left[\kappa_{it}p_{t+j,*}(s_{i,t+j})x_{i,t+j} - h_{3,t,j*}(s_{it})\right],
\end{aligned}$$

and define two block upper triangular matrices,

$$\begin{bmatrix} b'_{i,1,1} & b'_{i,1,2} & b'_{i,1,3} & \cdots & b'_{i,1,T-2} \\ & b'_{i,2,1} & b'_{i,2,2} & \cdots & b'_{i,2,T-3} \\ & & \ddots & \ddots & \vdots \\ 0 & & & \ddots & b'_{i,T-2,1} \end{bmatrix} \tag{36}$$

$$\begin{bmatrix} c'_{i,1,1} & c'_{i,1,2} & c'_{i,1,3} & \cdots & c'_{i,1,T-2} \\ & c'_{i,2,1} & c'_{i,2,2} & \cdots & c'_{i,2,T-3} \\ & & \ddots & \ddots & \vdots \\ 0 & & & \ddots & c'_{i,T-2,1} \end{bmatrix}. \tag{37}$$

Let

$$\begin{aligned} A_i &= \left(\sum_{j=1}^{T-2} a_{i,1,j}, \sum_{j=1}^{T-3} a_{i,2,j}, \cdots, \sum_{j=1}^{1} a_{i,T-2,j}, 0\right)', \\ B_i &= \left[\begin{array}{c|c} 0_{(T-1)\times d_\delta} & \begin{array}{c} \text{eq. (36)} \\ \hline 0_{1\times d_\alpha} \end{array} \end{array}\right], \\ C_i &= \left[\begin{array}{c|c|c} 0_{(T-2)\times 1} & \text{eq. (37)} & \\ \hline 0 & 0_{1\times(T-2)} & 0_{(T-1)\times d_\alpha} \end{array}\right]. \end{aligned}$$

Denote

$$\tilde{R}_{i*} = \tilde{R}_i(x_i, h_*(s_i; p_*))$$

Term (ii) equals the following,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{R}'_{i*}[A_i - (B_i + C_i)\theta_K] - \sqrt{n}\theta_K + o_p(1).$$

Let

$$E_i = \begin{bmatrix} \varphi'(p_{1*}(s_{i1}))(a_{i1} - p_{1*}(s_{i1})) \\ \vdots \\ \varphi'(p_{T-1*}(s_{i,T-1}))(a_{i,T-1} - p_{T-1*}(s_{i,T-1})) \end{bmatrix},$$

$$F_i = \begin{bmatrix} \sum_{j=1}^{T-2} \beta^j \frac{\partial \eta(p_{1+j}(s_{i,1+j}))}{\partial p} \kappa_{i1}(a_{i,1+j} - p_{1+j,*}(s_{i,1+j})) \\ \vdots \\ \beta \frac{\partial \eta(p_{T-1}(s_{i,T-1}))}{\partial p} \kappa_{i,T-1}(a_{i,T-1} - p_{T-1,*}(s_{i,T-1})) \\ 0 \end{bmatrix}.$$

Letting

$$g_{i,t,j} = \beta^j \Big[x_{i,t+j} \kappa_{it} \Big(a_{i,t+j} - p_{t+j,*}(s_{i,t+j})\Big)\Big],$$

and define a block triangular matrix,

$$\begin{bmatrix} g'_{i,1,1} & g'_{i,1,2} & g'_{i,1,3} & \cdots & g'_{i,1,T-2} \\ & g'_{i,2,1} & g'_{i,2,2} & \cdots & g'_{i,2,T-3} \\ & & \ddots & \ddots & \vdots \\ 0 & & & \ddots & g'_{i,T-2,1} \end{bmatrix} \tag{38}$$

and let

$$G_i = \left[\begin{array}{c|c|c} 0_{(T-1)\times 1} & \begin{array}{c} \text{eq. (38)} \\ \hline 0_{1\times(T-2)} \end{array} & 0_{(T-1)\times d_\alpha} \end{array} \right].$$

Term (iii) equals the following,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \tilde{R}'_{i*} M_{i,K} \begin{bmatrix} E_1 + F_1 - G_1\theta_K \\ \vdots \\ E_n + F_n - G_n\theta_K \end{bmatrix} - \sqrt{n}\theta_K + o_p(1).$$

5 Monte Carlo Studies

5.1 Infinite horizon stationary decision process

In this numerical example, both x_t and z_t are scalars. Letting the flow utility functions have the quadratic form,

$$\tilde{u}(x_t) = \delta_1 + \delta_2 x_t + \delta_3 x_t^2 \qquad \text{and} \qquad u(a_t = 0, x_t) = \alpha_1 + \alpha_2 x_t + \alpha_3 x_t^2,$$

and $u_t(a_t = 1, x_t) = \tilde{u}(x_t) + u(a_t = 0, x_t)$. In simulation, we let

$$u(a_t = 0, x_t) = \tilde{u}(x_t) = .5 + x_t - .5x_t^2.$$

The support for x_t is $[0, 3]$. The excluded variable z_t is time invariant (hence drop the time subscript t of z_t in the sequel). Conditional on (x_t, z), x_{t+1} is generated from

$$x_{t+1} = \min\Big(3,\ d_t\cdot(1+\xi_{t+1})x_t + (1-d_t)\cdot 2\xi_{t+1}\Big) \quad \text{with} \quad \xi_{t+1}|z \sim \text{Beta}\Big(\text{shape}_1(z), \text{shape}_2(z)\Big).$$

The shape parameters the beta distribution are determined in the following way so that $\mathrm{E}(\xi_{t+1} \,|\, z) = z$ and $\mathrm{Var}(\xi_{t+1}) = 1/20$:

$$\text{shape}_1(z) = 20z^2(1-z) - z \qquad \text{and} \qquad \text{shape}_2(z) = \frac{\text{shape}_1(z)\cdot(1-z)}{z}.$$

To ensure $\text{shape}_1(z), \text{shape}_2(z) > 0$, we let z follow a uniform distribution with support $[.053, .947]$. We set the discount factor $\beta = .8$. The utility shocks $\varepsilon_t(0), \varepsilon_t(1)$ follow in-

Table 1: Estimation of Infinite Horizon Stationary Markov Decision Process

Flow Utility Functions					Panel Data	
$\delta_1 = 0.5$	$\delta_2 = 1$	$\delta_3 = -0.5$	$\alpha_2 = 1$	$\alpha_3 = -0.5$	n	T
.525 (.065)	.940 (.112)	$-.501$ (.032)	.930 (.104)	$-.456$ (.027)	2e4	2
.544 (.101)	.903 (.169)	$-.504$ (.047)	.916 (.169)	$-.436$ (.045)	1e4	2
.621 (.182)	.771 (.294)	$-.500$ (.079)	.905 (.296)	$-.389$ (.084)	5e3	2

dependent type-I extreme value distribution. Given these structural parameters, we can determine the true CCP.[6]

In the numerical examples, we simulate *two* periods ($T = 2$) dynamic discrete choices by n number of agents. The first period states $\{x_{i1}, z_i : i = 1, ..., n\}$ were drawn uniformly from their support. We estimate the model when $n = 5,000, 10,000$ and $20,000$. Assume the discount factor is known in the estimation. Table 1 reports the performance of our estimator based on $1,000$ replications.[7] The performance is satisfying. There is bias due to the smoothing in the nonparametric ingredient of our three-step semiparametric estimator. As the cross-section sample size increases, the bias decreases. There is no information about α_1, because the intercept of $u(a_t = 0, x_t)$ is not identifiable.

5.2 General Markov decision process

Here, the data are still generated from the above stationary model, however, in the estimation, we do not assume the stationarity. Instead, we let

$$\tilde{u}_t(x_t) = \delta_{1,t} + \delta_{2,t}x_t + \delta_{3,t}x_t^2 \qquad \text{and} \qquad u_t(a_t = 0, x_t) = \alpha_{1,t} + \alpha_{2,t}x_t + \alpha_{3,t}x_t^2,$$

and estimate δ_t and α_t for each sampling period.

In the numerical examples, we simulate *six* periods ($T = 6$) dynamic discrete choices by n number of agents. As we did for the stationary model, we draw the first period states $\{x_{i1}, z_i : i = 1, ..., n\}$ uniformly from their support and estimate the model when $n = 5,000, 10,000$ and $20,000$. Assume the discount factor is known in the estimation. Table 2 reports the performance of our estimator based on $1,000$ replications. The performance

[6] Readers, who are interested in the algorithm of determining CCP from the structural parameters, can find the details in the documentation of our codes.

[7] The Monte Carlo simulation studies used the ALICE High Performance Computing Facility at the University of Leicester.

Table 2: Estimation of General Markov Decision Process

Flow Utility Fun.	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	n
$\delta_1=0.5$	.571 (.101)	.532 (.066)	.520 (.061)	.510 (.054)	.523 (.060)	
$\delta_2=1$	.893 (.179)	.949 (.135)	.975 (.132)	.992 (.119)	.953 (.134)	2e4
$\delta_3=-0.5$	−.470 (.046)	−.486 (.044)	−.499 (.043)	−.507 (.038)	−.498 (.044)	
$\alpha_2=1$		1.078 (.114)	1.058 (.112)	1.017 (.121)	.993 (.119)	
$\alpha_3=-0.5$		−.515 (.041)	−.510 (.037)	−.496 (.038)	−.490 (.034)	
$\delta_1=0.5$	.582 (.141)	.545 (.091)	.525 (.086)	.507 (.080)	.538 (.080)	
$\delta_2=1$	.882 (.246)	.929 (.190)	.969 (.188)	1.001 (.175)	.926 (.179)	1e4
$\delta_3=-0.5$	−.467 (.062)	−.484 (.063)	−.502 (.062)	−.514 (.057)	−.500 (.058)	
$\alpha_2=1$		1.092 (.161)	1.071 (.167)	1.019 (.177)	.978 (.172)	
$\alpha_3=-0.5$		−.520 (.058)	−.508 (.053)	−.492 (.053)	−.483 (.051)	
$\delta_1=0.5$	.610 (.208)	.567 (.136)	.539 (.116)	.523 (.109)	.559 (.111)	
$\delta_2=1$	.846 (.366)	.892 (.276)	.942 (.256)	.967 (.241)	.874 (.240)	5e3
$\delta_3=-0.5$	−.466 (.089)	−.483 (.090)	−.501 (.086)	−.514 (.082)	−.504 (.082)	
$\alpha_2=1$		1.105 (.231)	1.075 (.228)	1.027 (.240)	.964 (.228)	
$\alpha_3=-0.5$		−.513 (.084)	−.497 (.075)	−.484 (.074)	−.466 (.069)	

is satisfying. There is bias due to the smoothing in the nonparametric ingredient of our three-step semiparametric estimator. As the cross-section sample size increases, the bias decreases. Again, $\alpha_{1,t}$ is not identifiable.

6 Conclusion

We propose a three-step semiparametric estimator of the expected flow utility functions of structural dynamic programming discrete choice models. Our estimator is an extension of the CCP estimators in the literature. The main advantage of our estimator is that it does not require the estimation of the state transition distributions, which could be difficult when the dimension of the observable state variables is even moderately large. Our estimator can be applied to both infinite horizon stationary discrete choice models and the finite horizon nonstationary model with time varying expected flow utility functions and/or state transition distributions.

Appendix

A Proofs for the approximation formulas in infinite horizon stationary problem

Proof of Lemma 3. We prove by induction. For $J = 1$, we have

$$\begin{aligned}
\mathrm{E}(g(s_{t+1}) \mid s_t) &= \mathrm{E}(q^K(s_{t+1})'\rho + \nu(s_{t+1}) \,\big|\, s_t) \\
&= \mathrm{E}(q^K(s_{t+1})'\rho \,\big|\, s_t) + \mathrm{E}(\nu(s_{t+1}) \mid s_t) \\
&= \sum_{k=1}^{K} \rho_k \bar{q}_k(s_t) + \mathrm{E}(\nu(s_{t+1}) \mid s_t) \\
&= \sum_{k=1}^{K} \rho_k \big[q^K(s_t)'\gamma_k + \omega_k(s_t)\big] + \mathrm{E}(\nu(s_{t+1}) \mid s_t) \\
&= \sum_{k=1}^{K} \rho_k q^K(s_t)'\gamma_k + \sum_{k=1}^{K} \rho_k \omega_k(s_t) + \mathrm{E}(\nu(s_{t+1}) \mid s_t) \\
&= q^K(s_t)'\mathit{\Gamma}\rho + \omega^K(s_t)'\rho + \mathrm{E}(\nu(s_{t+1}) \mid s_t),
\end{aligned}$$

which equals to the formula in this lemma.

Suppose the lemma holds for $J = J_*$, we need to show that it holds for $J = J_* + 1$. It follows from the Markov property that

$$\begin{aligned}
\mathrm{E}(g(s_{t+J_*+1}) \mid s_t) &= \mathrm{E}(\mathrm{E}(g(s_{t+J_*+1}) \mid s_{t+1}) \mid s_t) \\
&= \mathrm{E}\Big(q^K(s_{t+1})'\mathit{\Gamma}^{J_*}\rho + \sum_{j=1}^{J_*} \mathrm{E}(\omega^K(s_{t+J_*+1-j})'\mathit{\Gamma}^{j-1}\rho \,\big|\, s_{t+1}) + \mathrm{E}(\nu(s_{t+J_*+1}) \mid s_{t+1}) \,\Big|\, s_t\Big) \\
&= \mathrm{E}(q^K(s_{t+1})'\mathit{\Gamma}^{J_*}\rho \mid s_t) + \sum_{j=1}^{J_*} \mathrm{E}(\omega^K(s_{t+J_*+1-j})'\mathit{\Gamma}^{j-1}\rho \,\big|\, s_t) + \mathrm{E}(\nu(s_{t+J_*+1}) \mid s_t).
\end{aligned}$$

We can organize $\mathrm{E}(q^K(s_{t+1})'\mathit{\Gamma}^{J_*}\rho \mid s_t)$ as follows,

$$\begin{aligned}
\mathrm{E}(q^K(s_{t+1})'\mathit{\Gamma}^{J_*}\rho \mid s_t) &= \mathrm{E}(q^K(s_{t+1})' \mid s_t)\mathit{\Gamma}^{J_*}\rho \\
&= (\bar{q}_1(s_t), \ldots, \bar{q}_K(s_t))\mathit{\Gamma}^{J_*}\rho \\
&= \Big[\Big(q^K(s_t)'\gamma_1, \ldots, q^K(s_t)'\gamma_K\Big) + \omega^K(s_t)'\Big]\mathit{\Gamma}^{J_*}\rho \\
&= q^K(s_t)'\mathit{\Gamma}^{J_*+1}\rho + \omega^K(s_t)'\mathit{\Gamma}^{J_*}\rho.
\end{aligned}$$

Together, we have

$$\mathrm{E}(g(s_{t+J_*+1})\,|\,s_t) - q^K(s_t)'\Gamma^{J_*+1}\rho = \sum_{j=1}^{J_*+1}\mathrm{E}(\omega^K(s_{t+J_*+1-j})'\Gamma^{j-1}\rho\,|\,s_t) + \mathrm{E}(\nu(s_{t+J_*+1})\,|\,s_t).$$

By induction, we conclude that the lemma is correct. ■

Proof of Proposition 1. We prove by showing the order of

$$\sum_{j=1}^{J}\mathrm{E}(\omega^K(s_{t+J-j})'\Gamma^{j-1}\rho\,|\,s_t) + \mathrm{E}(\nu(s_{t+J})\,|\,s_t).$$

First, viewing $\mathrm{E}(\cdot\,|\,s_t)$ as a linear operator (whose norm is 1), we know

$$\|\mathrm{E}(\nu(s_{t+J})\,|\,s_t)\|_2 \le \|\nu(s)\|_2 = O(K^{-m}).$$

Second, by the same token,

$$\|\mathrm{E}(\omega^K(s_{t+J-j})'\Gamma^{j-1}\rho\,|\,s_t)\|_2 \le \|\omega^K(s_{t+J-j})'\Gamma^{j-1}\rho\|_2.$$

Let $\tilde{\rho}_j = \Gamma^j\rho$. It follows from the triangle inequality, we have

$$\|\omega^K(s_{t+J-j})'\Gamma^{j-1}\rho\|_2 = \|\omega^K(s_{t+J-j})'\tilde{\rho}_{j-1}\|_2 \le \sum_{k=1}^{K}|\tilde{\rho}_{j-1,k}|\|\omega_k(s_t)\|_2. \qquad (39)$$

It then follows from Cauchy's inequality that

$$\text{eq. (39)} \le \sqrt{\sum_{k=1}^{K}|\tilde{\rho}_{j-1,k}|^2}\cdot\sqrt{\sum_{k=1}^{K}\|\omega_k(s_t)\|^2} = \|\tilde{\rho}_{j-1}\|\sqrt{\sum_{k=1}^{K}\|\omega_k(s_t)\|^2} \sim \|\tilde{\rho}_{j-1}\|\cdot O(K^{1/2-m}).$$

For $\tilde{\rho}_j = \Gamma^j\rho$, we have $\|\tilde{\rho}_j\| = \|\Gamma^j\rho\| \le \|\Gamma^j\|\|\rho\|$. Because $\|\rho\| \le \|g\|_2 < \infty$, we have $\|\tilde{\rho}_j\| \sim \|\Gamma^j\|$. As a result,

$$\text{eq. (39)} \le \|\Gamma^j\|\cdot O(K^{1/2-m}).$$

The order of the error term is $O(K^{1/2-m})\cdot\sum_{j=1}^{J}\|\Gamma^j\|$. For any finite J, the order is of course $O(K^{1/2-m})$.

We are also interested in the case when $J\to\infty$. Lemma A.1 shows that $\|\Gamma\| \le 1$ with equality only when $\omega^K(s_t) = 0$ or $\Gamma = 0$. For the two trivial cases, if $\Gamma = 0$, then the bound is simply zero; if $\omega^K(s_t) = 0$, the bound is simply $\|\mathrm{E}(\nu(s_{t+J})\,|\,s_t)\|_2 \le \|\nu(s)\|_2 = O(K^{-m})$. If $\|\Gamma\| < 1$, we know that the above display is simply

$$O(K^{1/2-m})\cdot\sum_{j=1}^{\infty}\|\Gamma^j\| \le O(K^{1/2-m})\cdot\sum_{j=1}^{\infty}\|\Gamma\|^j = O(K^{1/2-m})\frac{1}{1-\|\Gamma\|} = O(K^{1/2-m}).$$

So we conclude that the approximation bound is $O(K^{1/2-m})$ regardless of the value of J. ■

Lemma A.1. *We have* $\|\Gamma\| \leq 1$ *with equality only when* $\omega^K(s_t) = 0$ *or* $\Gamma = 0$.

Proof. For any $\rho_0 \in \mathbb{R}^K$, we have

$$\begin{aligned}\mathrm{E}(q^K(s_{t+1})'\Gamma\rho_0 \mid s_t) &= \mathrm{E}(q^K(s_{t+1})' \mid s_t)\Gamma\rho_0 \\ &= q^K(s_t)'\Gamma^2\rho_0 + \omega^K(s_t)'\Gamma\rho_0.\end{aligned}$$

Note that $q^K(s) = (q_1(s), \ldots, q_K(s))'$ is orthogonal to $\omega^K(s) = (\omega_1(s), \ldots, \omega_K(s))'$, because $\omega_k(s) = \sum_{j=K+1}^{\infty} \gamma_{k,j} q_j(s)$. So we have

$$\|q^K(s_t)'\Gamma^2\rho_0 + \omega^K(s_t)'\Gamma\rho_0\|_2 = \|q^K(s_t)'\Gamma^2\rho_0\|_2 + \|\omega^K(s_t)'\Gamma\rho_0\|_2,$$

and

$$\|q^K(s_t)'\Gamma^2\rho_0\|_2 \leq \|\mathrm{E}(q^K(s_{t+1})'\Gamma\rho_0 \mid s_t)\|_2,$$

with equality only when $\omega^K(s_t)'\Gamma\rho_0 = 0$. Since $q_1(s), \ldots, q_K(s)$ are orthonormal, we have

$$\|q^K(s_t)'\Gamma^2\rho_0\|_2 = \|\Gamma^2\rho_0\|.$$

We also know

$$\|\mathrm{E}(q^K(s_{t+1})'\Gamma\rho_0 \mid s_t)\|_2 \leq \|q^K(s_{t+1})'\Gamma\rho_0\| = \|\Gamma\rho_0\|.$$

We conclude that

$$\|\Gamma^2\rho_0\| \leq \|\Gamma\rho_0\|$$

for all $\rho_0 \in \mathbb{R}^K$. Thus, $\|\Gamma\| \leq 1$ with equality only when $\omega^K(s_t) = 0$ or $\Gamma = 0$. ■

Proof of Proposition 2. We write

$$\begin{aligned}\mathrm{E}(g(s_{t+J}) \mid s_t, a_t = a) &= \mathrm{E}(\mathrm{E}(g(s_{t+J}) \mid s_{t+1}) \mid s_t, a_t = a) \\ &= \mathrm{E}(q^K(s_{t+1})'\Gamma^{J-1}\rho + O(K^{1/2-m}) \mid s_t, a_t = a) \\ &= \mathrm{E}(q^K(s_{t+1})'\Gamma^{J-1}\rho \mid s_t, a_t = a) + O(K^{1/2-m}).\end{aligned}$$

We then write

$$\begin{aligned}\mathrm{E}(q^K(s_{t+1})'\Gamma^{J-1}\rho \mid s_t, a_t = a) &= \mathrm{E}(q^K(s_{t+1})' \mid s_t, a_t)\Gamma^{J-1}\rho \\ &= q^K(s_t)'\Gamma(a)\Gamma^{J-1}\rho + \omega^K(s_t, a)'\Gamma^{J-1}\rho \\ &= q^K(s_t)'\Gamma(a)\Gamma^{J-1}\rho + O(K^{1/2-m}).\end{aligned}$$

The last line follows from the proof of proposition 1. The two displays above together shows the result. ■

B Order of the bias term in infinite horizon stationary Markov decision process

We first have that

$$y_{it} - y_{it,J} = \sum_{j=J+1}^{\infty} \beta^j h_{1j}(s_{it}),$$

$$r'_{it} - r'_{it,J} = \Big(\sum_{j=J+1}^{\infty} \beta^j h_{3j}(s_{it})', \sum_{j=J+1}^{\infty} \beta^j h_{2j}(s_{it})' \Big).$$

It follows from Assumption 6.(i),

$$|y_{it} - y_{it,J}| \leq \beta^{J+1} 2\zeta/(1-\beta), \tag{40a}$$

$$\|r_{it} - r_{it,J}\| \leq \beta^{J+1} \sqrt{d_x}\zeta/(1-\beta). \tag{40b}$$

Proposition B.1. *Given Assumption 6, we have* $\sqrt{N}(\theta_J - \theta_{J*}) = O_p(\beta^{J+1})$.

Proof. As an M-estimator, the influence function of θ_J is

$$r_{it,J*}(y_{it,J*} - r'_{it,J*}\theta_{J*}).$$

For the simplicity of exposition, we omit the subscript "*" in $r_{it,J*}$ and $y_{it,J*}$ throughout the proof. To prove the proposition, it suffices to show that

$$r_{it,J}(y_{it,J} - r'_{it,J}\theta_{J*}) = O_p(\beta^{J+1}).$$

Because $y_{it} - r'_{it}\theta_* = 0$, it can be shown that

$$\begin{aligned} r_{it,J}(y_{it,J} - r'_{it,J}\theta_{J*}) &= r_{it,J}(y_{it,J} - r'_{it,J}\theta_{J*}) - r_{it,J}(y_{it} - r'_{it}\theta_*) \\ &= r_{it,J}(y_{it,J} - y_{it}) \qquad (41) \\ &\quad + r_{it,J}(r_{it} - r_{it,J})'\theta_{J*} \qquad (42) \\ &\quad + r_{it,J} r'_{it}(\theta_* - \theta_{J*}). \qquad (43) \end{aligned}$$

It follows from eq. (40) that eq. (41) and (42) are both $O_p(\beta^{J+1})$. To complete the proof, we only need to show that $\theta_* - \theta_{J*} = O_p(\beta^{J+1})$.

We have $\mathrm{E}(r_{it,J}(y_{it,J} - r'_{it,J}\theta_{J*})) = 0$, and

$$\mathrm{E}(r_{it,J}(y_{it,J} - r'_{it,J}\theta_{J*} - y_{it} + r'_{it}\theta_*)) = \mathrm{E}(r_{it,J}(y_{it,J} - r'_{it,J}\theta_{J*})) = 0.$$

It then can be shown that

$$\mathrm{E}(r_{it,J} r'_{it,J})(\theta_{J*} - \theta_*) = \mathrm{E}(r_{it,J}(y_{it,J} - y_{it})) - \mathrm{E}(r_{it,J}(r_{it,J} - r_{it})')\theta_*.$$

Applying eq. (40), we have $\theta_* - \theta_{J*} = O_p(\beta^{J+1})$. So we conclude that $r_{it,J}(y_{it,J} - r'_{it,J}\theta_{J*})$ is $O_p(\beta^{J+1})$. ■

Proposition B.2. *Given Assumption 6, we have* $\sqrt{N}(\theta_{J*} - \theta_*) = O_p(\beta^{J+1})$.

Proof. We write $\sqrt{N}(\theta_{J*} - \theta_*) = \sqrt{N}(\theta_{J*} - \theta_J) + \sqrt{N}(\theta_J - \theta_*)$. Proposition B.1 has shown that $\sqrt{N}(\theta_{J*} - \theta_J) = O_p(\beta^{J+1})$. Lemma B.1 below shows that $\sqrt{N}(\theta_J - \theta_*) = O_p(\beta^{J+1})$. ■

Lemma B.1. *Given Assumption 6, we have* $\sqrt{N}(\theta_J - \theta_*) = O_p(\beta^{J+1})$.

Proof. For the simplicity of exposition, we omit the subscript "*". We decompose $\theta_J - \theta$ as the following sum,

$$(R_J{}'R_J/N)^{-1}(R_J{}'Y_J)/N - (R_J{}'R_J/N)^{-1}R'Y/N \tag{44}$$

$$+(R_J{}'R_J/N)^{-1}R'Y/N - (R'R/N)^{-1}R'Y/N. \tag{45}$$

Below, we derive the orders of eq. (44) and (45).

We first derive the order of eq. (44), which equals

$$\mathrm{E}(r_{it,J}r'_{it,J})^{-1}[(R_J{}'Y_J)/N - R'Y/N] + o_p(1),$$

if $(R_J{}'Y_J)/N - R'Y/N = O_p(1)$, which will be shown. Because Assumption 6.(ii) has assumed the smallest eigenvalue of $\mathrm{E}(r_{it,J}r'_{it,J})$ is bounded from zero, it suffices to derive the order of $(R_J{}'Y_J)/N - R'Y/N$ to find the order of eq. (44). Decompose $(R_J{}'Y_J)/N - R'Y/N$ as the following sum

$$(R'_JY_J - R'_JY)/N + (R'_JY - R'Y)/N.$$

To derive the order of $(R'_JY_J - R'_JY)/N$, we consider its squared mean and then use the Markov inequality. Recall that $r_{it,J}$ is a d_θ-dimensional vector. We have

$$\mathrm{E}(\|(R'_JY_J - R'_JY)/N\|^2) = \sum_{\ell=1}^{d_\theta} \mathrm{E}\bigg(\Big[N^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T} r_{\ell,it,J}(y_{it,J} - y_{it})\Big]^2\bigg).$$

Since d_θ is finite, it suffices to consider the order of an individual term of the above sum. Since T is finite, without loss of generality, we consider the case, where $T = 2$. The term $\mathrm{E}([N^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T=2} r_{\ell,it,J}(y_{it,J} - y_{it})]^2)$ equals

$$(2n)^{-1}\,\mathrm{E}([r_{\ell,it,J}(y_{it,J} - y_{it})]^2) + (2n)^{-1}\,\mathrm{E}(r_{\ell,i1,J}(y_{i1,J} - y_{i1})r_{\ell,i2,J}(y_{i2,J} - y_{i2})).$$

Because the regressors $r_{\ell,it,J}$ are bounded, and the difference $|y_{it} - y_{it,J}|$ is of order β^{J+1}, we conclude

$$\mathrm{E}\bigg(\Big[N^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T=2} r_{\ell,it,J}(y_{it,J} - y_{it})\Big]^2\bigg) = O(n^{-1}\beta^{2(J+1)}).$$

By the Markov inequality, $(R_J'Y_J - R_J'Y)/N = O_p(N^{-1/2}\beta^{(J+1)})$. Similar arguments show that $(R_J'Y - R'Y)/N = O_p(N^{-1/2}\beta^{(J+1)})$. Hence, we conclude $(R_J'Y_J - R'Y)/N$, henceforth eq. (44), is $O_p(N^{-1/2}\beta^{(J+1)})$.

We next derive the order of eq. (45), which rewritten as follows,

$$[(R_J{}'R_J/N)^{-1} - (R'R/N)^{-1}]R'Y/N.$$

The term $R'Y/N = O_p(N^{-1/2})$. Letting

$$U_J \equiv 2R_J'(R - R_J)/N + (R - R_J)'(R - R_J)/N,$$

the term $(R_J{}'R_J/N)^{-1} - (R'R/N)^{-1}$ equals

$$(R_J{}'R_J/N)^{-1}U_J(R_J{}'R_J/N)^{-1}[I + U_J(R_J{}'R_J/N)^{-1}]^{-1}.$$

This can be shown by the Woodbury formula in linear algebra. It is easy to see that $U_J = O_p(\beta^{J+1})$, hence the above display is also $O_p(\beta^{J+1})$. We then conclude that eq. (45), is $O_p(N^{-1/2}\beta^{(J+1)})$. ∎

C Order of the bias term in general Markov decision process

Proof of Proposition 3. Let $\nu(s_T) = \bar{V}_T(s_T) - \sum_{j=1}^{K} \rho_j q_j(s_T)$ be the residual of series approximation. Let $\nu = H - Q_K\rho^K$ denote the vector consisting of $\beta^{T-t}\,\tilde{\mathrm{E}}(\nu(s_T)\,|\,s_{it})$ for i and t. Here $\rho^K = (\rho_1, \dots, \rho_K)'$.

We will show that $\|\sqrt{N}(\theta_K - \theta_*)\| \le \sqrt{\lambda_{max}}\|\nu\|$, where λ_{max} is the largest eigenvalue of $R_* \equiv (N^{-1}R'M_KR)^{-1}$. Replacing Y in $\theta_K = (R'M_KR)^{-1}R'M_KY$ with $Y = R\theta_* + H$, we have $\theta_K - \theta_* = (R'M_KR)^{-1}R'M_K\nu$. Then

$$\begin{aligned}
\|N(\theta_K - \theta_*)\|^2 &= \|(R_*)^{-1}R'M_K\nu\|^2 \\
&= \nu'M_KR(R_*)^{-1}(R_*)^{-1}R'M_K\nu \\
&= \nu'M_KR(R_*)^{-1/2}(R_*)^{-1}(R_*)^{-1/2}R'M_K\nu \\
&\le \lambda_{max}\cdot\nu'M_KR(R_*)^{-1/2}(R_*)^{-1/2}R'M_K\nu \\
&= \lambda_{max}\cdot\nu'M_KR(R_*)^{-1}R'M_K\nu \\
&= N\lambda_{max}\cdot\nu'M_KR(R'M_KR)^{-1}R'M_K\nu \\
&\le N\lambda_{max}\|\nu\|^2.
\end{aligned}$$

We have

$$\|\sqrt{N}(\theta_K - \theta_*)\| \leq \sqrt{\lambda_{max}}\|\nu\|. \tag{46}$$

By the condition that $\sup_{s\in\mathcal{S}}\left|\bar{V}_T(s_T) - \sum_{j=1}^K \rho_j q_j(s)\right| = \sup_{s\in\mathcal{S}}|\nu(s)| \leq O(K^{-\xi})$, we have $\|\nu\| = \sqrt{N}O(K^{-\xi})$. Also, $\sqrt{\lambda_{max}} < \infty$ with probability 1. So we have $\|(\theta_K - \theta_*)\| = O_p(K^{-\xi})$. ■

D Asymptotic variance of three-step semiparametric M-estimators with generated dependent variables

This appendix derives the asymptotic variance for the three-step semiparametric M-estimators with generated dependent variabels in the second step nonparametric regressions. First, we describe the three-step estimator for which our proposed formula applies. Second, we derive the influence function and the asymptotic variance. Third, we derive the asymptotic variance formula when the second step nonparametric regressions are "partial means" in Newey's (1994b) terminology. The general formula can be applied to our three-step CCP estimator.

Assume that we have random sample $\{y_i, x_i, z_i : i = 1, \dots, n\}$ about random variables y, x and z. In the first step, we obtain an estimator $\hat{p}(x, z)$ of $p_*(x, z) \equiv \mathrm{E}(y \mid x, z)$ by nonparametric regression of y on x and z. The variable $w_* = \varphi(x, z, p_*(x, z))$ is the dependent variable in the second step nonparametric regression. The function φ is known.

In the second step, the goal is to estimate

$$h_*(x; w_*) \equiv \mathrm{E}(w_* = \varphi(x, z, p_*(x, z)) \,|\, x).$$

Obviously, when $\varphi(x, z, p)$ is linear in p, the first step and second step can be merged by the law of iterated expectation. So we are interested in the case where φ is nonlinear in p. Because $p_*(x, z)$, hence w_*, is not observable, we use $\hat{w}_i = \varphi(x_i, z_i, \hat{p}(x_i, z_i))$ in the nonparametric regression. The nonparametric regression estimator of $\hat{w}$ on x is denoted by $\hat{h}(x; \hat{w})$.

The third step defines a semiparametric M-estimator $\hat{\theta}$ that solves a moment equation of the form

$$n^{-1}\sum_{i=1}^n m(x_i, \theta, \hat{h}(x_i; \hat{w})) = 0.$$

Suppose that $\mathrm{E}\big(m(x, \theta_*, h_*(x; w_*))\big) = 0$. Moreover, assume that m depends on h only through its value $h(x)$. Our interest is to characterize the first-order asymptotic properties of $\hat{\theta}$.

We derive the influence function of $\hat{\theta}$ using the pathwise derivative approach in Newey (1994a, page 1360-1361). For a path $\{F_\tau\}$, that equals the true distribution of (y,x,z) when $\tau=\tau_*$, let $\mu(F_\tau)$ be the probability limit of $\hat{\theta}$ if data were generated from F_τ. We have

$$\mathrm{E}_\tau(m(x,\mu(F_\tau),h_\tau(x;w_\tau)))=0,$$

where $w_\tau=\varphi(x,z,p_\tau(x,z))$, $p_\tau=\mathrm{E}_\tau(y\mid x,z)$, and $h_\tau(x;w_\tau)=\mathrm{E}_\tau(w_\tau\mid x)$. Let $M\equiv \partial\,\mathrm{E}(m(x,\theta,h_*(x;w_*)))\big/\partial\theta\big|_{\theta=\theta_*}$, and $S(y,x,z)=\partial\ln(\mathrm{d}\,F_\tau)/\partial\tau$ be score function (the derivatives with respect to τ are always evaluated at $\tau=\tau_*$). For expositional simplity, we write $\partial f(x_*)/\partial x=\partial f(x)/\partial x|_{x=x_*}$ for a generic function $f(x)$.

It can be shown by Newey's pathwise derivative approach that $-M\partial\mu(F_\tau)/\partial\tau$ equals the sum of the following three terms

$$\mathrm{E}(m(x,\theta_*,h_*(x;w_*))S(y,x,z)), \tag{i}$$

$$\mathrm{E}\bigg(\frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}(w_*-h_*(x))S(y,x,z)\bigg), \tag{ii}$$

$$\mathrm{E}\bigg(\frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}\frac{\partial\varphi(x,z,p_*(x,z))}{\partial p}(y-p_*(x,z))S(y,x,z)\bigg). \tag{iii}$$

Here term (i) corresponds to the leading term that captures the uncertainty left if we know h_* and p_*, hence w_*. Term (ii) accounts for the sampling variation in estimating $\mathrm{E}(w_*\mid x)$, when w_* is known. Term (iii) reflects the sampling variation in $\hat{p}$. It is easy to derive term (ii) from (Newey, 1994a, page 1360-1361). Term (iii) needs more explanation.

Term (iii) is derived from the pathwise derivative

$$\partial\,\mathrm{E}(m(x,\theta_*,h_*(x;w_\tau)))/\partial\tau \tag{47}$$

We have

$$\begin{aligned}
\text{eq. (47)} &= \partial\,\mathrm{E}\bigg(\frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}\frac{\partial h_*(x;w)}{\partial w}\bigg[\frac{\partial\varphi(x,z,p_*(x,z))}{\partial p}p_\tau\bigg]\bigg)\bigg/\partial p\\
&= \partial\,\mathrm{E}\bigg(\frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}\mathrm{E}\bigg(\frac{\partial\varphi(x,z,p_*(x,z))}{\partial p}p_\tau\,\bigg|\,x\bigg)\bigg)\bigg/\partial p\\
&= \partial\,\mathrm{E}\bigg(\frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}\frac{\partial\varphi(x,z,p_*(x,z))}{\partial p}p_\tau\bigg)\bigg/\partial p.
\end{aligned} \tag{48}$$

The second line follows from the fact that $h_*(x;w)=\mathrm{E}(w\mid x)$ is a linear functional, so its Fréchet derivative is itself. The last line follows from the law of iterated expectation. Since $p_\tau=\mathrm{E}_\tau(y\mid x,z)$ is projection, term (iii) can be derived from equation (4.5) of Newey (1994a).

From term (i), (ii) and (iii), we conclude that the influence function of the three-step estimator $\hat{\theta}$ is

$$-M^{-1}\Big[m(x,\theta_*,h_*(x;w_*)) + \frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}(w_* - h_*(x)) + \frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}\frac{\partial \varphi(x,z,p_*(x,z))}{\partial p}(y - p_*(x,z))\Big].$$

Remark 4 (Difference between generated regressors and generated dependent variables). Hahn and Ridder (2013) derives the asymptotic variance for semiparametric estimators with generated regressors. The estimation there also proceeds in three steps. The estimator in the first step is used to generate regressors for the nonparametric regression in the second step. An interesting conclusion of their paper is that when third step estimator is linear in the conditional mean estimated in the second step, the sampling variation in the first step estimation does not affect the asymptotic variance of the three-step estimator. However, it easy to see from term (iii) here that if the first step is to generate dependent variables rather than regressors, the sampling variation in the first step will always affect the asymptotic variance of the three-step estimator. ■

We next consider the case where the second step nonparametric regression involves partial means instead of full means. In particular, suppose there is a discrete variable a taking value from 1,...,K. In our dynamic discrete choice model, a is y itself. The second step is to estimate

$$h_{k*}(x;w_*) = \mathrm{E}(w_* = \varphi(x,z,p_*(x,z))\,|\,x, a = k).$$

The leading term does not change. Using the following identity,

$$\begin{aligned} h_{k*}(x;w_*) &= \mathrm{E}(w_* = \varphi(x,z,p_*(x,z))\,|\,x, a = k) \\ &= \mathrm{E}\left(\frac{1(a=k)w_*}{\Pr(a=k\mid x)}\,\middle|\, x\right), \end{aligned}$$

we can show that second term now becomes

$$\mathrm{E}\left(\frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}\left(\frac{1(a=k)w_*}{\Pr(a=k\mid x)} - h_{k*}(x)\right) S(y,x,z)\right).$$

For the third term, we have

$$\begin{aligned} \text{eq. (47)} &= \partial\,\mathrm{E}\left(\frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}\frac{\partial h_*(x;w)}{\partial w}\left[\frac{\partial \varphi(x,z,p_*(x,z))}{\partial p}p_\tau\right]\right)\Big/\partial p \\ &= \partial\,\mathrm{E}\left(\frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}\,\mathrm{E}\left(\frac{\partial \varphi(x,z,p_*(x,z))}{\partial p}p_\tau\,\middle|\, x, a=k\right)\right)\Big/\partial p \\ &= \partial\,\mathrm{E}\left(\frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}\frac{1(a=k)}{\Pr(a=k\mid x)}\frac{\partial \varphi(x,z,p_*(x,z))}{\partial p}p_\tau\right)\Big/\partial p. \end{aligned}$$

Hence the third adjustment term of the influence function is

$$\mathrm{E}\bigg(\frac{\partial m(x,\theta_*,h_*(x;w_*))}{\partial h}\frac{1(a=k)}{\Pr(a=k\mid x)}\frac{\partial\varphi(x,z,p_*(x,z))}{\partial p}(y-p_*(x,z))S(y,x,z)\bigg).$$

References

Aguirregabiria, V. (2010). Another look at the identification of dynamic discrete decision processes: An application to retirement behavior. *Journal of Business & Economic Statistics 28*(2), 201–218.

Aguirregabiria, V. and P. Mira (2002). Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models. *Econometrica 70*(4), 1519–1543.

Aguirregabiria, V. and P. Mira (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics 156*(1), 38–67.

Altuğ, S. and R. A. Miller (1998). The effect of work experience on female wages and labour supply. *The Review of Economic Studies 65*(1), 45–85.

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 79*(6), 1823–1867.

Arcidiacono, P. and R. A. Miller (2016). Nonstationary dynamic models with finite dependence. Technical report, Mimeo.

Blundell, R., M. Costa Dias, C. Meghir, and J. Shaw (2016). Female labor supply, human capital, and welfare reform. *Econometrica 84*(5), 1705–1753.

Chou, C. (2016). Identification and linear estimation of general dynamic programming discrete choice models. Mimeo, University of Leicester, Leicester.

Donald, S. G. and W. K. Newey (1994). Series estimation of semilinear models. *Journal of Multivariate Analysis 50*(1), 30–40.

Eckstein, Z. and K. I. Wolpin (1989). Dynamic labour force participation of married women and endogenous work experience. *The Review of Economic Studies 56*(3), 375–390.

Fang, H. and Y. Wang (2015). Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions. *International Economic Review 56*(2), 565–596.

Hahn, J. and G. Ridder (2013). Asymptotic variance of semiparametric estimators with generated regressors. *Econometrica 81*(1), 315–340.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2003). Fifty years of mincer earnings regressions. Working Paper 9732, National Bureau of Economic Research.

Helmberg, G. (2008). *Introduction to spectral theory in Hilbert space*. Mineola, N.Y.: Dover Publications, Inc.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies 60*(3), 497–529.

Hotz, V. J., R. A. Miller, S. Sanders, and J. Smith (1994). A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies 61*(2), 265–289.

Huggett, M., G. Ventura, and A. Yaron (2011). Sources of lifetime inequality. *The American Economic Review 101*(7), 2923–2954.

Kalouptsidi, M., P. T. Scott, and E. Souza-Rodrigues (2015). Identification of counterfactuals and payoffs in dynamic discrete choice with an application to land use. Working paper, NBER.

Kristensen, D., L. Nesheim, and A. de Paula (2014). CCP and the estimation of nonseparable dynamic models. Mimeo, University College London.

Magnac, T. and D. Thesmar (2002). Identifying dynamic discrete decision processes. *Econometrica 70*(2), 801–816.

Matzkin, R. L. (1992). Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. *Econometrica*, 239–270.

Miller, R. A. (1997). Estimating models of dynamic optimization with microeconomic data. In M. H. Pesaran and P. Schmidt (Eds.), *Handbook of Applied Econometrics*, Volume 2, Chapter 6, pp. 246–299. Blackwell Oxford.

Newey, W. K. (1994a). The asymptotic variance of semiparametric estimators. *Econometrica*, 1349–1382.

Newey, W. K. (1994b). Kernel estimation of partial means and a general variance estimator. *Econometric Theory 10*(02), 1–21.

Norets, A. and X. Tang (2014). Semiparametric inference in dynamic binary choice models. *The Review of Economic Studies 81*, 1229–1262.

Pesendorfer, M. and P. Schmidt-Dengler (2008). Asymptotic least squares estimators for dynamic games. *The Review of Economic Studies 75*(3), 901–928.

Srisuma, S. and O. Linton (2012). Semiparametric estimation of markov decision processes with continuous state space. *Journal of Econometrics 166*(2), 320–341.

Wasserman, L. (2006). *All of nonparametric statistics.* New York: Springer.