The Specification and Estimation of Dynamic Stochastic Discrete Choice Models

A Survey

Zvi Eckstein
Kenneth I. Wolpin

I. Introduction

This paper is a survey of a rapidly growing literature on methods for solving and estimating dynamic stochastic discrete choice models. The major characteristics of this literature are: (1) the optimization problem is forward looking and contains some stochastic elements; (2) the choice set only takes on discrete values; (3) the approach is structural, that is the parameters to be estimated come from objective functions (tastes and/or technology) and constraints; and (4) the econometrics is closely connected to the theory, and specifically, the error structure is motivated in the problem as an integral part of the optimization.

The importance and usefulness of discrete choice models is now well established and such models are routinely applied. While it has always seemed that dynamic optimization under uncertainty would provide a better description (and prediction) of behavior, until recently methods did not exist to consistently (structurally) incorporate dynamic elements into discrete choice models. It is this task that the new literature addresses.

The major tool for these new methods is dynamic programming (Bellman 1957). Dynamic programming is a recursive solution method for optimization problems which have a dynamic structure, and can be ap-

Zvi Eckstein is professor of economics at Tel-Aviv University and Boston University. Kenneth I. Wolpin is a professor of economics at the University of Minnesota. *The authors would like to thank Fatum Gonul, Hide Ichimura, Chuck Manski, Bob Miller, Ariel Pakes, John Rust, Steven Stern, and the editors for their useful comments on an earlier draft of this survey.*

plied with a discrete, continuous or mixed discrete-continuous choice set. The major insight of dynamic programming is that the solution of a multidimensional problem can be reduced to a recursive solution of a sequence of two period problems.

To illustrate the idea of dynamic programming and, more importantly, to illustrate the contribution of the literature we survey, consider the very simple problem of when to consume an indivisible good given an endowment of the good. Let $d(t) = 1$ if the good is consumed at time t, $d(t) = 0$ otherwise, and $R(d(1), \ldots, d(T))$ be the reward or utility from consuming the good at the T possible times. The direct solution of the problem would require that we make T comparisons of utility, i.e., $R(1, 0, \ldots, 0)$, $R(0, 1, 0, \ldots, 0), \ldots, R(0, 0, \ldots, 1, 0)$, $R(0, 0, \ldots, 0, 1)$. However, suppose we proceed as follows. If by the beginning of period T the good has not been consumed the only possibility is to consume it at T, so let $V(T) = R(0, 0, \ldots, 0, 1)$ which is the maximal lifetime utility at T given the state at T. Suppose now that we have reached $T - 1$ without consuming the good. If we consume it at $T - 1$ then utility is $R(0, 0, \ldots, 1, 0) = R(T - 1)$, while if we do not then utility is $R(0, 0, \ldots, 0, 1) = V(T)$. Define maximal lifetime utility at $T - 1$ as $V(T - 1) = \max(R(T - 1), V(T))$, which is the maximum of utility if we consume the good at $T - 1$ and maximal utility if we do not. Similarly if we have reached $T - 2$ without consuming the good, then if we don't consume it at $T - 2$ we obtain $V(T - 1)$, while if we do, we obtain $R(0, 0, \ldots, 1, 0, 0) = R(T - 2)$. Thus our maximal utility at $T - 2$ is $V(T - 2) = \max(R(T - 2), V(T - 1))$. In general, then we can continue to work backwards recursively making pairwise comparisons, with maximal utility at any $t < T$ obeying the recursive equations $V(t) = \max(R(t), V(t + 1))$. The dimensionality of the problem has been reduced from a T-wise comparison to $T - 1$ pairwise comparisons.

This example illustrates a very simple optimal stopping problem. The dynamics arise because the decision at time t explicitly depends on prior decisions, namely we can only consume the good now if we did not previously consume it. The individual decides whether or not to consume the good at t, given that it has not been consumed up to t, according to whether $R(t) \geq V(t + 1)$. One might think of $V(t + 1)$ as the reservation reward or utility at $t + 1$, i.e., the reward if one postpones consumption of the good.

The above model predicts that if all individuals have the same reward function, they will all make the same decision. Suppose we have data on when a sample of observably homogeneous individuals each consumed the good. It is quite likely that the data will reveal variation in the timing of consumption, in which case the model would clearly be rejected. One way to allow for heterogeneity in observed behavior is to assume that the reward function is stochastic, namely given by $R(t) + \epsilon(t)$, and to assume that individuals draw at each t from the density function for $\epsilon(t)$.[1] The analogue of the lifetime utility function under uncertainty is $V(t) = \max(R(t) + \epsilon(t), E(V(t+1)|\Omega(t)))$ where $\Omega(t)$ is the information set at t and $E(V(t+1)|\Omega(t))$ is the conditional expectation of $V(t+1)$ taken over all future ϵ's. Now, under the further simplifying assumption that the $\epsilon(t)$'s are i.i.d. over time, the model predicts that the proportion of individuals who choose to consume the good at t (given the good has not yet been consumed prior to t) is given by the probability that each of their draws of $\epsilon(t)$ exceeded $E(V(t+1)|\Omega(t)) - R(t)$. The assumption that $\epsilon(t)$ is known to the individual but not to us makes the observed decision random from our perspective. A comparison of these predicted probabilities over time to the corresponding observed individual choices or sample proportions forms the basis for the estimation of the reward function. It is this extension of the use of dynamic programming to estimation that is the contribution of this new literature.

We begin the paper in Section II with a general specification of a dynamic stochastic discrete choice model that encompasses all existing studies in this literature. We show how the general framework nests, in order of presentation, a job search model (Wolpin 1987), a patent renewal model (Pakes 1986), an engine replacement model (Rust 1987), an armed services retention model (Gotz and McCall 1987), a labor force participation model (Eckstein and Wolpin 1986, Gonul 1987), a retirement model (Berkovec and Stern 1987, Rust 1987), a fertility model (Wolpin 1984, Montgomery 1988, Hotz and Miller 1989), and a job matching model (Miller 1984).[2] Section III discusses alternative solution methods for these types of dynamic programming problems. We explicitly demonstrate solution methods for several of the preceding models. In Section IV we explain the maximum likelihood method of estimation as it is applied to these models and briefly touch on an alternative estimation method, the method of simulated moments. Section V briefly surveys two semi-reduced form approaches to estimating the structural parameters recently advanced in Hotz and Miller (1988) and in Manski (1988 and 1990 forthcoming). We also suggest several alternative approximation methods. A standard search model is used as an example in each of the sections, so that the reader can follow a familiar single model from its initial specification through its estimation. The final section summarizes the paper and concludes with a few additional topics in labor economics that can be explored using the methods discussed here.

1. The additivity of the disturbance in the current return function is not necessary.

2. We begin with the search model because it is the example we use to illustrate the solution and estimation methods throughout the paper. The rest of the ordering is for pedagogical reasons, in part due to the complexity of the models, and does not indicate the relative importance of the papers.

II. A General Model

We consider a general model of I discrete choices over T discrete periods of time, where T is either finite or infinite. In each period an individual chooses one of the I possible alternatives (e.g., employment status, occupation, etc.), where the indicator $d_i(t) = 1$ if alternative i is chosen at time t and $d_i(t) = 0$ otherwise, that is, if alternative i is not chosen by the agent. Alternatives are mutually exclusive, i.e., $\Sigma d_i(t) = 1$.[3] The objective of the individual at any time t, $t = 0, 1, \ldots, T$, is to maximize

$$(1)\quad E\left[\sum_{j=t}^{T} \beta^{j-t} \sum_{i \in I} R_i(j) d_i(j) | \Omega(t)\right]$$

where $0 < \beta < 1$ is the individual's discount rate, $E(\cdot)$ is the mathematical expectations operator, $\Omega(t)$ is the individual's information set at time t which includes all past and current realizations of the variables that directly or indirectly affect the value of Equation (1) and $R_i(t)$ is a random variable representing the individual's reward if alternative i is chosen at time t. $R_i(s)$ for $s \le t$ obviously belongs to the individual's information set at time t, $\Omega(t)$.

Maximization of (1) is accomplished by choice of the optimal sequence of control variables $\{d_i(t)\}_{i \in I}$ for $t = 0, 1, \ldots, T$ which are functions of information that is available when the decision is made. Define the maximal expected value of the reward at time t

$$(2)\quad V(\Omega(t)) = \sup_{\{d_i(t)\}_{i \in I}} E\left[\sum_{j=t}^{T} \beta^{j-t} R(j) | \Omega(t)\right]$$

where $R(j) = \Sigma_{i \in I} R_i(j) d_i(j)$ is the actual reward at time j. The function V depends only on the information set at time t, and obeys the dynamic programming equation[4]

$$(3)\quad V(\Omega(t)) = \max_{i \in I} \{L_i V(\Omega(t))\}$$

where L_i is the alternative-specific operator defined by

$$(4)\quad L_i V(\Omega(t)) = R_i(t) + \beta E[V(\Omega(t+1)) | d_i(t) = 1], \qquad t = 0, \ldots, T.$$

3. Alternatives can always be redefined to satisfy this assumption.

4. The sup operator is equal to the max operator if the maximum exists.

The dynamics of the problem are due to the dependence of the function V at time $t + 1$ on the choice of d_i, $i = 1, \ldots, I$, at time t and possibly before.

This general discrete choice problem describes all of the different structurally estimated dynamic discrete choice models found in the literature.[5] We now turn to a description of these models.

A. *Optimal Stopping Models*

Optimal stopping models are special cases of discrete choice dynamic programming models. We shall now show how the models of optimal stopping that exist in the literature fit the above general model. The most familiar optimal stopping model in economics is that of job search (Lippman and McCall 1976, Mortensen 1970).

1. A Job Search Model (Wolpin 1987)

Wolpin (1987) structurally estimated a standard two-state job search model in which a wealth-maximizing individual faces a known wage offer distribution, a constant cost of search, and a known per-period probability of receiving a wage offer. In each period of a finite horizon the individual decides whether to accept an offer if one is received. If an offer is rejected or if none is received the individual continues to search. Rejected offers cannot be recalled.

To place the search model in the framework of the previous section, note first that there are two alternatives, $I = 2$. Let $d_1(t) = 1$ if the individual is not employed (searching) and $d_1(t) = 0$ otherwise, and $d_2(t) = 1$ if the individual is employed and $d_2(t) = 0$ otherwise.[6] The reward function is given by

(5a) $R_1(t) = b$,

(5b) $$R_2(t) = \begin{cases} w(t) & \text{if } d_1(t-1) = 1 \\ R_2(t-1) & \text{if } d_2(t-1) = 1 \end{cases}$$

where b is net income if the individual searches (unemployment compensation minus search costs) and $w(t)$ is the wage at time t. If the individual is not employed at t, a wage offer is drawn with a known probability p

5. A formal treatment of this type of problem exists, for example, in Whittle (1983), as well as in other texts on dynamic stochastic programming.
6. In the case of $I = 2$ (two alternatives) it is more convenient to use one indicator because $d_1(t) + d_2(t) = 1$. But, for the general case ($I > 2$) the general notation is less cumbersome and so we employ it throughout the paper.

from a time-independent distribution function $F(w(t))$ which is also known.[7] With probability $1 - p$ no wage offer is received and net income is equal to b.

The search model described above is an example of an optimal stopping problem, i.e., once an offer is accepted the individual will never return to the nonemployment state. The stopping property results from the assumption in Equation (5b) that no new wage offers are received after a job is accepted as indicated by the fact that the reward at t is the previous period wage if the individual is employed at $t - 1$.[8] Properties of the solution both in the finite and infinite horizon cases are presented in Lippman and McCall (1976) and the economics literature contains numerous extensions of the basic model.

2. Patent Renewal (Pakes 1986)

Pakes formulated and estimated an optimal stopping model of patent renewal. Each period over a finite horizon a patentee must decide whether or not to pay an annual renewal fee in order to keep the patent in force. If the renewal fee is not paid then the patent is permanently canceled.

In terms of the general model there are two choice variables ($I = 2$), and $d_2(t) =$ $d_1(t) = 1$ if the patent is not renewed and $d_1(t) = 0$ otherwise, 1 if the patent is renewed and $d_2(t) = 0$ otherwise. The reward for each choice is given by

(6a) $R_1(t) = 0$,

(6b) $$R_2(t) = \begin{cases} \max[\delta R_2(t-1), z(t)] - c(t) & \text{with probability } 1 - e^{-\theta R_2(t-1)} \\ 0 & \text{with probability } e^{-\theta R_2(t-1)}; \end{cases}$$

$c(t)$ is the renewal cost of the patent at time t, $\theta > 0$ and $0 < \delta < 1$ are known parameters, $z(t)$ is a random variable with density function $g(z, t) = \sigma(t)^{-1} \exp[-(\gamma + z/\sigma(t))]$ and $\sigma(t) = \phi^{t-1}\sigma$, where σ, γ and ϕ are known parameters. In addition $R_2(1)$ is assumed lognormal with mean μ and standard deviation σ_R. Pakes allows this initial draw to be different for each firm. The idea is that with some probability which depends on the previous period return, $e^{-\theta R_2(t-1)}$, the patent is determined to have no value, while with probability $1 - e^{-\theta R_2(t-1)}$ the patent either depreciates in value or a new use is found and the patent value is $z(t)$.

7. Wolpin actually assumes that p is duration dependent in a known way, but for expositional purposes we ignore this minor complication.
8. For an example of a model in which offers are received while on the job, see Burdett (1978). See Flinn and Heckman (1982) for a discussion of search models in a continuous time setting.