

Conditional Choice Probabilities and the Estimation of Dynamic Models

V. JOSEPH HOTZ
University of Chicago

and

ROBERT A. MILLER
Carnegie Mellon University

First version received April 1989; final version accepted January 1993 (Eds.)

This paper develops a new method for estimating the structural parameters of (discrete choice) dynamic programming problems. The method reduces the computational burden of estimating such models. We show the valuation functions characterizing the expected future utility associated with the choices often can be represented as an easily computed function of the state variables, structural parameters, and the probabilities of choosing alternative actions for states which are feasible in the future. Under certain conditions, nonparametric estimators of these probabilities can be formed from sample information on the relative frequencies of observed choices using observations with the same (or similar) state variables. Substituting the estimators for the true conditional choice probabilities in formulating optimal decision rules, we establish the consistency and asymptotic normality of the resulting structural parameter estimators. To illustrate our new method, we estimate a dynamic model of parental contraceptive choice and fertility using data from the National Fertility Survey.

1. INTRODUCTION

Over the last several years there has been increasing interest in estimating structural models of dynamic discrete choice. Empirical applications have been undertaken in the areas of fertility (Wolpin (1984)), job search (Kiefer and Newmann (1979, 1981), Flinn and Heckman (1982), Lancaster and Chesher (1983), Wolpin (1987)), job matching (Miller (1982, 1984)), labour force participation (Eckstein and Wolpin (1989a), Gönül (1989)), Berkovec and Stern (1991)), patent renewal (Pakes (1986)) and the replacement of bus engines (Rust (1987))¹. These studies derive the stochastic process generating an agent's choice sequence from the solution to a dynamic optimization problem, which depends upon structural parameters characterizing the agent's preferences and her constraints. The estimation problem is to identify and consistently estimate the structural parameters from data on choices and other observed variables. Such estimates enable one to examine and forecast how exogenous changes in economic constraints affect choices.

In contrast to models with continuous choices which can be estimated from the first-order conditions, the optimal decision rules for dynamic discrete choice models are characterized by inequality conditions. This has prompted researchers to (numerically) solve the valuation function characterizing the optimal sequence of choices in order to

1. See Eckstein and Wolpin (1989b) for a recent survey of this fast growing field.

Hotz-Miller inversion proposition:

There exists an invertible mapping ψ from CCPs (p 's) to $(v_j - v_k)$'s

$$V_{t+1} = v_{0t+1} + \mathbb{E} \max\{\epsilon_{0t+1}, v_{1t+1} - v_{0t+1} + \epsilon_{1t+1}, \dots, v_{Jt+1} - v_{0t+1} + \epsilon_{Jt+1}\}$$

$$= v_{0t+1} + \mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p_{t+1}) + \epsilon_{1t+1}, \dots, \psi_0^J(p_{t+1}) + \epsilon_{Jt+1}\}$$

The p 's come from data, but we need:

The p 's come from data, but we need:

- The mapping $\psi(\cdot)$

The p 's come from data, but we need:

- The mapping $\psi(\cdot)$
- To calculate expectations of ϵ 's

The p 's come from data, but we need:

- The mapping $\psi(\cdot)$
- To calculate expectations of ϵ 's
- To handle the v_{0t+1} 's

The p 's come from data, but we need:

- The mapping $\psi(\cdot)$
- To calculate expectations of ϵ 's
- To handle the v_{0t+1} 's

Now let's go through an example with $\mathcal{J} = \{0, 1\}$ and stochastic X 's

The p 's come from data, but we need:

- The mapping $\psi(\cdot)$
- To calculate expectations of ϵ 's
- To handle the v_{0t+1} 's

Now let's go through an example with $\mathcal{J} = \{0, 1\}$ and stochastic X 's

Suppose option 0 is a **terminal choice**—no further choices after choosing 0

How can we simplify the conditional value function v_{1t} in this setting?

$$v_{1t}(X_t) = u_1(X_t) + \beta \sum_{X_{t+1}} v_{t+1}(X_{t+1}) f_1(X_{t+1}|X_t)$$

How can we simplify the conditional value function v_{1t} in this setting?

$$v_{1t}(X_t) = u_1(X_t) + \beta \sum_{X_{t+1}} v_{t+1}(X_{t+1}) f_1(X_{t+1}|X_t)$$

$$= u_1(X_t) + \beta \sum_{X_{t+1}} \left[v_{0t+1}(X_{t+1}) + \mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p(X_{t+1})) + \epsilon_{1t+1}\} \right] f_1(X_{t+1}|X_t)$$

How can we simplify the conditional value function v_{1t} in this setting?

$$\begin{aligned} v_{1t}(X_t) &= u_1(X_t) + \beta \sum_{X_{t+1}} v_{t+1}(X_{t+1}) f_1(X_{t+1}|X_t) \\ &= u_1(X_t) + \beta \sum_{X_{t+1}} \left[v_{0t+1}(X_{t+1}) + \mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p(X_{t+1})) + \epsilon_{1t+1}\} \right] f_1(X_{t+1}|X_t) \end{aligned}$$

If $v_{0t+1}(X_{t+1}) = X_{t+1}\alpha_0$, no backwards recursion needed

How can we simplify the conditional value function v_{1t} in this setting?

$$\begin{aligned} v_{1t}(X_t) &= u_1(X_t) + \beta \sum_{X_{t+1}} v_{t+1}(X_{t+1}) f_1(X_{t+1}|X_t) \\ &= u_1(X_t) + \beta \sum_{X_{t+1}} \left[v_{0t+1}(X_{t+1}) + \mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p(X_{t+1})) + \epsilon_{1t+1}\} \right] f_1(X_{t+1}|X_t) \end{aligned}$$

If $v_{0t+1}(X_{t+1}) = X_{t+1}\alpha_0$, no backwards recursion needed

... so long as we can handle the expectation term

With T1EV errors:

$$\mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p(X_{t+1})) + \epsilon_{1t+1}\} = -\log(p_0(X_{t+1})) + c$$

With T1EV errors:

$$\mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p(X_{t+1})) + \epsilon_{1t+1}\} = -\log(p_0(X_{t+1})) + c$$

How?

With T1EV errors:

$$\mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p(X_{t+1})) + \epsilon_{1t+1}\} = -\log(p_0(X_{t+1})) + c$$

How? We showed before that $V_{t+1} = v_{j_{t+1}} - \log(p_{j_{t+1}}) + c$ for any j when $\epsilon \sim \text{T1EV}$

With T1EV errors:

$$\mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p(X_{t+1})) + \epsilon_{1t+1}\} = -\log(p_0(X_{t+1})) + c$$

How? We showed before that $V_{t+1} = v_{j_{t+1}} - \log(p_{j_{t+1}}) + c$ for any j when $\epsilon \sim \text{T1EV}$

Therefore:

$$v_1(X_t) = u_1(X_t) + \beta \sum_{X_{t+1}} \left[v_{0t+1}(X_{t+1}) + \underbrace{\mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p(X_{t+1})) + \epsilon_{1t+1}\}}_{-\log(p_0(X_{t+1})) + c} \right] f_1(X_{t+1}|X_t)$$

With T1EV errors:

$$\mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p(X_{t+1})) + \epsilon_{1t+1}\} = -\log(p_0(X_{t+1})) + c$$

How? We showed before that $V_{t+1} = v_{jt+1} - \log(p_{jt+1}) + c$ for any j when $\epsilon \sim \text{T1EV}$

Therefore:

$$v_1(X_t) = u_1(X_t) + \beta \sum_{X_{t+1}} \left[v_{0t+1}(X_{t+1}) + \underbrace{\mathbb{E} \max\{\epsilon_{0t+1}, \psi_0^1(p(X_{t+1})) + \epsilon_{1t+1}\}}_{-\log(p_0(X_{t+1})) + c} \right] f_1(X_{t+1}|X_t)$$

Expressed as function of CCPs:

$$v_{1t}(X_t) = u_1(X_t) + \beta \sum_{X_{t+1}} \left[v_{0t+1}(X_{t+1}) - \log(p_0(X_{t+1})) \right] f_1(X_{t+1}|X_t) + \beta c$$

Implementation steps when choice 0 is terminal:

- v_{0t+1} is linear in parameters (i.e., no V_{t+2} recursion)

Implementation steps when choice 0 is terminal:

- v_{0t+1} is linear in parameters (i.e., no V_{t+2} recursion)
- Estimate $p_0(X_{t+1})$ from data (e.g., bin estimator)

Implementation steps when choice 0 is terminal:

- v_{0t+1} is linear in parameters (i.e., no V_{t+2} recursion)
- Estimate $p_0(X_{t+1})$ from data (e.g., bin estimator)
- Similar to first-stage estimation of $f_j(X_{t+1}|X_t)$

Implementation steps when choice 0 is terminal:

- v_{0t+1} is linear in parameters (i.e., no V_{t+2} recursion)
- Estimate $p_0(X_{t+1})$ from data (e.g., bin estimator)
- Similar to first-stage estimation of $f_j(X_{t+1}|X_t)$
- Problem reduces to logit with adjustment term for $\log(p_{0t+1})$

Implementation steps when choice 0 is terminal:

- v_{0t+1} is linear in parameters (i.e., no V_{t+2} recursion)
- Estimate $p_0(X_{t+1})$ from data (e.g., bin estimator)
- Similar to first-stage estimation of $f_j(X_{t+1}|X_t)$
- Problem reduces to logit with adjustment term for $\log(p_{0t+1})$
 - Stage 1: estimate p 's and f 's via bin estimator

Implementation steps when choice 0 is terminal:

- v_{0t+1} is linear in parameters (i.e., no V_{t+2} recursion)
- Estimate $p_0(X_{t+1})$ from data (e.g., bin estimator)
- Similar to first-stage estimation of $f_j(X_{t+1}|X_t)$
- Problem reduces to logit with adjustment term for $\log(p_{0t+1})$
 - Stage 1: estimate p 's and f 's via bin estimator
 - Stage 2: estimate α 's via logit with constructed adjustment term

Hotz-Miller application: Couples' fertility and sterilization decisions

Hotz-Miller application: Couples' fertility and sterilization decisions

- State: H_t = number of children, wife's age, wife's education

Hotz-Miller application: Couples' fertility and sterilization decisions

- State: H_t = number of children, wife's age, wife's education
- Choices: $j = 0$ (sterilize), $j = 1$ (no sterilization)

Hotz-Miller application: Couples' fertility and sterilization decisions

- State: H_t = number of children, wife's age, wife's education
- Choices: $j = 0$ (sterilize), $j = 1$ (no sterilization)
- Sterilization is **terminal**—no further fertility decisions

Hotz-Miller application: Couples' fertility and sterilization decisions

- State: H_t = number of children, wife's age, wife's education
- Choices: $j = 0$ (sterilize), $j = 1$ (no sterilization)
- Sterilization is **terminal**—no further fertility decisions
- If no sterilization: probability α of birth next period

Hotz-Miller application: Couples' fertility and sterilization decisions

- State: H_t = number of children, wife's age, wife's education
- Choices: $j = 0$ (sterilize), $j = 1$ (no sterilization)
- Sterilization is **terminal**—no further fertility decisions
- If no sterilization: probability α of birth next period

e.g. $f_1(H_{t+1}|H_t) = \alpha \cdot 1[H_{t+1} = H_t + 1] + (1 - \alpha) \cdot 1[(H_{t+1} = H_t)]$