Public-Policy Uses of Discrete-Choice Dynamic Programming Models

By KENNETH I. WOLPIN *

The quantitative assessment of proposals for altering existing public policies or for initiating new ones is most often conducted either by extrapolating from existing policies or by conducting and evaluating experimental pilot projects. The former approach relies on the existence of current or past policy variation that encompasses policies that are "close" to the proposed policy. Often, this condition is not met, sometimes because the policy proposal involves a large change, one outside the observed variation, and other times because the policy change is along a dimension that has never varied. Similarly, the use of experiments relies on the experimental setting (including policy variations built into the experimental design) being "close" to the policy that is actually adopted.

An example will help to clarify the point. In the recent debate about the impact of the minimum wage on teenage employment, it is argued by those who advocate increasing the minimum wage that, based on available estimates, disemployment effects of proposed policy changes would be small. It is generally accepted, however, that a large change, raising the hourly minimum wage to, say, $20 would, if enforced, have a devastating impact on teenage employment, well beyond the effect extrapolated from many of the estimates based on observed policy variation. Although few would seriously entertain raising the minimum wage to $20, a similar problem can be expected to arise in the evaluation of current policy proposals that would substantially alter the generosity of social programs such as Medicare, Medicaid, or AFDC, or that would place restrictions on eligibility that have not existed before.

All economic and social programs were at one time untried and there is always an abundance of current policy proposals that have no obvious analogues in existing policy regimes. Moreover, most new policies are not subjected to experimental trials, and those that are almost universally lead to debatable inferences (for example, see the literature on the evaluation of job-training programs such as the Job Training Partnership Act [JTPA]) due to perceived flaws in design and implementation.

The question of how to assess quantitatively the behavioral impact of a new policy or of a large policy change is certainly not a new one. Jacob Marschak (1952 p. 26) for example, makes the following argument: "if among the policies considered there are some that involve structural changes, then the choice of the policy best calculated to achieve given ends presupposes knowledge of the structure that has prevailed before." He defines structure as "a set of conditions which did not change while observations were made but which might change in the future" and the conditions that constitute the structure as "... (1) a set of relations describing human behavior and institutions as well as technological laws and involving, in general, nonobservable random disturbances and nonobservable random errors of measurement; (2) the joint probability distribution of these random quantities." Knowledge of structure thus supplies a third method for the assessment of policy proposals.

Marschak's definition of structure describes standard practice in economic modeling augmented by assumptions about factors that are not observable to us. The question is one of implementation, namely, what structure should be specified and estimated. Clearly, this is a matter of judgment, data availability, and estimation feasibility, and it is not my intention to contrast alternative structural modeling strategies (e.g., static vs. dynamic, imperfect

* Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104. I have received useful comments from Christopher Flinn, Michael Keane, and Mark Rosenzweig. Support is gratefully acknowledged from NSF grant SBR95-96246 and NIH grant HD-30516.

vs. perfect foresight, etc.). Instead, I will concentrate on illustrating the public-policy uses of a single class of models.

Over the last decade, there has emerged a growing literature on techniques for estimating behavioral models that involve discrete (or discretized) choices and that have dynamic programming structures (see Zvi Eckstein and Wolpin [1989] or John Rust [1994] for surveys). Considerable progress has been made in implementing behavioral models that incorporate ever more complex dynamic interactions. The state of the art is quite advanced, and I would argue that these efforts should be taken seriously as tools for the assessment of public-policy initiatives, at least as seriously as feasible alternative approaches.

I. The General Stochastic Dynamic Discrete-Choice Model

The types of models that have been estimated have a common structure. Specifically, an agent is assumed to decide among K possible alternatives in each of T (finite or infinite) discrete periods of time. Alternatives are defined as mutually exclusive so that if $d_k(t) = 1$ indicates that alternative k is chosen at time t and $d_k(t) = 0$ indicates otherwise, then $\Sigma_{k=1}^{K} d_k(t) = 1$. Associated with each choice at time t is a current-period reward (e.g., income, utility, profits), $R(t)$, that is known to the agent at time t but is random from the perspective of periods prior to t.

The objective of the agent at any time t is to maximize the expected discounted value of rewards by choice of the optimal sequence of control variables $\{d_k(t)\}_{k \in \mathcal{K}}$ where $\mathcal{K}$ is the set of possible alternatives. Letting $\delta > 0$ be the discount factor, $E(\cdot)$ the expectations operator, and $S(t)$ the state space at time t, the maximal expected value of discounted lifetime rewards at time t is

$$(1) \quad V(S(t), t) = \max_{d_k(t)} E\left[\sum_{\tau=t}^{T} \delta^{\tau-t} \sum_{k=1}^{K} R_k(\tau) d_k(\tau) \mid S(t) \right].$$

The state space includes all factors known to the agent that affect current rewards, including government policies and the probability distribution of future rewards. The value function, $V(S(t), t)$, can also be represented as the maximum over alternative-specific value functions (expected discounted lifetime rewards), $V_k(S(t), t)$, that satisfy the Bellman equation, namely,

$$(2) \quad V(S(t), t) = \max_{k \in \mathcal{K}} \{V_k(S(t), t)\}$$

and

$$(3) \quad V_k(S(t), t) = R_k(S_k(t), t) + \delta E[V(S(t+1), t+1) \mid S(t), d_k(t) = 1]$$

where $S_k(t)$ is the subset of the state space that affects the alternative-specific current-period reward.[1] If the horizon is finite, then (3) holds for $t < T$ and $V_k(S(T), T) = R_k(S_k(T), T)$. The representation in (3) assumes that future choices are made optimally. This assumption of "rationality" does not imply that all agents forecast exogenous, forcing variables (e.g., future policy regimes) in the same way or that those expectations necessarily conform to a rational-expectations equilibrium (see e.g., Charles F. Manski, 1988).

The expectation in (3) is taken over the distribution of $S(t+1)$ conditional on $S(t)$ and $d_k(t) = 1$. The randomness in rewards arises from the existence of state variables at time $t + 1$ observable to agents at $t + 1$, but not observable at t or before. The random elements in the state space in general may be contemporaneously or serially correlated. In principle, the state space may include the entire history of stochastic realizations for the agent. The deterministic or predetermined elements are represented by equations of motion.

II. Estimation

Solving and estimating the model is computationally demanding. The solution involves computing the multiple integral that forms the expectation in (3) at every state-space ele-

[1] $S_k(t)$ may span the entire set $S(t)$.

ment.[2] The solution serves as input into estimating the parameters of the model (those in the reward function and in the joint distribution of the random elements of the state space). To understand the connection between solution and estimation, consider having data on a homogeneous sample of agents. For each agent one observes the sequence of choices made over some part of the decision horizon, not necessarily beginning with the first decision period, and the sequence of a subset of the current-period rewards.

For the agent, the current decision is deterministic, namely, choose the alternative with the maximal alternative-specific value function [the value of k for which $V_k(S(t), t) = V(S(t), t)$]. For the researcher, because not all of the current rewards are observed (or, are observed without error) the current decision is probabilistic. The probability of observing any particular sequence of decisions and rewards can easily be written; it clearly depends on the parameters of the model. The sample likelihood can be formed and maximum-likelihood estimates obtained. Depending on the assumption that was made in formulating the optimization problem with respect to the joint distribution of the random components of the rewards (and additional assumptions about measurement errors), computing the likelihood can involve high-dimensional integrations as in any multinomial discrete-choice model. In this case, however, each parameter iteration involves resolving the dynamic optimization problem.

The generality of this framework is an important strength, allowing researchers to contemplate the estimation of behavioral models with a wide variety of structures and encompassing a wide variety of issues. The burden placed on the researcher is to devise useful models that are computationally, rather than analytically, tractable. The researcher has a number of dimensions over which tractability can be manipulated. They are (i) the size of the choice set, (ii) the size of the state space, (iii) the functional form of the reward functions, and (iv) the joint distribution of unobservables. Because these models can be analytically solved only under restrictive assumptions, within a numerical solution paradigm the form of the utility function and the placement of the stochastic elements and assumptions about their joint distribution can be quite general (subject to overall computational tractability).[3] Moreover, it is straightforward, subject to computational feasibility and identification, to allow for observable exogenous variables or for permanent unobservable heterogeneity.

III. Recent Applications

Most recently, this framework has served as the basis for addressing issues in health policy (Donna Gilleskie, 1995), welfare policy (Seth Sanders, 1993; Anne Kerttula, 1995; Christopher A. Swann, 1995), educational policy (Michael Keane and Wolpin, 1994; Todd R. Stinebrickner, 1995), employment policy (Christopher Ferrall, 1994; Silvio Rendon, 1995), environmental policy (Geoffrey Rothwell and Rust, 1995), agricultural development policy (Mark R. Rosenzweig and Wolpin, 1993), and industrial policy (Yuval Nachton, 1995). Given space limitations, I will concentrate attention on only three of these applications and, even then, only in brief. My intention is to provide a flavor of their richness and to demonstrate their applicability to policy questions for which knowledge of structure on a priori grounds would appear to be most useful if not essential.

A. *Employment Policy*

Ferrall (1994) provides a comprehensive structural analysis of the Canadian unemployment

[2] The number of contemporaneous random elements in the state space determines the dimension of the multiple integration. The model could of course be solved without any stochastic state-space elements (no uncertainty). The choice of the number is conditioned on estimation issues as discussed in what follows.

[3] Numerical solution requires that continuous state variables be discretized. This is the major complication in allowing for general serial correlation in shocks in terms of solving the optimization problem. Of course, because distributional assumptions made in the optimization problem carry over to estimation, the computational burden of introducing serial correlation is greater than simply having continuous observable state variables.

insurance (UI) program that incorporates all of the important UI program features: (i) a waiting period before the receipt of benefits; (ii) the dependence of the benefit level on the previous wage; (iii) an eligibility requirement based on prior employment; and (iv) the dependence of length of eligibility on the length of time spent in the previous employment spell. A critical additional aspect of the Canadian system in terms of evaluating alternative programmatic changes is that the program is uniformly administered throughout Canada.[4]

Policymakers seem to be concerned about the length of time it takes individuals to transit into employment after leaving school, and policies have been contemplated or initiated to "ease" the transition. According to the data presented by Ferrall (1994), in Canada the mean duration to employment upon leaving school varies from approximately four months for those with some postsecondary education to about six months for those without. Of course, these unemployment spells are generally not insured, and so it would seem that this transition is unconnected to UI policy. However, a forward-looking individual would take into account that the value of accepting employment includes not only the current pecuniary reward, but also the (expected) value of future UI eligibility. Thus, the existence of the UI system may in fact be masking an even longer underlying school-to-work transition process.[5]

Ferrall (1994) specifies a structural job-search model in which agents decide on acceptance criteria for both insured and uninsured unemployment spells, intermediated by employment spells determined by exogenous layoff rates. Given estimates of the model's parameters (job offer rates, layoff rates, the discount factor, and the wage-offer distribution), Ferrall finds, by simulating the most extreme counterfactual, that the length of the school-to-work transition is essentially unaffected by the existence of the UI program. However, this result is not due to some inherent feature of job-search behavior. Ferrall also simulates the impact of reducing the current 20-week work requirement for UI eligibility to four weeks. This experiment indicates that the mean duration of the school-to-work transition would fall by at least two months for the less educated and by 1.5 months for the more educated group, regardless of sex.

B. *Agricultural Development Policy*

In rural India, ownership of work animals, bullocks, is required to ensure the timeliness of planting operations. However, in Rosenzweig and Wolpin (1993), a high incidence of turnover was found in bullock ownership, along with considerable underinvestment. A consistent explanation is that Indian farmers face significant borrowing constraints and cannot both maintain productive assets and smooth consumption levels. Among policy interventions that have been proposed to ameliorate such production inefficiencies are the provision of weather insurance and the creation of opportunities for alternative and assured income flows. A quantitative cost–benefit assessment and comparison of these "new" policies requires knowledge of the behavioral response of farmers to their introduction.

The structural model estimated in Rosenzweig and Wolpin (1993) incorporates the major features of low-income agricultural environments: income uncertainty, constraints on borrowing, incomplete rental markets, returns to farmer experience, and the use of productive assets to smooth consumption. In each period, farmers maximize their expected present value of remaining lifetime utility by choosing how many bullocks to buy and sell, whether or not to breed a bullock (bullocks die according to an exogenous process), and whether or not to buy a water pump (which cannot be resold). Profits are subject to village-level weather shocks as well as farmer-specific shocks.

The estimates of the model imply that the accumulation or maintenance of bullocks is seriously impeded by the presence of weather shocks when there are borrowing constraints. Even farmers holding initially two bullocks

[4] The only program parameters that vary are those related to the length of eligibility, and they depend only on regional variation in the unemployment rate.

[5] I assume that UI benefits do not increase so rapidly with the prior wage that the UI system induces individuals to search longer after leaving school.

(the profit-maximizing number) will optimally sell them off when faced with an adverse weather shock in order to smooth consumption. It might appear, therefore, that the provision of weather insurance could lead to increased holdings of bullocks and welfare gains to farmers. However, while weather insurance does yield a welfare gain (farmers are risk-averse), discounted expected utility is no higher when farmers pay actuarially fair insurance premiums than it is without insurance, and average bullock stocks are unaffected. Farmers are already partly insured via a consumption floor maintained through informal transfer arrangements, and weather shocks are only a part of the uncertainty. On the other hand, it is estimated that an assured supplement to their income of approximately 30 percent, say, through the creation of nonagricultural earnings opportunities, would lead farmers to hold within 10 percent of the efficient bullocks stock. The estimates indicate that credit-constrained Indian farmers are too poor to be efficient.

C. *Welfare Policy*

The current legislative perspective adopted with respect to reforming the U.S. welfare system is President Bill Clinton's vow "to change welfare as we know it." Along with proposals to reduce benefit levels, some of the more radical proposals include establishing limits on time spent receiving welfare, eliminating welfare for women under the age of 18, and eliminating welfare for second and higher-order births. The literature on assessing the impact of welfare programs on behavior is extensive (see Robert Moffitt [1992] for a comprehensive survey); yet little, if any, of that literature can be used to evaluate the set of current proposals.

The availability of welfare programs, and in particular AFDC, influences choices well beyond the simple decision about welfare participation. Participation in AFDC is intimately connected with decisions about fertility, work, schooling, and marriage. Moreover, each of these decisions has ramifications for future well-being and thus for future decisions. Modeling these joint decisions, allowing for dynamics and uncertainties, within an estimable framework useful for policy analysis is a challenge.

Several recent papers (Sanders, 1993; Kerttula, 1995; Swann, 1995) illustrate the current frontier. Although their specific models differ in terms of the choice set, state space, and distributional- and functional-form assumptions, each is able to analyze and compare quantitatively several of the policy alternatives mentioned above. It is especially noteworthy that a policy restricting lifetime participation or individual spell lengths inescapably introduces dynamical considerations for the woman. To take one set of results, Swann's estimates imply that reducing AFDC benefits by 50 percent would reduce the time spent on AFDC in the aggregate by 15 percent. Further, restricting any AFDC spell length to at most three years would reduce the percentage of time spent on welfare in the aggregate by almost one-half.

V. Conclusions

Policy analysis should always be based on the best available research. All of the approaches to quantitative assessment have flaws. Some may be better suited for assessing particular types of interventions, and judgments may vary as to the weight that should be attached to particular studies within and among genres. However, informed judgments should not be based on methodological predispositions, but on evidence of performance. Ultimately, policy analysts would wish to rely on studies that forecast policy effects most accurately. Rarely is such a standard applied, and because of that it is impossible to draw general conclusions about the efficacy of different evaluation approaches. Attention should be turned to this matter.[6]

REFERENCES

Eckstein, Zvi and Wolpin, Kenneth I. "The Specification and Estimation of Dynamic

[6] Rust (1994), in the context of choosing among models of retirement behavior, and Keane and Moffitt (1995), in a static model of program participation, provide excellent examples of this kind of analysis.

Stochastic Discrete Choice Models.'' *Journal of Human Resources,* Fall 1989, *24*(4), pp. 562–98.

Ferrall, Christopher. ''Unemployment Insurance and Youth Labor Market Behavior in Canada and the United States.'' Mimeo, Queens University, 1994.

Gilleskie, Donna. ''A Dynamic Stochastic Model of Medical Care Use and Work Absence.'' Mimeo, University of North Carolina, 1995.

Keane, Michael and Moffitt, Robert. ''A Structural Model of Multiple Welfare Program Participation and Labor Supply.'' Mimeo, University of Minnesota, 1995.

Keane, Michael and Wolpin, Kenneth I. ''Career Decisions of Young Men.'' Mimeo, New York University, 1994.

Kerttula, Anne. ''A Dynamic Model of Welfare Participation.'' Mimeo, University of Pennsylvania, 1995.

Manski, Charles F. ''Path Utility Analysis of Dynamic Choice.'' Mimeo, University of Wisconsin, 1988.

Marschak, Jacob. ''Economic Measurements for Policy and Prediction,'' in W. C. Hood and T. C. Koopmans, eds., *Studies in econometric method*. New York: Wiley, 1952, pp. 1–26.

Moffitt, Robert. ''Incentive Effects of the U.S. Welfare System: A Review.'' *Journal of Economic Literature*, March 1992, *30*(1), pp. 1–61.

Nachton, Yuval. ''Optimal Policy of Investment, R&D, and Exit: An Application to Israeli High-Tech Industries.'' Mimeo, Tel Aviv University, 1995.

Rendon, Silvio. ''Job Search and Asset Accumulation Under Borrowing Constraints.'' Mimeo, New York University, 1995.

Rosenzweig, Mark R. and Wolpin, Kenneth I. ''Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India.'' *Journal of Political Economy*, April 1993, *101*(2), pp. 223–44.

Rothwell, Geoffrey and Rust, John. ''A Dynamic Programming Model of U.S. Nuclear Power Plant Operations.'' Mimeo, University of Wisconsin, 1995.

Rust, John. ''Estimation of Dynamic Structural Models, Problems and Prospects: Discrete Decision Processes,'' in C. Sims, ed., *Advances in econometrics: Sixth World Congress*, Vol. 2. Cambridge: Cambridge University Press, 1994, pp. 119–70.

Sanders, Seth. ''A Dynamic Model of Welfare and Work.'' Mimeo, Carnegie Mellon University, 1993.

Stinebrickner, Todd R. ''A Dynamic Discrete Choice Model of Teacher Attrition.'' Mimeo, University of Virginia, 1995.

Swann, Christopher A. ''A Dynamic Analysis of Marriage, Labor Force Participation, and Participation in the AFDC Program.'' Mimeo, University of Virginia, 1995.