Review of Economic Studies (1993) 60, 497-529 0034-6527/93/00260497$02.00

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 (1989*a*), 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 (1989*b*) for a recent survey of this fast growing field.

Key insight: Differenced $v_j - v_{j'}$ can be mapped into CCPs (p_j's)

Key insight: Differenced $v_j - v_{j'}$ can be mapped into CCPs (p_j's)

Method: Extract p_j's from data in first stage, avoid solving backwards recursion

Key insight: Differenced $v_j - v_{j'}$ can be mapped into CCPs (p_j's)

Method: Extract p_j's from data in first stage, avoid solving backwards recursion

Application: Couples' fertility decisions (optimal stopping problem)

Consider an individual with two choices and Type 1 Extreme Value errors

Consider an individual with two choices and Type 1 Extreme Value errors

Probability of choice 1:

$$p_1 = \frac{\exp(v_1)}{\exp(v_0) + \exp(v_1)}$$

Consider an individual with two choices and Type 1 Extreme Value errors

Probability of choice 1:

$$p_1 = \frac{\exp(v_1)}{\exp(v_0) + \exp(v_1)}$$

Ratio of probabilities:

$$\frac{p_1}{p_0} = \frac{\exp(v_1)}{\exp(v_0)} = \exp(v_1 - v_0)$$

Consider an individual with two choices and Type 1 Extreme Value errors

Probability of choice 1:

$$p_1 = \frac{\exp(v_1)}{\exp(v_0) + \exp(v_1)}$$

Ratio of probabilities:

$$\frac{p_1}{p_0} = \frac{\exp(v_1)}{\exp(v_0)} = \exp(v_1 - v_0)$$

Taking logs:

$$\log\left(\frac{p_1}{p_0}\right) = \log(p_1) - \log(p_0) = v_1 - v_0$$

Consider an individual with two choices and Type 1 Extreme Value errors

Probability of choice 1:

$$p_1 = \frac{\exp(v_1)}{\exp(v_0) + \exp(v_1)}$$

Ratio of probabilities:

$$\frac{p_1}{p_0} = \frac{\exp(v_1)}{\exp(v_0)} = \exp(v_1 - v_0)$$

Taking logs:

$$\log\left(\frac{p_1}{p_0}\right) = \log\left(p_1\right) - \log\left(p_0\right) = v_1 - v_0$$

Works for any two alternatives!

Let's look at it equivalently from the standpoint of V_{t+1}, assuming ϵ's are T1EV:

Let's look at it equivalently from the standpoint of V_{t+1}, assuming ϵ's are T1EV:

$$V_{t+1} = \mathbb{E} \max_{j \in \mathcal{J}} \left\{ v_{jt+1} + \epsilon_{jt+1} \right\}$$

Let's look at it equivalently from the standpoint of V_{t+1}, assuming ϵ's are T1EV:

$$V_{t+1} = \mathbb{E} \max_{j \in \mathcal{J}} \{ v_{jt+1} + \epsilon_{jt+1} \}$$

$$= \log \left(\sum_k \exp \left(v_{kt+1} \right) \right) + c$$

Let's look at it equivalently from the standpoint of V_{t+1}, assuming ϵ's are T1EV:

$$V_{t+1} = \mathbb{E} \max_{j \in \mathcal{J}} \{ v_{jt+1} + \epsilon_{jt+1} \}$$

$$= \log \left(\sum_k \exp \left(v_{kt+1} \right) \right) + c$$

Can do some algebra to show that $V_{t+1} = v_{jt+1} - \log \left(p_{jt+1} \right) + c$ for any alternative j

Key derivation trick: multiply and divide inside the log sum by $\exp(v_j)$

$$V_{t+1} = \log\left(\sum_k \exp(v_{kt+1})\right) + c$$

Key derivation trick: multiply and divide inside the log sum by $\exp(v_j)$

$$V_{t+1} = \log\left(\frac{\exp(v_{jt+1})}{\exp(v_{jt+1})}\sum_k \exp(v_{kt+1})\right) + c$$

Key derivation trick: multiply and divide inside the log sum by $\exp(v_j)$

$$V_{t+1} = \log\left(\exp(v_{jt+1}) \cdot \frac{\sum_k \exp(v_{kt+1})}{\exp(v_{jt+1})}\right) + c$$

Key derivation trick: multiply and divide inside the log sum by $\exp(v_j)$

$$V_{t+1} = \underbrace{\log \exp(v_{jt+1})}_{v_{jt+1}} + \log\left(\underbrace{\frac{\sum_k \exp(v_{kt+1})}{\exp(v_{jt+1})}}_{\left(\frac{1}{p_{jt+1}}\right)}\right) + c$$

Key derivation trick: multiply and divide inside the log sum by $\exp(v_j)$

$$V_{t+1} = v_{jt+1} + \log\left(\frac{1}{p_{jt+1}}\right) + c$$

Key derivation trick: multiply and divide inside the log sum by $\exp(v_j)$

$$V_{t+1} = v_{jt+1} - \log\left(p_{jt+1}\right) + c \qquad \forall\, j \in \mathcal{J}$$

For T1EV, we showed earlier:

$$v_j - v_0 = \log(p_j) - \log(p_0)$$

For T1EV, we showed earlier:

$$v_j - v_0 = \log(p_j) - \log(p_0)$$

Define the mapping:

$$\psi_0^j(p) = \log(p_j) - \log(p_0)$$

For T1EV, we showed earlier:

$$v_j - v_0 = \log(p_j) - \log(p_0)$$

Define the mapping:

$$\psi_0^j(p) = \log(p_j) - \log(p_0)$$

This is our inversion formula for the logit case

For T1EV, we showed earlier:

$$v_j - v_0 = \log(p_j) - \log(p_0)$$

Define the mapping:

$$\psi_0^j(p) = \log(p_j) - \log(p_0)$$

This is our inversion formula for the logit case

- Regular (no ψ): solve for $v_j - v_0 \to$ compute $p_j \to$ estimate parameters

For T1EV, we showed earlier:

$$v_j - v_0 = \log(p_j) - \log(p_0)$$

Define the mapping:

$$\psi_0^j(p) = \log(p_j) - \log(p_0)$$

This is our inversion formula for the logit case

- Regular (no ψ): solve for $v_j - v_0 \to$ compute $p_j \to$ estimate parameters
- Inversion (using ψ): observe $p_j \to$ recover $v_j - v_0 \to$ estimate parameters

Hotz-Miller inversion proposition:

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

For sake of example, let's consider $j = 0$

Hotz-Miller inversion proposition:

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

For sake of example, let's consider $j = 0$

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

Hotz-Miller inversion proposition:

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

For sake of example, let's consider $j = 0$

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

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

Hotz-Miller inversion proposition:

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

For sake of example, let's consider $j = 0$

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

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

Hotz-Miller inversion proposition:

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

For sake of example, let's consider $j = 0$

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

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

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

Hotz-Miller inversion proposition:

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

For sake of example, let's consider $j = 0$

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

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

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

Hotz-Miller inversion proposition:

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

For sake of example, let's consider $j = 0$

$$
\begin{aligned}
V_{t+1} &= \mathbb{E}\max\{v_{0t+1} + \epsilon_{0t+1}, v_{1t+1} + \epsilon_{1t+1}, \ldots, v_{Jt+1} + \epsilon_{Jt+1}\} \\
&= v_{0t+1} + \mathbb{E}\max\{\epsilon_{0t+1}, v_{1t+1} - v_{0t+1} + \epsilon_{1t+1}, \ldots, v_{Jt+1} - v_{0t+1} + \epsilon_{Jt+1}\} \\
&= v_{0t+1} + \mathbb{E}\max\{\epsilon_{0t+1}, \psi_0^1(p_{t+1}) + \epsilon_{1t+1}, \ldots, \psi_0^J(p_{t+1}) + \epsilon_{Jt+1}\}
\end{aligned}
$$

This mapping $\psi(\cdot)$ works for *any* distribution of ϵ!