Journal of Econometrics 188 (2015) 435–446

Contents lists available at ScienceDirect

Journal of Econometrics

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

Estimation of dynamic discrete models from time aggregated data[☆]

Han Hong[a], Weiming Li[b,*], Boyu Wang[c]

[a] Department of Economics Stanford University, Stanford, CA 94305, USA
[b] ISEM, Capital University of Economics and Business, Beijing, 100070, China
[c] RadiumOne Inc. 55 Second Street, San Francisco, CA 94105, USA

ARTICLE INFO

Article history:
Available online 19 March 2015

JEL classification:
E00
G12
C51
C52

Keywords:
Dynamic discrete choice models
Maximum likelihood estimator
Semiparametric methods

ABSTRACT

An important component in dynamic discrete choice models and dynamic discrete games is the transition density of state variables from the current period to the next period. Most empirical dynamic discrete choice models identify the theoretical time interval in the behavioral model with that observed in the data set. However, many empirical data sets are time aggregated. In this paper, we show that when the time interval in the behavioral theory model differs from that in the observed data, difficulties with nonparametric identification and specification arise. In addition, we study the properties of parametric maximum likelihood estimators and flexible semiparametric estimators of the transition density in dynamic discrete models with time aggregated data sets.

1. Introduction

In dynamic discrete choice models and discrete games, developed in Rust (1987) and Hotz and Miller (1993), agents solve an optimal dynamic action as a function of the current state variables. The state variables evolve according to a law of motion that also depends on the actions by the agents. A recent literature generalizes to dynamic games and develops nonparametric identification results and flexible semiparametric estimators, including contributions from, among others, Bajari et al. (2007), Aguirregabiria and Mira (2007), Berry et al. (2003), Pesendorfer and Schmidt-Dengler (2010), Jenkins et al. (2004), Magnac and Thesmar (2002), Kasahara and Shimotsu (2008), Hu and Shum (2012), Arcidiacono and Miller (2011), Norets (2009), Imai et al. (2009) and Bajari et al. (2009).

A potential important issue in empirically estimating dynamic discrete choice models is that there is often a discrepancy between the frequency of observations in the data set and the decision frequency in the behavior model. Often times individual behaviors are reported in prespecified time intervals in each data set. This can be daily, monthly, quarterly or annually. For example, the Health and Retirement Survey was conducted once every two years. This requires that a researcher adapts the frequency of the decision making in the dynamic discrete choice model to the observed time intervals reported in the data set (Fang and Wang, 2010). Such an approach will lead to inconsistent estimates of the utility parameters when the decision time model is misspecified. Other applications of dynamic discrete models include Rust (1987), Jenkins et al. (2004) and Ryan (2012).

Exceptions are Doraszelski and Judd (2012) and Arcidiacono et al. (2012), who formulate and estimate continuous time dynamic discrete models. In this paper we consider equal-spaced discrete time models, in which the data frequency is coarser than the model frequency. Our results complement those in Blevins (2013, 2014), who provide identification results for continuous time models using discretely observed data using the logarithmic root of the discrete time state transition matrix.

The structural parameters in a dynamic discrete choice model include the discount rate, the error distribution, the period utility function and the state variable transition density. When the discount rate and the error distribution are known as is conventionally assumed in the literature, the dynamic discrete model is exactly identified. In particular, the state variable transition density is directly identified and can be estimated from the data. The period utility function can then be recovered from the reduced form choice probabilities under a set of exclusion restrictions.

These results no longer hold when only time aggregated data is available where the data frequency differs from the model frequency. With time aggregated data, only multi-period state variable transition density can be directly identified and recovered

☆ We acknowledge generous support by the National Science Foundation (SES 1024504, 1164589) and SIEPR. We also thank the associate editor and two anonymous referees for helpful comments.

* Corresponding author.
E-mail address: liweiming291@163.com (W. Li).

http://dx.doi.org/10.1016/j.jeconom.2015.03.009
0304-4076/

from the data set. Multi-period transition density, however, is a convoluted function of the single-period transition density and the conditional choice probabilities of actions as a function of the state variables. The inverse mapping of this relation might not exist or might admit multiple solutions. As a consequence, the single period state transition density can be either correctly specified and identified, or misspecified, or unidentified. Even when the population model is correctly specified, nonparametric methods of identification and estimation that are applicable for disaggregated time data still cannot be applied directly to aggregate time data sets.

Parametric and semiparametric maximum likelihood methods can be applied to aggregated time data when the population model is correctly specified and identified. We discuss the adaptation of the estimation methods for dynamic discrete models developed in Rust (1987) and Hotz and Miller (1993) to time aggregate data sets, and present their statistical and computational properties. The EM algorithm (Arcidiacono and Jones, 2003; Arcidiacono and Miller, 2011) can be utilized to efficiently compute the MLE estimators. We find that an estimator based on aggregated time data might achieve more statistical efficiency than an estimator based on regularly spaced observations. Numerical simulations show that the performance of likelihood based estimators in finite samples accords with the predictions of the asymptotic theory.

2. Model and data

We consider a stationary dynamic environment. Let $t = 0, 1, \ldots, \infty$ denote time. The notations to be introduced below will apply to both single agent dynamic discrete choice settings and multi-agent dynamic discrete games with n agents, in which case the decision making of forward looking rational agents is assumed to be consistent with a dynamic program as in Rust (1987) or a Markov perfect dynamic equilibrium.

In each time period t, agent j chooses an action a_{jt} from $\mathcal{A} = \{1, \ldots, L\}$. Its observed state variable is denoted by x_{jt}, taking the value from a finite set of $\mathcal{X} = \{0, 1 \cdots K\}$. Further, let $a_t = (a_{jt}, \forall j) \in \mathcal{A}^m$ and $x_t = (x_{jt}, \forall j) \in \mathcal{X}^m$ be the action and state profile vectors of all agents. Also let a_{-j} be a vector of strategies for all players excluding j, and similarly for x_{-j}. While x_{jt} can be perfectly observed by both the agents and the econometrician subject to the data periodicity to be discussed below, there is a set of *unobserved state variable* $\epsilon_{jt}(a_{jt})$ observable only by agent j as private information. Throughout the paper we maintain the conventional conditional independence and distribution assumptions following Rust (1987).

Assumption 1. The error terms $\epsilon_{jt}(a_{jt})$ are identically and independently distributed across t, j and a_{jt} and follow Type I extreme value distribution.

Assumption 2. Conditional on x_t and a_t, x_{t+1} is independent of ϵ_t.

The complete state variable of agent j at time t is $s_{jt} = (x_{jt}, \epsilon_{jt})$. In each period agent j derives utility $u_j(a_t, x_t, \epsilon_{jt}) = \Pi_j(a_{jt}, a_{-jt}, x_t; \theta_u) + \epsilon_{jt}(a_{jt})$. The period mean utility function $\Pi_j(a_{jt}, a_{-jt}, x_t; \theta_u)$ is described by a parameter vector θ_u, which includes the nonparametric specification as a special case when the state variables are discrete. Agents are forward looking with a given discount rate β. In equilibrium, they choose best response a_{jt} for a given x_t and ϵ_{jt} to maximize a discounted sum of expected future utility functions,

$$E\left[\sum_{\tau=0}^{\infty} \beta^{\tau} u_j(a_{t+\tau}, x_{t+\tau}, \epsilon_{t+\tau}) | a_{jt}, x_t, \epsilon_{jt}\right]. \quad (1)$$

The Markov perfect equilibrium conditional choice probabilities $\sigma_j(a_j|x)$ are obtained by integrating out ϵ_{jt} in the optimal response function $a_{jt}(x_t, \epsilon_{jt})$.

Because of the private information assumption, the dynamic discrete game is observationally equivalent to a single agent dynamic discrete model, in which the period expected utility function of agent j is given by $\Pi_j(a_j, x) = \sum_{a_{-j} \in A_{-j}} \Pi_j(a_j, a_{-j}, x)\sigma_{-j}(a_{-j}|x)$, and in which agent j follows the Bellman principle of stochastic dynamic programming

$$W_j(x_t, \epsilon_{jt}; \sigma_{-j}) = \max_{a_{jt} \in \mathcal{A}} \left\{ \Pi_j(a_{jt}, x_t) + \epsilon_{jt} + \beta \int_{x', \epsilon'} W_j(x', \epsilon'; \sigma_{-j}) \times dF(x', \epsilon' | a_{jt}, x_t, \epsilon_{jt}) \right\},$$

where the value function for player j, $W_j(\cdot, \cdot; \sigma_{-j})$, is defined as the fixed point of the above functional equation.

Because of Assumptions 1 and 2, the Bellman principle can be expressed through a *choice specific value function* $V_j(a_j, x)$ defined as

$$V_j(a_j, x) = \Pi_j(a_j, x; \theta_u) + \beta \int_{x'} V_j(x') dF(x'|a_j, x), \quad (2)$$

where the *social surplus function* or *ex ante value function* is defined by

$$V_j(x) = \int W_j(x, \epsilon_j; \sigma) f(\epsilon_j) d\epsilon_j.$$

Under Assumption 1, the social surplus function is in turn related to the choice specific value functions through the following relation

$$V_j(x) = G(V_j(a_j, x), \forall a_j) \equiv \log \sum_{a_i=0}^{L} \exp(V_j(a_j, x)). \quad (3)$$

In a Markov perfect equilibrium of a discrete game, the transition distribution of the observed state variables satisfies, by conditional independence

$$F(x_{t+1}|a_{jt}, x_t, \epsilon_{jt}) = F(x_{t+1}|a_{jt}, x_t) = \sum_{a_{-jt}} g(x_{t+1}|x_t, a_{jt}, a_{-jt}) \cdot \sigma_{-j}(a_{-jt}|x_t),$$

where $g(x_{t+1}|x_t, a_{jt}, a_{-jt})$ denotes the single period conditional density of the next period state variables given the current state variables and actions. In the following, we will focus on a single agent dynamic model with infinite horizon and stationarity in which $j = 1, \ldots, n$ denotes individuals operating in n independent markets.

Estimation of a dynamic discrete choice model requires a cross section of observations on the choice profile and state variables (a_t, x_t), which provides information about the conditional choice probabilities $p(a_t|x_t)$, and data that can be used to recover the transition density of the observed state variables $F(x_{t+1}|a_{it}, x_t, \epsilon_{it})$. In a typical stationary model in which the frequency of data availability coincides with the model decision frequency, this is given by the conditional distribution of x_{t+1} given a_t and x_t. We denote a balanced panel data of *non time aggregated* data set by

$$D = \{(a_{i,0}, x_{i,0}), (a_{i,1}, x_{i,1}), \ldots, (a_{i,T}, x_{i,T})\}_{1 \le i \le n}.$$

Instead of this conventional setting, in the rest of the paper we focus on *time aggregated data* in which the data frequency is coarser than the model frequency. In the model, time increments by each period. However, data is only available in every r periods. In other words, for $r \geq 1$, we assume availability of only a balanced time aggregated panel of observations

$$D_r = \{(a_{i,0}, x_{i,0}), (a_{i,r}, x_{i,r}), \ldots, (a_{i,Mr}, x_{i,Mr})\}_{1 \le i \le n}$$