Chapter 51

STRUCTURAL ESTIMATION OF MARKOV DECISION PROCESSES*

JOHN RUST

University of Wisconsin

Contents

*This is an abridged version of a monograph, *Stochastic Decision Processes: Theory, Computation, and Estimation* written for the Leif Johansen lectures at the University of Oslo in the fall of 1991. I am grateful for generous financial support from the Central Bank of Norway and the University of Oslo and comments from John Dagsvik, Peter Frenger and Steinar Strøm.

Handbook of Econometrics, Volume IV, Edited by R.F. Engle and D.L. McFadden

1. Introduction

Markov decision processes (MDP) provide a broad framework for modelling sequential decision making under uncertainty. MDP's have two sorts of variables: *state variables* s_t and *control variables* d_t, both of which are indexed by time $t = 0, 1, 2, 3, \ldots, T$, where the horizon T may be infinity. A decision-maker or agent can be represented by a set of *primitives* (u, p, β) where $u(s_t, d_t)$ is a utility function representing the agent's preferences at time t, $p(s_{t+1} | s_t, d_t)$ is a Markov transition probability representing the agent's subjective beliefs about uncertain future states, and $\beta \in (0, 1)$ is the rate at which the agent discounts utility in future periods. Agents are assumed to be rational: they behave according to an *optimal decision rule* $d_t = \delta(s_t)$ that solves $V_0^T(s) \equiv \max_\delta E_\delta\{\sum_{t=0}^T \beta^t u(s_t, d_t) | s_0 = s\}$ where E_δ denotes expectation with respect to the controlled stochastic process $\{s_t, d_t\}$ induced by the decision rule δ. The method of *dynamic programming* provides a constructive procedure for computing δ using the *value function* V_0^T as a "shadow price" to decentralize a complicated stochastic/multiperiod optimization problem into a sequence of simpler deterministic/static optimization problems.

MDP's have been extensively used in theoretical studies because the framework is rich enough to model most economic problems involving choices made over time and under uncertainty.[1] Applications include the pioneering work on optimal inventory policy by Arrow et al. (1951), investment under uncertainty [Lucas and Prescott (1971)] optimal intertemporal consumption/savings and portfolio selection under uncertainty [Phelps (1962), Hakansson (1970), Levhari and Srinivasan (1969), Merton (1969) and Samuelson (1969)], optimal growth under uncertainty [Brock and Mirman (1972), Leland (1974)], models of asset pricing [Lucas (1978), Brock (1982)], and models of equilibrium business cycles [Kydland and Prescott (1982), Long and Plosser (1983)]. By the early 1980's the use of MDP's had become widespread in both micro- and macroeconomic theory as well as in finance and operations research.

In addition to providing a normative theory of how rational agents "should" behave, econometricians soon realized that MDP's might provide good empirical models of how real-world decision-makers actually behave. Most data sets take the form $\{d_t^a, s_t^a\}$ where d_t^a is the decision and s_t^a is the state of an agent a at time t.[2] Reduced-form estimation methods can be viewed as uncovering agents' decision

[1] Stochastic control theory can also be used to model "learning" behavior in which agents update beliefs about unobserved stae variables and unknown parameters of the transition probabilities according to the Bayes rule.

[2] In time-series data, a is fixed at 1 and t ranges over $1, \ldots, T$. In cross-sectional data sets, T is fixed at 1 and a ranges over $1, \ldots, A$. In panel data sets, t ranges over $1, \ldots, T_a$, where T_a is the number of periods agent a is observed (possibly different for each agent) and a ranges over $1, \ldots, A$ where A is the total number of agents in the sample.

rules or, more generally, the stochastic process from which the realizations $\{d_t^a, s_t^a\}$ were "drawn", but are generally independent of any particular behavioral theory.[3] This chapter focuses on structural estimation of MDP's under the maintained hypothesis that $\{d_t^a, s_t^a\}$ is a realization of a controlled stochastic process. In addition to uncovering the form of this stochastic process (and the associated decision rule δ), structural methods attempt to uncover (estimate) the primitives (u, p, β) that generated it.

Before considering whether it is technically possible to estimate agents' preferences and beliefs, we need to consider whether this is even logically possible, i.e. whether (u, p, β) is identified. I discuss the identification problem in Section 3.5, and show that the question of identification depends on what type of data we have access to (i.e. experimental vs. non-experimental), and what kinds of a priori restrictions we are willing to impose on (u, p, β). If we only have access to non-experimental data (i.e. uncontrolled observations of agents "in the wild"), and if we are unwilling to impose any prior restrictions on (u, p, β) beyond basic measurability and regularity conditions on u and p, then it is impossible to consistently estimate (u, p, β), i.e. the class of all MDP's is non-parametrically unidentified. On the other hand, if we are willing to restrict u and p to a finite-dimensional parametric family, say $\{u = u_\theta, p = p_\theta | \theta \in \Theta \subseteq R^K\}$, then the primitives (u, p, β) are identified (generically). If we are willing to impose an even stronger prior restriction, stationarity and rational expectations (RE), then we only need parametric restrictions on u in order to identify (u, p, β) since stationarity and the RE hypothesis allow us to use non-parametric methods to consistently estimate agents' subjective beliefs from observations of their past states and decisions. Given that we are already imposing strong prior assumptions by modelling agents' behavior as an optimal decision rule to an MDP, it would be somewhat schizophrenic to be unwilling to impose any additional prior restrictions on (u, p, β). In the sequel, I assume that the econometrician is willing to bring to bear prior knowledge in the form of a parametric representation for (u, p, β). This reduces the problem of structural estimation to the technical issue of estimating a parameter vector $\theta \in \Theta$ where Θ is a compact subset of R^K.

The appropriate econometric method for estimating θ depends critically on whether the control variable d_t is continuous or discrete. If d_t can take on a continuum of possible values we say that the MDP is a *continuous decision process* (CDP), and if d_t can take on a finite or countable number of values then the MDP is a *discrete decision process* (DDP). The predominant estimation method for CDP's is generalized method of moments (GMM) using the first order conditions from the MDP problem (stochastic Euler equations) as orthogonality conditions [Hansen (1982), Hansen and Singleton (1982)]. Hansen's chapter (this volume) and Pakes's (1994) survey provide excellent introductions to the literature on structural estimation methods for CDP's.

[3] For an overview of this literature, see Billingsley (1961), Chamberlain (1984), Heckman (1981a), Lancaster (1990) and Basawa and Prakasa Rao (1980).

Thus chapter focuses on structural estimation of DDP's. DDP's are appropriate for decision problems such as whether not to quit a job [Gotz and McCall (1984)], search for a new job [Miller (1984)], have a child [Wolpin (1984)], renew a patent [Pakes (1986)], replace a bus or airplane engine [Rust (1987), Kennet (1994)] or retire a cement kiln [Das (1992)]. Although most of the early empirical applications of DDP's have been for binary decision problems, this chapter shows that most of the estimation methods extend naturally to DDP's with any finite number of possible decisions. Examples of multiple choice DDP's include Rust's (1989, 1993) model of retirement behavior where workers decide each period whether to work full-time, work part-time, or quit, and whether or not to apply for Social Security, and Miller's (1984) multi-armed-bandit model of occupation choice.

Since the control variable in a DDP model assumes at most a finite number of possible values, the optimal decision rule is determined by the solution to a system of inequalities rather than as a zero to a first order condition. As a result there is no analog of stochastic Euler equations to serve as orthogonality conditions for GMM estimation of θ as in the case of CDP's. Instead, most structural estimation methods for DDP's require explicit calculation of the optimal decision rule δ, typically via numerical methods since analytic solutions for δ are quite rare. Although we also discuss simulation estimators that rely on Monte Carlo simulations of the controlled stochastic process $\{s_t, d_t\}$ rather than on explicit numerical calculation of δ, all of these methods can be conceptualized as forms of nonlinear regression that search for an estimate $\hat{\theta}$ whose implied decision rule $d_t = \delta(s_t, \hat{\theta})$ "best fits" the data $\{d_t^a, s_t^a\}$ according to some metric. Unfortunately straightforward application of nonlinear regression methods is not possible due to three complications: (1) the "dependent variable" d_t is discrete rather than continuous; (2) the functional form of δ is generally not known a priori but rather must be derived from the solution to the stochastic control problem; (3) the "error term" ε_t in the "regression function" δ is typically multi-dimensional and enters in a non-additive, non-separable fashion: $d_t = \delta(x_t, \varepsilon_t, \theta)$.

The basic motivation for including an error term in the DDP model is to obtain a "statistically non-degenerate" econometric model. The degeneracy of DDP models without error terms is due to a basic result of MDP theory reviewed in Section 2: the optimal decision rule δ is a deterministic function of the state s_t. Section 3.1 offers several possible interpretations for the error terms in a DDP model, but argues that the most natural and internally consistent interpretation is that ε_t is an unobserved state variable. Under this interpretation, we partition the full state variable $s_t = (x_t, \varepsilon_t)$ into a subvector x_t that is observed by the econometrician, and a subvector ε_t that is observed only by the agent. If we are willing to impose two additional restrictions on u and p, namely, that ε_t enters u in an additive separable (AS) fashion and that p satisfies a conditional independence (CI) condition, we can apply a number of powerful results from the literature on estimation of static discrete choice models [McFadden (1981, 1984)] to yield estimators of θ with desirable asymptotic properties. In particular, the AS–CI assumption allows us to "integrate out" ε_t from the decision rule δ, yielding a non-degenerate system of *conditional choice probabilities* $P(d_t|x_t, \theta)$ for estimating θ by the method of maximum likelihood. Under the further restriction that $\{\varepsilon_t\}$ is an IID extreme value process we obtain a dynamic generalization of the well-known multinomial logit model,

$$P(d_t|x_t, \theta) = \frac{\exp\{v_\theta(x_t, d_t)\}}{\sum_{d' \in D(x_t)} \exp\{v_\theta(x_t, d')\}}. \tag{1.1}$$

As far as estimation is concerned, the main difference between the static and dynamic logit models is the interpretation of the v_θ function: in the static logit model it is a one period utility function that is typically specified as a linear-in-parameters function of θ, whereas in the dynamic logit model it is the sum of a one period utility function plus the expected discounted utility in all future periods. Since the functional form of v_θ in DDP is generally not known a priori, its values must be computed numerically for any particular value of θ. As a result, maximum likelihood estimation of DDP models requires a "nested numerical solution algorithm" consisting of an "outer" optimization algorithm that searches over the parameter space Θ to maximize the likelihood function and an "inner" dynamic programming algorithm that solves (or approximately solves) the stochastic control problem and computes the choice probabilities $P(d|x, \theta)$ and derivatives $\partial P(d|x, \theta)/\partial\theta$ for each trial value of θ. There are a number of fast algorithms for solving finite- and infinite-horizon stochastic control problems, but space constraints prevent more than a cursory discussion of the main methods in this chapter.

Section 3.3 presents other econometric specifications for the error term that allow ε_t to enter u in a nonlinear, non-additive fashion, and, also, specifications with more complicated patterns of serial dependence in $\{\varepsilon_t\}$ than is allowed by the CI assumption. Section 3.4 discusses the simulation estimator proposed by Hotz et al. (1993) that avoids the computational burden of the nested numerical solution methods, and the associated "curse of dimensionality", i.e. the exponential rise in the amount of computer time/space required to solve a DDP problem as its "size" (measured in terms of number of possible values the state and control variables can assume) increases. However, the curse of dimensionality also has implications for the "data" and "estimation" complexity of a DDP model: as the size (i.e. the level of realism or detail) of a DDP model increases, the amount of data needed to estimate the model with an acceptable degree of precision increases more than proportionately. The problems are most severe for estimating beliefs, p. Subjective beliefs can be very slippery, high-dimensional objects to estimate. Since the optimal decision rule δ is generally quite sensitive to the specification of p, an innaccurate or inconsistent estimate of p will contaminate the estimates of u and β. Even under the assumption of rational expectations (which allows us to estimate p non-parametrically), the number of observations required to calculate estimates of p of specified accuracy increases exponentially with the number of state and control variables included in the model. The simulation estimator is particularly data-dependent in that it requires accurate non-parametric estimates of agents' conditional choice probabilities P as well as their beliefs p.

Given all the difficulties involved in structural estimation, the reader might wonder why not simply estimate agents' conditional choice probabilities P using simpler flexible parametric and non-parametric estimation methods. Of course, reduced-form methods can be used, and are quite useful for initial exploratory data analysis and judging whether more tightly parameterized structural models are misspecified. Nevertheless there is considerable interest in structural estimation methods for both intellectual and practical reasons. The intellectual reason is that structural estimation is the most direct way to assess the empirical validity of a specific MDP model: in the process of solving, estimating, and testing a particular MDP model we learn not only about the data, but the detailed implications of the theory. The practical motivation is that structural models can generate more accurate predictions of the impacts of policy changes than reduced-form models. As Lucas (1976) noted, reduced-form econometric techniques can be thought of as uncovering the form of an agent's historical decision rule. The resulting estimate $\hat{\delta}$ can then be used to predict the agent's behavior in the future, provided that the environment is stationary. Lucas showed that reduced-form estimates can produce very misleading forecasts of the effects of policy changes that alter the stochastic environment that agents will face in the future.[4] The reason is that a policy α (such as government rules for payment of Social Security or welfare benefits) can affect an agent's preferences, beliefs and discount factor. If we denote the dependence of primitives on policy as $(u_\alpha, p_\alpha, \beta_\alpha)$, then under a new decision rule α' the agent's behavior will be given by a new decision rule $\delta(u_{\alpha'}, p_{\alpha'}, \beta_{\alpha'})$ rather than the historical decision rule $\delta(u_\alpha, p_\alpha, \beta_\alpha)$. Unless there has been a lot of historical variation in policies α, reduced-form models won't be able to estimate the independent effect of α on δ, and, therefore, we won't be able to predict how agents will react to a hypothetical policy α'. However if we are able to parameterize the way in which policy affects the primitives, $(u_\alpha, p_\alpha, \beta_\alpha)$, then it is a typically straightforward exercise to compute the new decision rule $\delta(u_{\alpha'}, p_{\alpha'}, \beta_{\alpha'})$ for a hypothetical policy α'.

One can push this line of argument only so far, since its validity depends on the assumption that agents really are rational expected-utility maximizers and the structural model is correctly specified. If we admit that a tightly parameterized structural model is at best an abstract and approximate representation of reality, there is no reason why a structural model necessarily yields more accurate forecasts than reduced-form models. Furthermore, because of the identification problem it is possible that we could have a situation where two distinct sets of primitives fit an historical data set equally well, but yield very different predictions about the impact of a hypothetical policy. Under such circumstances there is no objective basis for choosing one prediction over another, and we may have to go to the expense of

[4] The limitations of reduced-form models have also been pointed out in an earlier paper by Marschak (1953), although his exposition pertained more to the static econometric models of that period. These general ideas can be traced back even further to the work of Haavelmö (1944) and others at the Cowles Commission.

conducting a controlled experiment to help identify the primitives and predict the impact of a new policy α'.[5] In spite of these problems, the final section of this chapter provides some empirical applications that demonstrate the ability of simple structural models to make much more accurate predictions of the effects of various policy changes than reduced-form models.

Readers who are familiar with the theory of stochastic control are free to skip the brief review of theory and solution methods in Section 2 and move directly to the econometric implementation of the theory in Section 3. A general observation about the current state of the art in this literature is that, while it is easy to formulate very general and detailed MDP's, Bellman's "curse of dimensionality" implies that our ability to actually solve and estimate these problems is much more limited.[6] However, recent research [Rust (1995b)] shows that use of random Monte Carlo integration methods does succeed in breaking the curse of dimensionality for the subclass of DDP's. This result offers the promise that fairly realistic and detailed DDP models will be estimable in the near future. The approach of this chapter is to start with a presentation of the general theory of MDP's and then show how various restrictions on the general theory lead to subclasses of econometric models that are feasible to estimate.

The first general restriction is to exclude MDP's formulated in continuous time. Although many of the results described in Section 3 can be generalized to continuous-time semi-Markov processes [Ahn (1993b)], there has been little progress on extending the theory to cover other types of continuous-time objects such as controlled diffusion processes. The rationale for using discrete-time models is that solutions to continuous-time problems can be arbitrarily closely approximated by solutions to corresponding discrete-time versions of the problem [cf. Gihman and Skorohod (1979, Chapter 2.3), van Dijk (1984)]. Indeed the standard approach to solving continuous-time stochastic control problems involves solving an approximate version of the problem in discrete time [Kushner (1990)].

The second restriction is implicit in the theory of stochastic control, namely the assumption that agents conform to the von Neumann–Morgenstern axioms for choice under uncertainty so that their preferences can be represented by the expected value of a cardinal utility function. A number of experiments have indicated that human decision-making under uncertainty may not always be consistent with the von Neumann–Morgenstern axioms.[7] In addition, expected-utility models imply that agents are indifferent about the timing of the resolution of uncertain events, whereas human decision-makers seem to have definite preferences over the time at which uncertainty is resolved [Kreps and Porteus (1978), Chew and Epstein (1989)]. The justification for focusing on expected utility is that it remains the most tractable

[5] Experimental data are subject to their own problems, and it would be a mistake to think of controlled experiments as the only reliable way to predict the response to a new policy. See Heckman (1991, 1994) for an enlightening discussion of some of these limitations.

[6] See Rust (1994, Section 2) for a more detailed discussion of some of the problems faced in estimating MDP's.

[7] Machina (1982) identifies the "independence axiom" as the source of many of the discrepancies.

framework for modelling choice under uncertainty.[8] Furthermore, Section 3.5 shows that, from an econometric standpoint, the expected-utility framework is sufficiently rich to model virtually any type of observed behavior. Our ability to discriminate between expected utility and the more subtle non-expected-utility theories of choice under uncertainty may require quasi-econometric methods such as controlled experiments.[9]

2. Solving MDP's via dynamic programming: A brief review

This section reviews the main results on dynamic programming in finite-horizon problems, and the functional equations that must be solved in infinite-horizon problems. Due to space constraints I only give a cursory outline of the main numerical methods for solving these functional equations, referring the reader to Puterman (1990) or Rust (1995a, 1996) for more in-depth surveys.

Definition 2.1

A (discrete-time) Markovian decision process consists of the following objects:

- A time index $t \in \{0, 1, 2, \ldots, T\}$, $\quad T \leqslant \infty$;
- A state space S;
- A decision space D;
- A family of constraint sets $\{D_t(s_t) \subseteq D\}$;
- A family of transition probabilities $\{p_{t+1}(\cdot|s_t, d_t): \mathscr{B}(S) \Rightarrow [0, 1]\}$;[10]
- A family of discount functions $\{\beta_t(s_t, d_t) \geqslant 0\}$ and single period utility functions $\{u_t(s_t, d_t)\}$ such that the utility functional U has the additively separable decomposition[11]

$$U(\boldsymbol{s}, \boldsymbol{d}) = \sum_{t=0}^{T} \left[\prod_{j=0}^{t-1} \beta_j(s_j, d_j) \right] u_t(s_t, d_t). \tag{2.1}$$

[8]Recent work by Epstein and Zin (1989) and Hansen and Sargent (1992) on models with non-separable, non-expected-utility functions shows that certain specifications are computationally and analytically tractable. Epstein and Zin have already used their specification of preferences in an empirical investigation of asset pricing. Despite these promising beginnings, the theory and computational methods for these more general problems are in their infancy, and due to space constraints, we are unable to cover these methods in this survey.

[9]An example of the ability of laboratory experiments to uncover discrepancies between human behavior and the predictions of expected-utility theory is the "Allias paradox" described in Machina (1982, 1987).

[10]$\mathscr{B}(S)$ is the Borel σ-algebra of measurable subsets of s. For simplicity, the rest of this chapter avoids measure-theoretic details since they are superfluous in the most commonly encountered case where both the state and control variables are discrete. See Rust (1996) for a statement of the required regularity conditions for problems with continuous state and control variables.

[11]The boldface notation denotes sequences: $s = (s_0, \ldots, s_T)$. Also, define $\Pi_{j=0}^{-1} \beta_j(s_j, d_j) = 1$ in formula (2.1).

The agent's optimization problem is to choose an optimal decision rule $\delta^* = (\delta_0, \ldots, \delta_T)$ to solve the following problem:

$$\max_{\delta = (\delta_0, \ldots, \delta_T)} E_\delta\{U(\boldsymbol{s}, \boldsymbol{d})\}. \tag{2.2}$$

2.1. *Finite-horizon dynamic programming and the optimality of Markovian decision rules*

In finite-horizon problems ($T < \infty$), the optimal decision rule $\delta^* = (\delta_0^T, \ldots, \delta_T^T)$ can be computed by backward induction starting at the terminal period, T. In principle, the optimal decision at each time t can depend not only on the current state s_t, but on the entire previous history of the process, $d_t = \delta_t^T(s_t, H_{t-1})$ where $H_t = (s_0, d_0, \ldots, s_{t-1}, d_{t-1})$. However in carrying out the process of backward induction it is easy to see that the Markovian structure of p and the additive separability of U imply that it is unnecessary to keep track of the entire previous history: the optimal decision rule depends only on the current time t and the current state s_t: $d_t = \delta_t^T(s_t)$. For example, starting in period T we have

$$\delta_T(H_{T-1}, s_T) = \underset{d_T \in D_T(s_T)}{\operatorname{argmax}} \; U(H_{T-1}, s_T, d_T), \tag{2.3}$$

where U can be rewritten as

$$\begin{aligned} U(H_{T-1}, s_T, d_T) &= \sum_{t=0}^{T} \left(\prod_{j=0}^{t-1} \beta_j(s_j, d_j) \right) u_t(s_t, d_t) \\ &= \sum_{t=0}^{T-1} \left(\prod_{j=0}^{t-1} \beta_j(s_j, d_j) \right) u_t(s_t, d_t) + \left(\prod_{j=0}^{T-1} \beta_j(s_j, d_j) \right) u_T(s_T, d_T). \end{aligned} \tag{2.4}$$

From (2.4) it is clear that previous history H_{T-1} does not affect the optimal decision of d_T in (2.3) since d_T appears only in the final term $u_T(s_T, d_T)$ on the right hand side of (2.4). Since the final term is affected by H_{T-1} only by the multiplicative discount factor $\prod_{j=0}^{T-1} \beta_j(s_j, d_j)$, it's clear that δ_T depends only on s_T. Working backwards recursively, it is straightforward to verify that at each time t the optimal decision rule δ_t depends only on s_t. A decision rule that depends on the past history of the process only via the current state s_t is called *Markovian*. Notice also that the optimal decision rule will generally be a deterministic function of s_t because randomization can only reduce expected utility if the optimal value of d_T in (2.3) is unique. This is a generic property, since if there are two distinct values of $d \in D_T(S_T)$ that attain the maximum in (2.3) by a slight perturbation of U, we obtain a similar model where the maximizing value is unique.

The *value function* is the expected discounted value of utility over the remaining horizon assuming an optimal policy is followed in the future. The method of dynamic programming calculates the value function and the optimal policy recursively as follows. In the terminal period V_T^T and δ_T^T are defined by

$$\delta_T^T(s_T) = \underset{d_T \in D_T(s_T)}{\operatorname{argmax}}\, u_T(s_T, d_T), \tag{2.5}$$

$$V_T^T(s_T) = \max_{d_T \in D_T(s_T)} u_T(s_T, d_T). \tag{2.6}$$

In periods $t = 0, \ldots, T-1$, V_t^T and δ_t^T are recursively defined by

$$\delta_t^T(s_t) = \underset{d_t \in D_t(s_t)}{\operatorname{argmax}} \left\{ u_t(s_t, d_t) + \beta_{t-1}(s_{t-1}, d_{t-1}) \int V_{t+1}^T[s_{t+1}, \delta_{t+1}^T(s_{t+1})] p_{t+1}(\mathrm{d}s_{t+1} | s_t, d_t) \right\}, \tag{2.7}$$

$$V_t^T(s_t) = \max_{d_t \in D_t(s_t)} \left\{ u_t(s_t, d_t) + \beta_{t-1}(s_{t-1}, d_{t-1}) \int V_{t+1}^T[s_{t+1}, \delta_{t+1}^T(s_{t+1})] p_{t+1}(\mathrm{d}s_{t+1} | s_t, d_t) \right\}. \tag{2.8}$$

It's straightforward to verify that at time $t = 0$ the value function $V_0^T(s_0)$ represents the conditional expectation of utility over all future periods. Since dynamic programming has recursively generated the optimal decision rule $\delta^* = (\delta_0^T, \ldots, \delta_T^T)$, it follows that

$$V_0^T(s) = \max_{\delta} E_\delta \{ U(\tilde{s}, \tilde{d}) | s_0 = s \}. \tag{2.9}$$

These results can be formalized as follows.

Theorem 2.1

Given an MDP that satisfies certain weak regularity conditions [see Gihman and Skorohod (1979)],

1. An optimal, non-randomized decision rule δ^* exists,
2. An optimal decision rule can be found within the subclass of non-randomized Markovian strategies,
3. In the finite-horizon case ($T < \infty$) an optimal decision rule δ^* can be computed by backward induction according to the recursions (2.5), ..., (2.8),
4. In the infinite-horizon case ($T = \infty$) an optimal decision rule δ^* can be approximated arbitrarily closely by the optimal decision rule δ_N^* to an N-period problem in the sense that

$$\lim_{N \to \infty} E_{\delta_N^*} \{ U_N(\tilde{s}, \tilde{d}) \} = \lim_{N \to \infty} \sup_{\delta} E_\delta \{ U_N(\tilde{s}, \tilde{d}) \} = \sup_{\delta} E_\delta \{ U(\tilde{s}, \tilde{d}) \}. \tag{2.10}$$

2.2. Infinite-horizon dynamic programming and Bellman's equation

Further simplifications are possible in the case of stationary MDP's. In this case the transition probabilities and utility functions are the same for all t, and the discount functions $\beta_t(s_t, d_t)$ are set equal to some constant $\beta \in [0, 1)$. In the finite-horizon case the time homogeneity of u and p does not lead to any significant simplifications since there still is a fundamental non-stationarity induced by the fact that remaining utility $\sum_{j=t}^{T} \beta^j u(s_j, d_j)$ depends on t. However in the infinite-horizon case, the stationary Markovian structure of the problem implies that the future looks the same whether the agent is in state s_t at time t or in state s_{t+k} at time $t+k$ provided that $s_t = s_{t+k}$. In other words, the only variable which affects the agent's view about the future is the value of his current state s. This suggests that the optimal decision rule and corresponding value function should be time invariant, i.e. for all $t \geqslant 0$ and all $s \in S$, $\delta_t^\infty(s) = \delta(s)$ and $V_t^\infty(s) = V(s)$. Analogous to equation (2.7), δ satisfies

$$\delta(s) = \underset{d \in D(s)}{\operatorname{argmax}} \left[u(s, d) + \beta \int V(s') p(\mathrm{d}s' | s, d) \right], \tag{2.11}$$

where V is defined recursively as the solution to *Bellman's equation*,

$$V(s) = \max_{d \in D(s)} \left[u(s, d) + \beta \int V(s') p(\mathrm{d}s' | s, d) \right]. \tag{2.12}$$

It is easy to see that if a solution to Bellman's equation exists, then it must be unique. Suppose that $W(s)$ is another solution to (2.12). Then we have

$$\begin{aligned} |V(s) - W(s)| &\leqslant \beta \int \max_{d \in D(s)} |V(s') - W(s')| p(\mathrm{d}s' | s, d) \\ &\leqslant \beta \sup_{s \in S} |V(s) - W(s)|. \end{aligned} \tag{2.13}$$

Since $0 < \beta < 1$, the only solution to (2.13) is $\sup_{s \in S} |V(s) - W(s)| = 0$.

2.3. Bellman's equation, contraction mappings and optimality

To establish the existence of a solution to Bellman's equation, assume for the moment the following regularity conditions: (1) $u(s, d)$ is jointly continuous and bounded in (s, d), (2) $D(s)$ is a continuous correspondence. Let $C(S)$ denote the vector space of all continuous, bounded functions $f: S \to R$ under the supremum norm, $\| f \| = \sup_{s \in S} |f(s)|$. Then $C(S)$ is a Banach space, i.e. a complete normed linear space.[12] Define an operator $\Gamma: C(S) \to C(S)$ by

$$\Gamma(W)(s) = \max_{d \in D(s)} \left[u(s,d) + \beta \int W(s')p(\mathrm{d}s' \mid s,d) \right]. \tag{2.14}$$

Bellman's equation can then be rewritten in operator notation as

$$V = \Gamma(V), \tag{2.15}$$

i.e. V is a fixed point of the mapping Γ. Using an argument similar to (2.13) it is easy to show that given any $V, W \in C(S)$ we have

$$\| \Gamma(V) - \Gamma(W) \| \leqslant \beta \| V - W \|. \tag{2.16}$$

An operator that satisfies inequality (2.16) for some $\beta \in (0,1)$ is said to be a *contraction mapping.*

Theorem 2.2. (Contraction mapping theorem)

If Γ is a contraction mapping on a Banach space B, then Γ has a unique fixed point V.

The uniqueness of the fixed point can be established by an argument similar to (2.13). The existence of a solution is a result of the completeness of the Banach space B. Starting from any initial element of B (such as 0), the contraction property (2.16) implies that the following sequence of successive approximations forms a Cauchy sequence in B:

$$\{0, \Gamma(0), \Gamma^2(0), \Gamma^3(0), \ldots, \Gamma^n(0), \ldots\}. \tag{2.17}$$

Since the Banach space B is complete, the Cauchy sequence converges to a point $V \in B$, so existence follows by showing that V is a fixed point of Γ. To see this, note that a contraction Γ is (uniformly) continuous, so

$$V = \lim_{n \to \infty} \Gamma^n(0) = \lim_{n \to \infty} \Gamma[\Gamma^{n-1}(0)] = \Gamma(V), \tag{2.18}$$

i.e. V is indeed the required fixed point of Γ.

We now show that given the single period decision rule δ defined in (2.11) the stationary, infinite-horizon policy $\delta^* = (\delta, \delta, \ldots)$ does in fact constitute an optimal decision rule for the infinite-horizon MDP. This result follows by showing that the unique solution $V(s)$ to Bellman's equation coincides with the optimal value

[12] A space B is said to be complete if every Cauchy seuence in B converges to a point in B.

function V_0^∞ defined by

$$V_0^\infty(s) \equiv \max_\delta E_\delta\left\{\sum_{t=0}^{\infty} \beta^t u(s_t, d_t) | s_0 = s\right\}. \tag{2.19}$$

Consider approximating the infinite-horizon problem by the solution to a finite-horizon problem with value function

$$V_0^T(s) \equiv \max_\delta E_\delta\left\{\sum_{t=0}^{T} \beta^t u(s_t, d_t) | s_0 = s\right\}. \tag{2.20}$$

Since u is bounded and continuous, $\sum_{t=0}^{T} \beta^t u(s_t, d_t)$ converges to $\sum_{t=0}^{\infty} \beta^t u(s_t, d_t)$ for any sequences $\boldsymbol{s} = (s_0, s_1, \ldots)$ and $\boldsymbol{d} = (d_0, d_1, \ldots)$. Theorem 2.1(4) implies that for each $s \in S$, $V_0^T(s)$ converges to the infinite-horizon value function $V_0^\infty(s)$:

$$\lim_{T\to\infty} V_0^T(s) = V_0^\infty(s) \qquad \forall s \in S. \tag{2.21}$$

But the contraction mapping theorem also implies that this same sequence converges to V (since $V_0^T = \Gamma^T(0)$), so $V = V_0^\infty$. Since V is the expected present discounted value of utility under the policy δ^* (a result we demonstrate in Section 2.4), the fact that $V = V_0^\infty$ immediately implies the optimality of δ^*.

A similar result can be proved under weaker conditions that allow $u(s, d)$ to be an unbounded function of the state variable. As we will see in Section 3, unbounded utility functions arise in DDP problems as a consequence of assumptions about the distribution of unobserved state variables. Although the contraction mapping theorem is no longer directly applicable, one can prove the following result, a generalization of Blackwell's theorem, under a weaker set of regularity conditions that allows for unbounded utility.

Theorem 2.3 (Blackwell's theorem)

Given an infinite-horizon, time homogeneous MDP that satisfies certain regularity conditions [see Bhattacharya and Majumdar (1989)];

1. A unique solution V to Bellman's equation (2.12) exists, and it coincides with the optimal value function defined in (2.19),
2. There exists a stationary, non-randomized, Markovian optimal control δ^* given by the solution to (2.11),
3. There is an optimal non-randomized, Markovian decision rule δ^* which can be approximated by the solution δ_N^* to an N-period problem with utility function $U_N(s, \boldsymbol{d}) = \sum_{t=0}^{N} \beta^t u(s_t, d_t)$:

$$\lim_{N\to\infty} E_{\delta_N^*}\{U_N(\tilde{s}, \tilde{\boldsymbol{d}})\} = \lim_{N\to\infty} \sup_\delta E_\delta\{U_N(\tilde{s}, \tilde{\boldsymbol{d}})\} = \sup_\delta E_\delta\{U(\tilde{s}, \tilde{\boldsymbol{d}})\} = E_{\delta^*}\{U(\tilde{s}, \tilde{\boldsymbol{d}})\}. \tag{2.22}$$

2.4. A geometric series representation for MDP's

Presently, the most commonly used solution procedure for MDP problems involves discretizing continuous state and control variables into a finite number of possible values. This resulting class of finite state DDP problems has a simple and beautiful algebraic structure that we now review.[13] Without loss of generality we can identify the state and decision spaces as finite sets of integers $\{1,\dots,S\}$ and $\{1,\dots,D\}$, and the constraint set as $\{1,\dots,D(s)\}$ where for notational simplicity we now let S,D and $D(s)$ denote positive integers rather than sets. It follows that a feasible stationary decision rule δ is an S-dimensional vector satisfying $\delta(s)\in\{1,\dots,D(s)\}$, $s=1,\dots,S$, and the value function V is an S-dimensional vector in the Euclidean space R^S. Given δ we can define a vector $u_\delta\in R^S$ whose ith component is $u[i,\delta(i)]$, and an $S\times S$ *transition probability matrix* E_δ whose (i,j) element is $p[i|j,\delta(j)]=\Pr\{s_{t+1}=i|s_t=j,d_t=\delta(j)\}$. Bellman's equation for a DDP reduces to

$$\Gamma(V)(s)=\max_{1\leqslant d\leqslant D(s)}\left[u(s,d)+\beta\sum_{s'=1}^{S}V(s')p(s'|s,d)\right]. \tag{2.23}$$

Given a stationary, Markovian decision rule δ, we define $V_\delta\in R^S$ as the vector of expected discounted utilities under policy δ. It is straightforward to show that V_δ is the solution to a system of linear equations,

$$V_\delta=u_\delta+\beta E_\delta V_\delta, \tag{2.24}$$

which can be solved by matrix inversion:

$$\begin{aligned}V_\delta&=G_\delta(V_\delta)\\&=[I-\beta E_\delta]^{-1}u_\delta\\&=u_\delta+\beta E_\delta u_\delta+\beta^2E_\delta^2u_\delta+\beta^3E_\delta^3u_\delta+\cdots.\end{aligned} \tag{2.25}$$

The last equation in (2.25) is simply a geometric series expansion for V_δ in powers of β and E_δ. As is well known, $E_\delta^N=(E_\delta)^N$ is simply the N-stage transition probability matrix, whose (i,j) element equals $\Pr\{s_{t+N}=i|s_t=j,\delta\}$, where the presence of δ as a conditioning argument denotes the fact that all intervening decisions satisfy $d_{t+j}=\delta(s_{t+j})$, $j=0,\dots,N$. Since $\beta^NE_\delta^Nu_\delta$ is the expected discounted utility received in period N under policy δ, formula (2.25) can be thought of as a vector generalization of a geometric series, showing explicitly how V_δ equals the sum of expected discounted utilities under δ in all future periods.[14] Since E_δ^N is a transition probability matrix (i.e. all elements are between 0 and 1, and its rows sum to unity), it

[13]The geometric representtion also holds for continous state MDP's, but in infinite-dimensional space instead of R^s.

[14]As Lucas (1978) notes, "a little knowledge of geometric series goes a long way".

follows that $\lim_{N\to\infty} \beta^N E_\delta^N = 0$, guaranteeing the invertibility of $[I - \beta E_\delta]$ for any Markovian decision rule δ and all $\beta \in [0, 1)$.[15]

2.5. *Overview of solution methods*

This section provides a brief review of solution methods for MDP's. For a more extensive review we refer the reader to Rust (1995a).

The main solution method for finite-horizon MDP's is backward recursion, which has already been described in Section 2.1. The amount of computer time/space required to solve a problem increases linearly in the length of the horizon T and quadratically in the number of possible state variables S, the latter result being due to the fact that the main work involved in dynamic programming is calculating the conditional expectation of future utility, which requires multiplying an $S \times S$ transition matrix by the $S \times 1$ value function.

In the infinite-horizon case there are a variety of solution methods, most of which can be viewed as different strategies for solving the Bellman functional equation. The method of *successive approximations* which we described in Section 2.2 is probably the most well-known solution method for infinite-horizon problems: it essentially amounts to using the solution to a finite-horizon problem with a large horizon T to approximate the solution to the infinite-horizon problem. In certain cases we can significantly accelerate successive approximations by employing the *McQueen–Porteus error bounds,*

$$\Gamma^k(V) + \underline{b}_k e \leqslant V^* \leqslant \Gamma^k(V) + \bar{b}_k e, \tag{2.26}$$

where V^* is the fixed point to Γ, e denotes an $S \times 1$ vector of 1's, and

$$\begin{aligned} \underline{b}_k &= \beta/(1-\beta)\min[\Gamma^k(V) - \Gamma^{k-1}(V)], \\ \bar{b}_k &= \beta/(1-\beta)\max[\Gamma^k(V) - \Gamma^{k-1}(V)]. \end{aligned} \tag{2.27}$$

The contraction property guarantees that $\underline{b}_k$ and $\bar{b}_k$ approach each other geometrically at rate β. The fact that the fixed point V^* is bracketed within these bounds suggests that we can obtain an improved estimate of V^* by terminating the contraction iterations when $|\bar{b}_k - \underline{b}_k| < \varepsilon$ and setting the final estimate of V^* to be the median bracketed value

$$\hat{V}_k = \Gamma^k(V_0) + \left(\frac{\bar{b}_k + \underline{b}_k}{2}\right)e. \tag{2.28}$$

[15] If there are continous state variables, the MDP problem still has the same representation as in (2.25), except that E_δ is a Markov operator (a bounded, positive linear operator with norm equal to 1) instead of an $S \times S$ transition probability matrix.

Bertsekas (1987, p. 195) shows that the rate of convergence of $\{\hat{V}_k\}$ to V^* is geometric at rate $\beta|\lambda_2|$, where λ_2 is the subdominant eigenvalue of E_{δ^*}. In cases where $|\lambda_2| < 1$, the use of the error bounds can lead to significant speed-ups in the convergence of successive approximations at essentially no extra computational cost. However in problems where E_{δ^*} has multiple ergodic sets, $|\lambda_2| = 1$, and the error bounds will not lead to an appreciable speed improvement as illustrated in computational results in Table 5.2 of Bertsekas (1987).

In relatively small scale problems ($S < 10\,000$) the method of *policy iteration* is generally the fastest method for computing V^* and the associated optimal decision rule δ^*, provided the discount factor is sufficiently large ($\beta > 0.95$). The method starts by choosing an arbitrary initial policy, δ_0.[16] Next a policy valuation step is carried out to compute the value function V_{δ_0} implied by the stationary decision rule δ_0. This requires solving the linear system (2.25). Once the solution V_{δ_0} is obtained, a policy improvement step is used to generate an updated policy δ_1,

$$\delta_1(s) = \underset{1 \leqslant d \leqslant D(s)}{\operatorname{argmax}} \left[u(s,d) + \beta \sum_{s'=1}^{S} V_{\delta_0}(s')p(s'|s,d)\right]. \tag{2.29}$$

Given δ_1, one continues the cycle of policy valuation and policy improvement steps until the first iteration k such that $\delta_k = \delta_{k-1}$ (or alternatively $V_{\delta_k} = V_{\delta_{k-1}}$). It is easy to see from (2.25) and (2.29) that such a V_{δ_k} satisfies Bellman's equation (2.23), so that by Theorem 2.3 the stationary Markovian decision rule $\delta^* = \delta_k$ is optimal. One can show that policy iteration always generates an improved policy:

$$V_{\delta_k} \geqslant V_{\delta_{k-1}}. \tag{2.30}$$

Since there are only a finite number $D(1) \times \cdots \times D(S)$ of feasible stationary Markov policies, it follows that policy iteration always converges to the optimal decision rule δ^* in a finite number of iterations.

Policy iteration is able to find the optimal decision rule after testing an amazingly small number of trial policies δ_k. However the amount of work per iteration is larger than for successive approximations. Since the number of algebraic operations needed to solve the linear system (2.25) for V_{δ_k} is of order S^3, the standard policy iteration algorithm becomes impractical for S much larger than 10 000.[17] To solve very large scale MDP problems, it seems that the best strategy is to use policy iteration, but to only attempt to approximately solve for V_δ in each policy evaluation step (2.25). There are a number of variants of policy iteration that avoid direct numerical solution of the linear system in (2.25), including modified policy iteration

[16] One obvious choice is $\delta_0(s) = \operatorname{argmax}_{1 \leqslant d \leqslant D(s)}[u(s,d)]$.

[17] Supercomputers using combinations of vector processing and multitasking can now routinely solve dense linear systems exceeding 1000 equations and unknowns in under 1 CPU second. See, for example, Dongara and Hewitt (1986).

[Puterman and Shin (1978)], and adaptive state aggregation algorithms [Bertsekas and Castañon (1989)].

Puterman and Brumelle (1978, 1979) have shown that policy iteration is identical to Newton's method for computing a zero to a nonlinear function. This insight turns out to be useful for computing fixed points to contraction mappings Ψ that are closely related to, but distinct from, the contraction mapping Γ defined by Bellman's equation (2.11). An example of such a mapping is $\Psi: B \to B$ defined by

$$\Psi(v)(s,d) = u(s,d) + \beta \int \log\Bigg[\sum_{d' \in D(s')} \exp\{v(s',d')\}\Bigg] p(\mathrm{d}s' \mid s,d). \tag{2.31}$$

In Section 3 we show that the fixed point to this mapping is identical to the value function v_θ entering the dynamic logit model (1.1). Rewriting the fixed point condition as $0 = v - \Psi(v)$, we can apply Newton's method, generating iterations of the form

$$v_{k+1} = v_k - [I - \Psi'(v_k)]^{-1}(I - \Psi)(v_k), \tag{2.32}$$

where I denotes the identity matrix and $\Psi'(v)$ is the gradient of Ψ evaluated at the point $v \in B$. An argument exactly analogous to the series expansion argument used to proved the existence of $[I - \beta E_\delta]^{-1}$ can be used to establish that the matrix $[I - \Psi'(v)]^{-1}$ is invertible, so the Newton iterations are always well-defined. Given a starting point v_0 in a domain of attraction sufficiently close to the fixed point v^* of Ψ, the Newton iterations will converge to v^* at a quadratic rate:

$$|V_{k+1} - V^*| \leqslant K|V_k - V^*|^2 \tag{2.33}$$

for a positive constant K.

Although Newton iterations yield rapid quadratic rates of convergence, it is only guaranteed to converge for initial estimates v_0 in a domain of attraction of v^* whereas the method of successive approximations yields much slower linear rates of convergence but are always guaranteed to converge to v^* starting from any initial point v_0.[18] This suggests the following hybrid method or polyalgorithm: start with successive approximations, and when the McQueen–Porteus error bounds indicate that one is sufficiently close to v^*, switch to Newton iterations to rapidly converge to the solution.

There is another class of methods, which Judd (1994) has termed *minimum weighted residual* (MWR) *methods*, that can be applied to solve general operator equations of the form

$$\Phi(v^*) = 0, \tag{2.34}$$

[18]Newton's method does exhibit global convergence in finite state DDP problems due to the fact that Newton's method and policy iteration are identical in this case, and policy iteration converges from any starting point. Thus the domain of attraction in this case is all of R^s.

where Φ: $B \to B$ is a nonlinear operator on a potentially infinite-dimensional Banach space B. For example, Bellman's equation is a special case of (2.34) for $\Phi(V) = [I - \Gamma](V)$. Similar to policy iteration, Newton's method becomes computationally burdensome in high-dimensional problems. To avoid this, MWR methods attempt to approximate the solution to (2.34) by restricting the search to a smaller-dimensional subspace B_N spanned by the basis elements $\{x_1, x_2, \ldots, x_N\}$. It follows that we can index any approximate solution $v \in B_N$ by a vector $c = (c_1, \ldots, c_N) \in R^N$:

$$v_c = c_1 x_1 + \cdots + c_N x_N. \tag{2.35}$$

Unless the true solution v^* is an element of B_N, $\Phi(v_c)$ will generally be non-zero for all vectors $c \in R^N$. The MWR method computes an estimate $v_{\hat{c}}$ of v^* using a value of $\hat{c}$ that solves

$$\hat{c} = \underset{c \in R^N}{\operatorname{argmin}} \, |\Phi(v_c)|. \tag{2.36}$$

Variants of MWR methods can be obtained by using different subspaces B_N (e.g., Legendre or Chebyshev polynomials, etc.) and different norms on $\Psi(v_c)$ (e.g., least squares or sup norm, etc.). In cases where B is an infinite-dimensional space (which occurs when the DDP problem contains continuous state variables), one must also choose a finite grid of points over which the norm in (2.36) is evaluated.

Although I have described MWR as parameterizing the value function in terms of a small number of unknown coefficients c, there are variants of this approach that are based on parameterizations of other features of the stochastic control problem such as the decision rule δ [Smith (1991)], or the conditional expectation operator E_δ [Marcet (1994)]. For simplicity, I refer to all these methods as MWR even though there are important differences in their computational implementation.

The advantage of the MWR approach is that it converts the problem of finding a zero of a high-dimensional operator equation (2.34) into the problem of finding a zero to a smaller-dimensional minimization problem (2.36). MWR methods may be particularly effective for solving DDP problems with several continuous state variables, since straightforward discretization methods quickly run into the curse of dimensionality. However a disadvantage of the procedure is the computational burden of solving (2.36) given that $|\Phi(v_c)|$ must be evaluated for each trial value of c. Typically, one uses approximate methods to evaluate $|\Phi(v_c)|$, such as Gaussian quadrature or Monte Carlo integration. Another disadvantage is that MWR methods are non-iterative, i.e. previous approximations $v_1, v_2, \ldots, v_{N-1}$ are not used to determine the next estimate v_N. In practice, one must make do with a single approximation v_N, however there is no analog of the McQueen–Porteus error bounds to tell us how far v_N is from the true solution. Indeed, there are no general theorems proving the convergence of MWR methods as the dimension N of the subspace increases. There are also problems to be faced in cases where Φ has multiple solutions V^*, and when the minimization problem (2.36) has multiple local minima. Despite these unresolved problems, versions of the MWR method have proved to be effective in a variety of applied problems. See, for example, Kortum (1993) (who has nested the MWR solution of (2.35) in an estimation routine), and Bansal et al. (1993) who have used Marcet's method of parameterized expectations to generate stochastic simulations of dynamic, stochastic models for use by their "non-parameteric simulation estimator".

A final class of methods uses Monte Carlo integration to avoid the computational burden of multivariate numerical integration that is the dominating factor that limits our ability to solve DDP problems. Keane and Wolpin (1994) developed a method that combines Monte Carlo integration and interpolation to dramatically reduce the solution time for large scale DDP problems with continuous multidimensional state variables. As we will see below, incorporation of unobservable state variables ε implies that DDP problems will always have these multidimensional continuous state variables. Recently, Rust (1995b) has introduced a "random multigrid algorithm" using a *random Bellman operator* $\tilde{\Gamma}$ that avoids the need for interpolation and repeated Monte Carlo simulations that is an inherent limiting future of Keane and Wolpin's method. Rust showed that his algorithm succeeds in breaking the curse of dimensionality of solving the DDP problem – i.e. the amount of computer time required to solve the DDP problem increases only polynomially rather than exponentially with the dimension d of the state variables using Rust's algorithms. These new methods offer the promise that substantially more realistic DDP models will be estimable in the near future.

3. Econometric methods for discrete decision processes

As we discussed in Section 1, structural estimation methods for DDP's are fundamentally different from the Euler equation methods used to estimate CDP's. Since the control variable is discrete, we cannot differentiate to derive first order necessary conditions characterizing the optimal decision rule $\delta^* = (\delta, \delta, \ldots)$. Instead each component function $\delta(s)$ is defined by a finite number of inequality conditions:[19]

$$d \in \delta(s) \Leftrightarrow \left\{ \forall d' \in D(s) \middle| u(s,d) + \beta \int V^*(s')p(\mathrm{d}s'|s,d) \geqslant u(s,d') + \beta \int V^*(s')p(\mathrm{d}s'|s,d') \right\}. \tag{3.1}$$

Econometric methods for DDP's borrow heavily from methods developed in the literature on estimation of static discrete choice models.[20] The primary difference between estimation of static versus dynamic models of discrete choice is that agents' choices are governed by the relative magnitude of the value function V rather than the single period utility function u. Even if the functional form of the latter is

[19] For notational simplicity, this section focuses on stationary infinite-horizon DDP problems and ignores the distinction between the optimal policy δ^* and its components $\delta^* = (\delta, \delta, \ldots)$.

[20] See McFadden (1981, 1984) for excellent surveys of the huge literature on estimation of static discrete choice models.

specified a priori, the value function is generally unknown, although it can be computed for any value of θ. To date, most empirical applications have used "nested numerical solution algorithms" that compute the best fitting estimate $\hat{\theta}$ by repeatedly solving the dynamic programming problem for each trial value of θ.

3.1. Alternative models of the "error term"

In addition to the numerical problems involved in computing the value function and optimal decision rule, we face the problem of how to incorporate "error terms" into the structural model. Error terms are necessary in light of "Blackwell's theorem" (Theorem 2.3) that the optimal decision rule $d = \delta(s)$ is a deterministic function of the agent's state s. Blackwell's theorem implies that if we were able to observe all components of s, then a correctly specified DDP model would be able to perfectly predict agents' behavior. Since no theory is realistically capable of perfectly predicting the behavior of human decision-makers, there are basically four ways to reconcile discrepancies between the predictions of the DP model and observed behavior: (1) optimization errors, (2) measurement errors, (3) approximation errors, and (4) unobserved state variables.[21]

An optimization error causes an agent who "intends" to behave according to the optimal decision rule δ to take an actual decision d given by

$$d = \delta(s) + \eta, \tag{3.2}$$

where η is interpreted as an error that prevents the agent from correctly calculating or implementing the optimal action $\delta(s)$. This interpretation of discrepancies between d and $\delta(s)$ seems logically inconsistent: if the agent knew that there were random factors that lead to ex post discrepancies between intended and realized decisions, he would re-optimize taking these uncertainties into account. The resulting decision rule will generally be different from the optimal decision rule δ when intended and realized decisions coincide. On the other hand, if η is simply a way of accounting for irrational or non-maximizing behavior, it is not clear why this behavior should take the peculiar form of random deviations from a rational decision rule δ. Given these logical difficulties, we ignore optimization errors as a way of explaining discrepancies between d and $\delta(s)$.

Measurement errors, due to response or coding errors, must surely be acknowledged in most empirical studies. Measurement errors are usually much more likely to occur in continuous components of s than in the discrete values of d, although significant errors can occur in the latter as a result of classification error (e.g. defining workers as choosing to work full-time vs. part-time based on noisy measurements of total hours of work). From an econometric standpoint, measurement

[21] Another method, *unobserved heterogeneity*, can be regarded as a special case of unobserved state variables in which certain components of the state vector vary over individuals but not over time.

errors in s create more serious difficulties since δ is typically a nonlinear function of s. Unfortunately, the problem of nonlinear errors-in-variables has not yet been satisfactorily resolved in the econometrics literature. In certain cases [Eckstein and Wolpin (1989b) and Christensen and Kiefer (1991b)], one can account for measurement error in a statistically and computationally tractable manner, although at the present time this approach seems to be highly problem-specific.

An approximation error is defined as the difference between the actual and predicted decision, $\varepsilon \equiv d - \delta(s)$. This approach amounts to an up-front admission that the DDP model is misspecified, and does not attempt to impose auxiliary statistical assumptions about the distribution of ε. The existence of such errors is hard to deny since by their very nature DDP models are simplified, abstract representations of human behavior and we would never expect their predictions to be 100% correct. Under this interpretation the econometric problem is to find a specification (u, p, β) that minimizes some metric of the approximation error such as mean squared prediction error. While this approach seems quite natural, it leads to a "degenerate" econometric model and estimators with poor asymptotic properties. The approximation error approach also suffers from ambiguity about the appropriate metric for determining whether a given model does or does not provide a good approximation to observed behavior.

The final approach, unobserved state variables, is the subject of Section 3.2.

3.2. *Maximum likelihood estimation of DDP's*

The remainder of this chapter focuses on structural estimation of DDP's with unobserved state variables. In these models the state variable s is partitioned into two components $s = (x, \varepsilon)$ where x is a state variable observed by both agent and econometrician and ε is observed only by the agent. The existence of unobserved state variables is quite plausible: it is unlikely that any survey could completely record all information that is relevant to the agent's decision-making process. It also provides a natural way to "rationalize" discrepancies between observed behavior and the predictions of the DDP model: even though the optimal decision rule $d = \delta(x, \varepsilon)$ is a deterministic function, if the specification of unobservables is sufficiently "rich" any observed (x, d) combination can be explained as the result of an optimal decision by an agent for an appropriate value of ε. Since ε enters the decision rule δ in a non-additive fashion, it is infeasible to estimate θ by nonlinear least squares. The preferred method for estimating θ is maximum likelihood using the *conditional choice probability*,

$$P(d|x) = \int I\{d = \delta(x, \varepsilon)\} q(\mathrm{d}\varepsilon | x), \tag{3.3}$$

where $q(\mathrm{d}\varepsilon | x)$ is the conditional distribution of ε given x (to be defined). Even though δ is a step function, integration over ε in (3.3) leads to a conditional choice probabilty that is a smooth function of θ provided that the primitives (u, p, β) are smooth functions of θ and the DDP problem satisfies certain general properties given in assumptions AS and CI below. These assumptions guarantee that the conditional choice probability has "full support":

$$d \in D(x) \Leftrightarrow P(d|x) > 0, \tag{3.4}$$

which is equivalent to saying that the set $\{\varepsilon | d = \delta(x, \varepsilon)\}$ has positive probability under $q(\mathrm{d}\varepsilon|x)$. We say that a specification for unobservables is *saturated* if (3.4) holds for all possible values of θ. The problem with an unsaturated specification is the possibility that the DDP model may be contradicted in a sufficiently large data set: i.e. one may encounter observations (x_t^a, d_t^a) which cannot be rationalized by any value of ε or θ, i.e. $P(d_t^a|x_t^a, \theta) = 0$ for all θ. This leads to practical difficulties in maximum likelihood estimation, causing the log-likelihood function to "blow up" when it encounters a "zero probability" observation. Although one might eliminate such observations to achieve convergence, the impact on the asymptotic properties of the estimator is unclear. In addition, an unsaturated specification may yield a likelihood function whose support depends on θ or which may be a non-smooth function of θ. Little is known about the general asymptotic properties of these "non-standard" maximum likelihood estimators.[22]

Borrowing from the literature on static discrete choice models [McFadden (1981)] we introduce two assumptions that are sufficient to generate a saturated specification for unobservables in a DDP model.

Assumption ***AS***

The choice sets depend only on the observed state variable $x: D(s) = D(x)$. The unobserved state variable ε is a vector with at least as many components as the number of elements in $D(x)$.[23] The utility function has the additively separable decomposition

$$u(s, d) = u(x, d) + \varepsilon(d), \tag{3.5}$$

where $\varepsilon(d)$ is the dth component of the vector ε.

[22] Results are available for certain special cases, such as Flinn and Heckman's (1982) and Christensen and Kiefer's (1991) analysis of the job search model. If wages are measured without error, this model generates the restriction that any accepted wage offer must be greater than the reservation wage (which is an implicit function of θ). This implies that the support of the likelihood function depends on θ, resulting in a non-normal limiting distribution with certain parameters converging faster than the $\sqrt{A}$ rate that is typical of standard maximum likelihood estimators. The basic result is analogous to estimating the upper bound θ of a uniform distribution $U[0, \theta]$. The support of this distribution clearly depends θ and, as well known (Cox and Hinckley, 1974), the maximum likelihood estimator is $\hat{\theta} = \max\{x_1, \ldots, x_A\}$, which converges at rate A to an exponential limiting distribution.

[23] For technical reasons ε may have a number of superfluous components so that we may formally embed the ε state vectors in a common state space e. For details, see Definition 3.1.

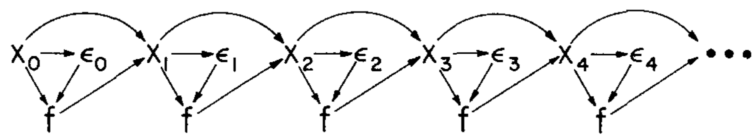

Figure 1. Pattern of dependence in controlled stochastic process implied by the CI assumption

Assumption ***CI***

The transition density for the controlled Markov process $\{x_t, \varepsilon_t\}$ factors as

$$p(\mathrm{d}x_{t+1}, \mathrm{d}\varepsilon_{t+1} | x_t, \varepsilon_t, d_t) = q(\mathrm{d}\varepsilon_{t+1} | x_{t+1})\pi(\mathrm{d}x_{t+1} | x_t, d_t), \tag{3.6}$$

where the marginal density of $q(\mathrm{d}\varepsilon | x)$ of the first $|D(x)|$ components of ε has support equal to $R^{|D(x)|}$ and finite absolute first moments.

CI is a conditional independence assumption which limits the pattern of dependence in the $\{x_t, \varepsilon_t\}$ process in two ways. First, x_{t+1} is a sufficient statistic for ε_{t+1} implying that any serial dependence between ε_t and ε_{t+1} is transmitted entirely through the observed state x_{t+1}.[24] Second, the probability density for x_{t+1} depends only on x_t and not on ε_t. Intuitively, CI implies that the $\{\varepsilon_t\}$ process is essentially a noise process superimposed on the main dynamics which are embodied by the transition probability $\pi(\mathrm{d}x' | x, d)$.

Under assumptions AS and CI Bellman's equation has the form

$$V(x, \varepsilon) = \max_{d \in D(x)} [v(x, d) + \varepsilon(d)], \tag{3.7}$$

where

$$v(x, d) = u(x, d) + \beta \int V(y, \varepsilon) q(\mathrm{d}\varepsilon | y) \pi(\mathrm{d}y | x, d). \tag{3.8}$$

Equation (3.8) is the key to subsequent results. It shows that the DDP problem has the same basic structure as a static discrete choice problem except that the value function v replaces the single period utility function u as an argument of the conditional choice probability. In particular, AS–CI yields a saturated specification for unobservables: (3.8) implies that the set $\{\varepsilon | d = \delta(x, \varepsilon)\}$ is a non-empty intersection of half-spaces in $R^{|D(x)|}$, and since ε is continuously distributed with unbounded support, it follows that regardless of the values of $\{v(x, d)\}$ the choice probability $P(d|x)$ is positive for each $d \in D(x)$.

In order to formally define the class of DDP's, we need to embed the unobserved state variables ε in a common space E. Without loss of generality, we can identify each choice set $D(x)$ as the set of integers $D(x) = \{1, \ldots, |D(x)|\}$, and let the decision space D be the set $D = \{1, \ldots, \sup_{x \in X} |D(x)|\}$. Then we define $E = R^{|D|}$, and whenever

[24]If $q(\mathrm{d}\varepsilon | x)$ is dependent of x then $\{\varepsilon_t\}$ is an IID process which is independent of $\{x_t\}$.

$|D(x)| < |D|$ then $q(d\varepsilon|x)$ assigns the remaining $|D| - |D(x)|$ "irrelevant" components of ε equal to some arbitrary value, say 0, with probability 1.

Definition 3.1

A discrete decision process (DDP) is an MDP satisfying the following restrictions:

- The decision space $D = \{1, \ldots, \sup_{s \in S} |D(s)|\}$, where $\sup_{s \in S} |D(s)| < \infty$.
- The state space S is the product space $S = X \times E$, where X is a Borel subset of R^J and $E = R^{|D|}$.
- For each $s \in S$ and $x \in X$ we have $D(s) = D(x) \subset D$.
- The utility function $u(s, d)$ satisfies assumption AS.
- The transition probability $p(ds_{t+1}|s_t, d_t)$ satisfies assumption CI.
- The component $q(d\varepsilon|x)$ of the transition probability $p(ds|s, d)$ is itself a product measure on $R^{|D(x)|} \times R^{|D|-|D(x)|}$ whose first component has support $R^{|D(x)|}$ and whose second component is a unit mass on a vector of 0's of length $|D| - |D(x)|$.

The conditional choice probability $P(d|x)$ can be defined in terms of a function McFadden (1981) has called the *social surplus*,

$$G[\{u(x, d), d \in D(x)\} | x] = \int_{R^{|D|}} \max_{d \in D(x)} [u(x, d) + \varepsilon(d)] q(d\varepsilon|x). \tag{3.9}$$

If we think of a population of consumers indexed by ε, then $G[\{u(x, d), d \in D(x)\} | x]$ is simply the expected indirect utility of choosing alternatives $d \in D(x)$. G has an important property, apparently first noted by Williams (1977) and Daly and Zachary (1979), that can be thought of as a discrete analog of Roy's identity.

Theorem 3.1

If $q(d\varepsilon|x)$ has finite first moments, then the social surplus function (3.9) exists, and has the following properties.

1. G is a convex function of $\{u(x, d), d \in D(x)\}$.
2. G satisfies the additivity property

$$G[\{u(x, d) + \alpha, d \in D(x)\} | x] = \alpha + G[\{u(x, d), d \in D(x)\} | x]. \tag{3.10}$$

3. The partial derivative of G with respect to $u(x, d)$ equals the conditional choice probability:

$$\frac{\partial G[\{u(x, d), d \in D(x)\} | x]}{\partial u(x, d)} = P(d|x). \tag{3.11}$$

From the definition of G in (3.9), it is evident that the proof of Theorem 3.1(3) is simply an exercise in interchanging integration and differentiation. Taking the partial derivative operator inside the integral sign we obtain[25]

$$\begin{aligned}\frac{\partial G[\{u(x,d), d\in D(x)\}|x]}{\partial u(x,d)} &= \int\left(\frac{\partial\{\max_{d\in D(y)}[u(x,d)+\varepsilon(d)]\}}{\partial u(x,d)}\right)q(\mathrm{d}\varepsilon|x)\\ &= \int I\{d = \operatorname*{argmax}_{d'\in D(x)}[u(x,d')+\varepsilon(d')]\}q(\mathrm{d}\varepsilon|x)\\ &= P(d|x). \end{aligned} \tag{3.12}$$

Note that the additivity property (3.10) implies that the conditional choice probabilities sum to 1, so $P(\cdot|x)$ is a well-defined probability distribution over $D(x)$.

The fact that the unobserved state variables ε have unbounded support implies that the objective function in the DDP problem is unbounded. We need to introduce three extra assumptions to guarantee the existence of a solution since the general results for MDP's given in Theorems 2.1 and 2.2 assumed that u is bounded above.

Assumption ***BU***

For each $d\in D(x)$, $u(x,d)$ is an upper semicontinuous function of x with bounded expectation:

$$\begin{aligned} R(x) &\equiv \sum_{t=1}^{\infty}\beta^t R_t(x) < \infty,\\ R_{t+1}(x) &= \max_{d\in D(x)}\int R_t(y)\pi(\mathrm{d}y|x,d),\\ R_1(x) &= \max_{d\in D(x)}\int\int \max_{d'\in D(y)}|u(y,d')+\varepsilon(d')|q(\mathrm{d}\varepsilon|y)\pi(\mathrm{d}y|x,d). \end{aligned} \tag{3.13}$$

Assumption ***WC***

$\pi(\mathrm{d}y|x,d)$ is a weakly continuous function of (x,d): for each bounded continuous function $h: X\to R$, $\int h(y)\pi(\mathrm{d}y|x,d)$ is a continuous function of x for each $d\in D(x)$.

Assumption ***BE***

Let B be the Banach space of bounded, Borel measurable functions $h: X\times D\to R$ under the essential supremum norm. Then $u\in B$ and for each $h\in B$, $Eh\in B$, where Eh is defined by

$$Eh(x,d) \equiv \int G[\{h(y,d), d\in D(y)\}|y]\pi(\mathrm{d}y|x,d). \tag{3.14}$$

[25]The interchange is justified by the Lebesgue dominated convergence theorem, since the derivative of $\max_{d\in D(x)}[u(x,d)+\varepsilon(d)]$ with respect to $u(x,d)$ is bounded (it equals either 0 or 1) for almost all ε.

Theorem 3.2

If $\{s_t, d_t\}$ is a DDP satisfying AS, CI and regularity conditions BU, WC and BE, then the optimal decision rule δ is given by

$$\delta(x,\varepsilon) = \underset{d \in D(x)}{\operatorname{argmax}}\,[v(x,d) + \varepsilon(d)], \tag{3.15}$$

where v is the unique fixed point to the contraction mapping $\Psi\colon B \to B$ defined by

$$\Psi(v)(x,d) = u(x,d) + \beta \int G[\{v(y,d'), d' \in D(y)\} \,|\, y]\pi(\mathrm{d}y \,|\, x, d). \tag{3.16}$$

Theorem 3.3

If $\{s_t, d_t\}$ is a DDP satisfying AS, CI and regularity conditions BU, WC and BE, then the controlled process $\{x_t, \varepsilon_t\}$ is Markovian with transition probability

$$\Pr\{\mathrm{d}x_{t+1}, d_{t+1} \,|\, x_t, d_t\} = P(d_{t+1} \,|\, x_{t+1})\pi(\mathrm{d}x_{t+1} \,|\, x_t, d_t), \tag{3.17}$$

where the conditional choice probability $P(d|x)$ is given by

$$P(d|x) = \frac{\partial G[\{v(x,d), d \in D(x)\} \,|\, x]}{\partial v(x,d)}, \tag{3.18}$$

where G is the social surplus function defined in (3.9), and v is the unique fixed point to the contraction mapping Ψ defined in (3.16).

The proofs of Theorems 3.2 and 3.3 are straightforward: under assumption AS–CI the value function is the unique solution to Bellman's equation given in (3.7) and (3.8). Substituting the formula for V given in (3.7) into the formula for v given in (3.8) we obtain

$$\begin{aligned} v(x,d) &= u(x,d) + \beta \int\!\!\int \max_{d' \in D(y)} [v(y,d') + \varepsilon(d')]q(\mathrm{d}\varepsilon \,|\, y)\pi(\mathrm{d}y \,|\, x, d) \\ &= u(x,d) + \beta \int G[\{v(y,d'), d' \in D(y)\} \,|\, y]\pi(\mathrm{d}y \,|\, x, d). \end{aligned} \tag{3.19}$$

The latter formula is the fixed point condition (3.16). It's a simple exercise to verify that Ψ is a contraction mapping, guaranteeing the existence and uniqueness of the function v. The fact that the observed components $\{x_t, d_t\}$ of the controlled process $\{x_t, \varepsilon_t, d_t\}$ is Markovian is a direct result of the CI assumption: the observed state x_{t+1} is a "sufficient statistic" for the agent's choice d_{t+1}. Without the CI assumption, lagged state and control variables would be useful for predicting the agent's choice at time $t+1$ and $\{x_t, d_t\}$ will no longer be Markovian. As we will see, this observation provides the basis for a specification test of CI.

For specific functional forms for q we obtain concrete formulas for the conditional choice probability $P(d|x)$, the social surplus function G and the contraction mapping Ψ. For example if $q(\mathrm{d}\varepsilon|x)$ is a multivariate extreme-value distribution we have[26]

$$q(\mathrm{d}\varepsilon|x) = \prod_{d\in D(x)} \exp\{-\varepsilon(d)+\gamma\}\exp[-\exp\{-\varepsilon(d)+\gamma\}] \qquad \gamma \doteq 0.577. \tag{3.20}$$

Then $P(d|x)$ is given by the well-known *multinomial logit* formula

$$P(d|x) = \frac{\exp\{v(x,d)\}}{\sum_{d'\in D(x)} \exp\{v(x,d')\}}, \tag{3.21}$$

where v is the fixed point to the contraction mapping Ψ:

$$\Psi(v)(x,d) = u(x,d) + \beta \int \log\left[\sum_{d'\in D(y)} \exp\{v(y,d')\}\right]\pi(\mathrm{d}y|x,d). \tag{3.22}$$

The extreme-value specification is especially attractive for empirical applications since the closed-form expressions for P and G avoid the need for multi-dimensional numerical integrations required for other distributions.[27] A simple consequence of the extreme value specification is that the log-odds ratio for two alternatives equals the utility differential:

$$\log\left\{\frac{P(d|x)}{P(1|x)}\right\} = u(x,d) - u(x,1). \tag{3.23}$$

Suppose that the utility function depends only on the attributes of the chosen alternative: $u(x,d) = u(x_d)$, where $x = (x_1, \ldots, x_D)$ is a vector of attributes of all the alternatives and x_d is the attribute of the dth alternative. In this case the log-odds ratio implies a property known as *independence from irrelevant alternatives* (IIA): the odds of choosing alternative d over alternative 1 depends only on the attributes of those two alternatives. The IIA property has a number of undesirable implications such as the "red bus/blue bus" problem noted by Debreu (1960). Note, however, that in the dynamic logit model the IIA property does not hold: the log-odds of

[26] The constant γ in (3.18) is Euler's constant, which shifts the extreme value distribution so it has unconditional mean zero.

[27] Closed-form solutions for the conditional choice probability are available for the larger family of multivariate extreme-value distributions [McFadden (1977)]. This family is characterized by the property that it is *max-stable*, i.e. it is closed under the operation of maximization. Dagsvik (1991) showed that this class is dense in the space of all distributions for ε in the sense that the conditinal choice probabilities for an arbitrary density q can be approximated arbitrarily closely by the choice probability for some multivariate extreme-value distribution.

choosing d over 1 equals the difference in the value functions $v(x,d) - v(x,1)$, but from the definition of $v(x,d)$ in (3.22) we see that it generally depends on the attributes of all of the other alternatives even when the single period utility function depends only on the attributes of the chosen alternative, $u(x,d) = u(x_d)$. Thus, the dynamic logit model benefits from the computational simplifications of the extreme-value specification but avoids the IIA problem of static logit models.

Although Theorems 3.2 and 3.3 appear to apply only to infinite-horizon stationary DDP problems, they actually include finite-horizon, non-stationary DDP problems as a special case. To see this, let the time index t be an additional component of x_t, and assume that the process enters an absorbing state with $u_t(x_t, d_t) \equiv u(x_t, t, d_t) = 0$ for $t > T$. Then Theorems 3.2 and 3.3 continue to hold, with the exception that δ, P, G, π and v all depend on t. The value functions $v_t, t = 1, \ldots, T$ are given by the same backward recursion formulas as in the finite-horizon MDP models described in Section 2:

$$
\begin{aligned}
&v_T(x,d) = u_T(x,d), \\
&v_t(x,d) = u_t(x,d) + \beta \int G_t[\{v_{t+1}(y,d'), d' \in D(y)\} \mid y] \pi_t(\mathrm{d}y \mid x,d).
\end{aligned}
\tag{3.24}
$$

Substituting these value functions into (3.18), we obtain choice probabilities P_t that depend on time. It is easy to see that the process $\{x_t, d_t\}$ is still Markovian, but with non-stationary transition probabilities.

Given panel data $\{x_t^a, d_t^a\}$ on observed states and decisions of a collection of individuals, the full information maximum likelihood estimator $\hat{\theta}^f$ is defined by

$$
\hat{\theta}^f = \underset{\theta \in R^N}{\operatorname{argmax}}\, L^f(\theta) \equiv \prod_{a=1}^{A} \prod_{t=1}^{T_a} P(d_t^a \mid x_t^a, \theta) \pi(x_t^a \mid x_{t-1}^a, d_{t-1}^a, \theta).
\tag{3.25}
$$

Maximum likelihood estimation is complicated by the fact that even in cases where the conditional choice probability has a closed-form solution in terms of the value functions v_θ, the latter function does not have an a priori known functional form and is only implicitly defined by the fixed point condition (3.16). Rust (1987, 1988b) developed a nested fixed point algorithm for estimating θ: an "inner" contraction fixed point algorithm computes v_θ for each trial value of θ, and an "outer" hill-climbing algorithm searches for the value of θ that maximizes L^f.

In practice, θ can be estimated by a simpler 2-stage procedure that yields consistent, asymptotically normal but inefficient estimates of θ^*, and a 3-stage procedure which is asymptotically equivalent to full information maximum likelihood. Suppose we partition θ into two components (θ_1, θ_2), where θ_1 is a subvector of parameters that appear only in π and θ_2 is a subvector of parameters that appear only in (u, q, β). In the first stage we estimate θ_1 using the partial likelihood estimator $\hat{\theta}_1^p$,[28]

[28] Cox (1975) has shown that under standard regularity conditions, the partial likelihood estimator will be consistent and asymptotically normally distributed.

$$\hat{\theta}_1^p = \underset{\theta_1 \in R^{N_1}}{\operatorname{argmax}}\, L_1^p(\theta_1) \equiv \prod_{a=1}^{A} \prod_{t=1}^{T_a} \pi(\mathrm{d}x_t^a | x_{t-1}^a, d_{t-1}^a, \theta_1). \tag{3.26}$$

Note that the first stage does not require a nested fixed point algorithm to solve the DDP problem. In the second stage we estimate the remaining parameters using the partial likelihood estimator $\hat{\theta}_2^p$ defined by

$$\hat{\theta}_2^p = \underset{\theta_2 \in R^{N_2}}{\operatorname{argmax}}\, L_2^p(\hat{\theta}_1^p, \theta_2) \equiv \prod_{a=1}^{A} \prod_{t=0}^{T_a} P(d_t^a | x_t^a, \hat{\theta}_1^p, \theta_2). \tag{3.27}$$

The second stage treats the consistent first stage estimates of $\pi(\mathrm{d}x_{t+1} | x_t, d_t, \hat{\theta}_1^p)$ as the "truth", reducing the problem to estimating the remaining parameters θ_2 of (u, q, β). It is well known that for any optimization method the number of likelihood function evaluations needed to find a maximum increases rapidly with the number of parameters being estimated. Since the second stage estimation requires a nested fixed point algorithm to solve the DDP problem at each likelihood function evaluation, any reduction in the number of parameters being estimated can lead to substantial computational savings.

Note that, due to the presence of estimation error in the first stage estimate of $\hat{\theta}_1^p$, the covariance matrix formed by inverting the information matrix for the partial likelihood function (3.27) will be inconsistent. Although there is a standard correction formula that yields a consistent estimate of the covariance matrix [Amemiya (1976)], in practice it is just as simple to use the consistent estimates $\hat{\theta}^p = (\hat{\theta}_1^p, \hat{\theta}_2^p)$ from stages 1 and 2 as starting values for one or more "Newton steps" on the full likelihood function (3.25):

$$\hat{\theta}^n = \hat{\theta}^p - \gamma \hat{S}(\hat{\theta}^p), \tag{3.28}$$

where the "search direction" $\hat{S}(\hat{\theta}^p)$ is given by

$$\hat{S}(\hat{\theta}^p) = -[\partial^2 \log L^f(\hat{\theta}^p)/\partial\theta\partial\theta']^{-1}[\partial \log L^f(\hat{\theta}^p)/\partial\theta] \tag{3.29}$$

and $\gamma > 0$ is a step-size parameter. Ordinarily the step size γ is set equal to 1, but one can also choose γ to maximize L^f without changing the asymptotic properties of $\hat{\theta}^n$. Using the well-known "information equality" we can obtain an alternative asymptotically equivalent version of (3.29) by replacing the Hessian matrix with the negative of the information matrix $\hat{I}(\hat{\theta}^p)$ defined by

$$\begin{aligned}\hat{I}(\hat{\theta}^p) = \sum_{a=1}^{A} \Bigg[& \sum_{t=1}^{T} \partial \log P(d_t^a | x_t^a, \hat{\theta}^p) \pi(x_{t-1}^a | x_{t-1}^a, d_{t-1}^a, \hat{\theta}^p)/\partial\theta \\ & \times \sum_{t=1}^{T} \partial \log P(d_t^a | x_t^a, \hat{\theta}^p) \pi(x_{t-1}^a | x_{t-1}^a, d_{t-1}^a, \hat{\theta}^p)/\partial\theta' \Bigg].\end{aligned} \tag{3.30}$$

We call $\hat{\theta}^n$ the Newton-step estimator. It is straightforward to show that this procedure results in parameter estimates that are asymptotically equivalent to full information maximum likelihood and, as a by-product, consistent estimates of the asymptotic covariance matrix $\hat{I}^{-1}(\hat{\theta}^f)$.[29]

The feasibility of the nested fixed point maximum likelihood procedure depends on our ability to rapidly compute the fixed point v_θ, for any given value of θ, and to find the maximum of L^f or L^p in as few likelihood function evaluations as possible. At a minimum, the likelihood function should be a smooth function of θ so that more efficient gradient optimization algorithms can be employed. The smoothness of the likelihood is also crucial for establishing the large sample properties of the maximum likelihood estimator. Since the primitives (u, p, β) are specified a priori, they can be chosen to be smooth functions of θ. The convexity of the social surplus function implies that the conditional choice probabilities are smooth functions of v_θ. Therefore the question of smoothness further reduces to finding sufficient conditions under which $\theta \to v_\theta$ is a smooth mapping from R^N into B. This follows from the implicit function theorem since the pair (θ, v_θ) is a zero of the nonlinear operator $F: R^N \times B \to B$ defined by

$$0 = F(v, \theta) \equiv (I - \Psi_\theta)(v). \tag{3.31}$$

Theorem 3.4

Under regularity conditions (A1) to (A13) given in Rust (1988b), $\partial v / \partial \theta$ exists and is a continuous function of θ given by

$$\frac{\partial v_\theta}{\partial \theta} = [I - \Psi'_\theta(v)]^{-1} \left[\frac{\partial \Psi'_\theta(v)}{\partial \theta} \right] \Bigg|_{v = v_\theta}. \tag{3.32}$$

The successive approximation/Newton iteration polyalgorithm described in Section 2.5 can be used to compute v_θ. Since Newton's algorithm involves inverting the operator $[I - \Psi'_\theta(v)]$, it follows that one can use it to compute $\partial v_\theta / \partial \theta$ using formula (3.32) at negligible marginal cost.

Once we have the derivatives $\partial v_\theta / \partial \theta$, it is a straightforward exercise to compute the derivatives of the likelihood, $\partial L^f / \partial \theta$. This allows us to employ more efficient quasi-Newton gradient optimization algorithms to search for the maximum likelihood estimate, $\hat{\theta}^f$. Some of these methods require second derivatives of the likelihood function, which are significantly harder to compute. However, the information equality implies that the information matrix (3.30) is a good approximation to the negative of the Hessian of L^f in large samples. This idea forms the basis of the BHHH optimization algorithm [Berndt, Hall, Hall and Hausman (1974)] which only

[29] In practice, the two-stage estimator $\hat{\theta}^p$ may be sufficiently far away from the maximum likelihood estimates that several Newton steps (3.28) are necessary. In this case, the Newton-step estimator is simply a way generating values for computing the full information maximum likelihood estimates in (3.25). Also, we haven't attempted to correct the estimated standard errors for possible misspecification as in White (1982) due to the fact that such corrections require second derivatives of the likelihood function which are difficult to compute in DDP models.

requires first derivatives of the likelihood function.[30] The nested fixed point algorithm combines the successive approximation/Newton iteration polyalgorithm and the BHHH/Broyden gradient optimization algorithm in order to obtain an efficient and numerically stable method for computing the maximum likelihood estimate $\hat{\theta}$.[31]

In order to derive the asymptotic properties of the maximum likelihood estimators $\hat{\theta}^i$, $i = f, p, n$, we need to make some additional assumption about the sampling process. First, we assume that the periods at which agents are observed coincide with the periods at which they make their decisions. In practice agents do not make decisions at exogenously spaced intervals of time, so it is unlikely that the particular points in time at which agents are interviewed coincide with the times they make their decisions. One way to deal with the problem is to use retrospective data on decisions made between survey dates. In order to minimize problems of time aggregation one should in principle formulate a sufficiently fine-grained model with "null actions" that allow one to model decision-making processes with randomly varying times between decisions. However if the DDP model has a significantly shorter decision interval than the observation interval, we may face the problem that the data set may not contain observations on the agent's intervening states and decisions. In principle, this problem can be solved by using a partial likelihood function that omits the intervening periods, or a full likelihood function that "integrates out" the unobserved states and decisions in the intervening periods. The practical limitation of this approach is the "curse of dimensionality" of solving very fine-grained DP models.

Next, we need to make some assumptions about the dependence between the realizations $\{x_t^a, d_t^a\}$ and $\{x_t^b, d_t^b\}$ for agents $a \neq b$. The standard assumption is that these realizations are independent, but this may not be plausible in models where agents are affected by "macroeconomic shocks" (examples of such shocks include price, unemployment rates and news announcements). We assume that the observed state variable can be partitioned into two components, $x_t = (m_t, z_t)$, where m_t represents a macroeconomic shock that is common to all agents and z_t represents an idiosyncratic component that is independently distributed across agents conditional on the realization of $\{m_t\}$. Sufficient conditions for such independence are given in the three assumptions below.

Assumption ***CI–X***

The transition probability for the observed state variable $x_t = (m_t, z_t)$ is given by

$$\pi(\mathrm{d}x_{t+1} \mid x_t, d_t) = \pi_1(\mathrm{d}z_{t+1} \mid z_t, m_t, d_t)\pi_2(\mathrm{d}m_{t+1} \mid m_t). \tag{3.33}$$

[30] Convergence of the BHHH method in small samples can be accelerated by Broyden and Davidon Fletcher Powell updating procedures that adaptively improve the accuracy of the information matrix approximation to the Hessian of L^f. The method also applies to maximization of the partial likelihood function L^p.

[31] A documented example of the algorithm written in the Gauss programming language is available from the author upon request.

Assumption ***SI–E***

For each $t \geqslant 0$ the distributions of the unobserved state variable ε_t^a are conditionally independent:

$$\Pr(d\varepsilon_t^1, \ldots, d\varepsilon_t^A | x_t^1, \ldots, x_t^A) = \prod_{a=1}^{A} q(d\varepsilon_t^a | x_t^a). \tag{3.34}$$

Assumption ***SI–Z***

For each $t \geqslant 1$ the transition probabilities for the idiosyncratic components z_t^a of the observed state variable x_t^a are conditionally independent:

$$\Pr(dz_{t+1}^1, \ldots, dz_t^A | x_t^1, \ldots, x_t^A, d_t^1, \ldots, d_t^A) = \prod_{a=1}^{A} \pi_1(dz_{t+1}^a | z_t^a, m_t, d_t^a) \tag{3.35}$$

and, when $t = 0$, the initial distributions of z_0^a are independent, conditional on m_0:

$$\Pr(dz_0^1, \ldots, dz_0^A | m_0) = \prod_{a=1}^{A} \pi_1(dz_0^a | m_0). \tag{3.36}$$

Assumption CI–X is an additional conditional independence assumption imposed when the observed state variable x_t includes macroeconomic shocks. It corresponds to an asymmetry that seems reasonable when individual agents are small relative to the economy: macroeconomic shocks can affect the evolution of agents' idiosyncratic states $\{z_t^a, \varepsilon_t^a\}$, but an individual's decision d_t^a has no effect on the evolution of the $\{m_t\}$ process, modelled as an exogenous Markov process with transition probability π_2.[32] SI–E and SI–Z require that any correlation between agents' idiosyncratic states $\{z_t^a, \varepsilon_t^a\}$ is a result of the common macroeconomic shocks. Together with assumption CI–X, these conditions imply that realizations $\{z_t^a, d_t^a\}$ and $\{z_t^b, d_t^b\}$ are independent conditional on $\{m_t\}$.

A final assumption is needed about relative sizes of time-series and cross-sectional dimensions of the data set. There are three cases to consider: (1) the number of time-series observations for each agent is fixed and $A \to \infty$, (2) the number of cross-section observations is fixed and $T_a \to \infty$, or (3) both A and T_a tend to infinity. In most panel data sets the cross-sectional dimension is much larger relative to the time-series dimension, so we will focus on case 1. If we further assume that the observation period T_a is fixed at a common value T for all agents a, then it is straightforward to show that, conditional on $\{m_0, \ldots, m_T\}$, the realizations of $\{x_t^a, d_t^a\}$ are IID. This allows us to use the simpler IID strong law of large numbers and Lindeberg–Levy central limit theorem to establish the consistency and asymptotic normality of $\hat{\theta}^i$, $i = p, f, n$, requiring only continuous second derivatives of the likelihood.[33]

[32] Note that while $\{m_t\}$ is independent of the decisions of individual agents, it will generally not be independent of their collective behavior. Thus, CI–X is not inconsistent with dynamic general equilibrium in large economies.

[33] In cases where the observations are INID or weakly dependent, somewhat stronger smoothness and boundedness conditions are required to establish the asymptotic properties of the MLE.

Theorem 3.5

Under assumptions AS, CI, BU, WC, BE, CI–X, SI–E and SI–Z and regularity conditions (A10) to (A35) of Rust (1988b), the full information and 3-stage maximum likelihood estimators $\hat{\theta}^i$, $i = f, n$ satisfy

1. $\hat{\theta}^i$ is a well-defined random variable,
2. $\hat{\theta}^i$ converges to θ^* with probability 1 as $A \to \infty$,
3. the distribution of $\sqrt{A}(\hat{\theta}^i - \theta^*)$ converges weakly to $N[0, I(\theta^*)^{-1}]$ where the information matrix $I(\theta^*)$ is given by

$$
\begin{aligned}
I(\theta^*) = & -E\left\{\sum_{t=1}^{T} \partial^2 \log P(\tilde{d}_t|\tilde{z}_t, m_t, \theta^*)/\partial\theta\partial\theta' \middle| (m_0, \dots, m_T)\right\} \\
& -E\left\{\sum_{t=1}^{T}{}' \partial^2 \log \pi_1(\tilde{z}_t|\tilde{z}_{t-1}, m_{t-1}, \tilde{d}_{t-1}, \theta^*)/\partial\theta\partial\theta' \middle| (m_0, \dots, m_T)\right\} \\
& -\left\{\sum_{t=1}^{T} \partial \log \pi_2(m_t|m_{t-1}, \theta^*)/\partial\theta\partial\theta'\right\} \\
= & \, E\left\{\sum_{t=1}^{T} \partial \log P(\tilde{d}_t|\tilde{z}_t, m_t, \theta^*)/\partial\theta \sum_{t=1}^{T} \partial \log P(\tilde{d}_t|\tilde{z}_t, m_t, \theta^*)/\partial\theta' \middle| (m_0, \dots, m_T)\right\} \\
& + E\left\{\sum_{t=1}^{T} \partial \log \pi_1(\tilde{z}_t|\tilde{z}_{t-1}, m_{t-1}, \theta^*)/\partial\theta \right. \\
& \left. \times \sum_{t=1}^{T} \partial \log \pi_1(\tilde{z}_t|\tilde{z}_{t-1}, m_{t-1}, \theta^*)/\partial\theta' \middle| (m_0, \dots, m_T)\right\} \\
& + 2E\left\{\sum_{t=1}^{T} \partial \log P(\tilde{d}_t|\tilde{z}_t, m_t, \theta^*)/\partial\theta \right. \\
& \left. \times \sum_{t=1}^{T} \partial \log \pi_1(\tilde{z}_t|\tilde{z}_{t-1}, m_{t-1}, \theta^*)/\partial\theta' \middle| (m_0, \dots, m_T)\right\} \\
& + 2E \sum_{t=1}^{T} \partial \log P(\tilde{d}_t|\tilde{z}_t, m_t, \theta^*)/\partial\theta \sum_{t=1}^{T} \partial \log \pi_2(m_t|m_{t-1}, \theta^*)/\partial\theta' \middle| (m_0, \dots, m_T)\right\} \\
& + 2E\left\{\sum_{t=1}^{T} \partial \log \pi_1(\tilde{z}_t|\tilde{z}_{t-1}, m_{t-1}, \theta^*)/\partial\theta \right. \\
& \left. \times \sum_{t=1}^{T} \partial \log \pi_2(m_t|m_{t-1}, \theta^*)/\partial\theta' \middle| (m_0, \dots, m_T)\right\} \\
& -\left\{\sum_{t=1}^{T} \partial \log \pi_2(m_t|m_{t-1}, \theta^*)/\partial\theta\partial\theta'\right\}.
\end{aligned}
\tag{3.37}
$$

Theorem 3.6

Under assumptions AS, CI, BU, WC, BE, CI–X, SI–E and SI–Z and regularity conditions (A10) to (A35) of Rust (1988b), the 2-stage partial likelihood estimator $\hat{\theta}^p = (\hat{\theta}_1^p, \hat{\theta}_2^p)'$ satisfies

1. $\hat{\theta}^p$ is a well-defined random variable,
2. $\hat{\theta}^p$ converges to θ^* with probability 1 as $A \to \infty$,
3. the distribution of $\sqrt{A}(\hat{\theta}^p - \theta^*)$ converges weakly to $N(0, \Sigma)$ where Σ is given by

$$\Sigma = \Delta^{-1} \Omega \Delta'^{-1}, \tag{3.38}$$

where Δ and Ω are given by

$$\Delta = \begin{bmatrix} \Delta_{11} & 0 \\ \Delta_{21} & \Delta_{22} \end{bmatrix} \qquad \Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{12}' & \Omega_{22} \end{bmatrix}, \tag{3.39}$$

where

$$\begin{aligned}
\Delta_{11} &= E\left\{ \sum_{t=1}^{T} \partial^2 \log \pi_1(\tilde{z}_t | \tilde{z}_{t-1}, m_{t-1}, \tilde{d}_{t-1}, \theta_1^*) \pi_2(m_t | m_{t-1}, \theta_1^*) / \right. \\
&\qquad \left. \partial \theta_1 \partial \theta_1' \,\middle|\, (m_0, \ldots, m_T) \right\}, \\
\Delta_{21} &= E\left\{ \sum_{t=1}^{T} \partial^2 P(\tilde{d}_t | \tilde{x}_t, \theta_1^*, \theta_2^*) / \partial \theta_2 \partial \theta_1' \,\middle|\, (m_0, \ldots, m_T) \right\}, \\
\Delta_{22} &= E\left\{ \sum_{t=1}^{T} \partial^2 P(\tilde{d}_t | \tilde{x}_t, \theta_1^*, \theta_2^*) / \partial \theta_2 \partial \theta_2' \,\middle|\, (m_0, \ldots, m_T) \right\}, \\
\Omega_{11} &= E\left\{ \sum_{t=1}^{T} \partial \log \pi_1(\tilde{z}_t | \tilde{z}_{t-1}, m_{t-1}, \tilde{d}_{t-1}, \theta_1^*) \pi_2(m_t | m_{t-1}, \theta_1^*) / \partial \theta_1 \right. \\
&\qquad \left. \times \sum_{t=1}^{T} \partial \log \pi_1(\tilde{z}_t | \tilde{z}_{t-1}, m_{t-1}, \tilde{d}_{t-1}, \theta_1^*) \pi_2(m_t | m_{t-1}, \theta_1^*) / \partial \theta_1' \,\middle|\, (m_0, \ldots, m_T) \right\}, \\
\Omega_{12} &= E\left\{ \sum_{t=1}^{T} \partial \log \pi_1(\tilde{z}_t | \tilde{z}_{t-1}, m_{t-1}, \tilde{d}_{t-1}, \theta_1^*) \pi_2(m_t | m_{t-1}, \theta_1^*) / \partial \theta_1 \right. \\
&\qquad \left. \times \sum_{t=1}^{T} \partial P(\tilde{d}_t | \tilde{x}_t, \theta_1^*, \theta_2^*) / \partial \theta_2' \,\middle|\, (m_0, \ldots, m_T) \right\}, \\
\Omega_{22} &= E\left\{ \sum_{t=1}^{T} \partial P(\tilde{d}_t | \tilde{x}_t, \theta_1^*, \theta_2^*) / \partial \theta_2 \sum_{t=1}^{T} \partial P(\tilde{d}_t | \tilde{x}_t, \theta_1^*, \theta_2^*) / \partial \theta_2' \,\middle|\, (m_0, \ldots, m_T) \right\}.
\end{aligned} \tag{3.40}$$

In addition to the distribution of the parameter vector $\hat{\theta}^i$, we are often interested in the distribution of utility and value functions u_θ and v_θ treated as random elements of B. This is useful, for example, in computing uniform confidence bands for these functions. The following result is essentially a Banach space version of the "delta theorem".

Theorem 3.7

If $\hat{\theta}$ converges with probability 1 to θ^* and $\sqrt{A}[\hat{\theta}-\theta^*]$ converges weakly to $N(0,\Sigma)$, and v_θ is a smooth function of θ, then

1. $v_{\hat{\theta}}$ is a B-valued random element,
2. $v_{\hat{\theta}}$ converges with probability 1 to v_{θ^*},
3. the distribution of $\sqrt{A}[v_{\hat{\theta}}-v_{\theta^*}]$ converges weakly to a Gaussian random element of B with mean 0 and covariance operator $[\partial v_{\theta^*}/\partial\theta]\Sigma[\partial v_{\theta^*}/\partial\theta']$.

We conclude this section with some comments on specification and hypothesis testing. Since structural models are often highly simplified and tightly parameterized, one has an obligation to go beyond simply reporting parameter estimates and their standard errors. If the structural model is to have any credibility, it needs to be subjected to a battery of in-sample specification tests, and if possible, out-of-sample predictive tests (a good example of the latter type of test is presented in Section 4.2). The maximum likelihood framework allows formulation of standard "holy trinity" test statistics, the Wald, Likelihood Ratio, and Lagrange Multiplier (LM) tests [see Engle (1984)]. Examples 1 and 2 below show how these statistics can be used to test the validity of various functional-form restrictions on the DDP model. Example 3 discusses the chi-square goodness-of-fit statistic, which is perhaps the most useful omnibus test of the correctness of the DDP specification.

Example 1

The holy trinity can be used to conduct a "heterogeneity test" of the null hypothesis that the parameter vector θ^* is the same for two or more subgroups of agents. If there are K subgroups, we can formulate the null hypothesis as $K-1$ linear restrictions on a KN-dimensional full parameter vector $(\theta_1,\ldots,\theta_K)$ where θ_k is the N-dimensional parameter vector for subgroup k. The likelihood ratio test involves computing -2 times the difference in the restricted and unrestricted log-likelihood functions, where we compute the restricted log-likelihood by pooling all K subgroups and estimating a single N-dimensional parameter vector θ. The Wald test statistic is a quadratic form in the $K-1$ differences in the group-specific coefficient estimates, $\hat{\theta}_{K+1}-\hat{\theta}_K$, $k=1,\ldots,K-1$. In this case the LM statistic is the easiest to compute since it only requires computation of a single N-dimensional parameter estimate $\hat{\theta}$ for the pooled sample under the null hypothesis of no heterogeneity. The LM statistic tests whether the score of the likelihood function is approximately zero for all K subgroups. All three test statistics have an asymptotic chi-square distribution under the null, with degrees of freedom equal to the number of restrictions being tested. In the example, there are $(K-1)N$ degrees of freedom. Computation of the Wald and LM statistics requires an estimate of the information matrix, $\hat{I}(\hat{\theta}_k)$ for each of the subgroups $k = 1, \ldots, K$.

Example 2

The holy trinity can be used to test the validity of the conditional independence assumption CI. Recall that CI implies that the unobserved state variable ε_t is independent of ε_{t-1} and is conditional on the value x_t of the observed state variable. This is a strong restriction although, as we will see in Section 4.6, it seems to be necessary to obtain a computationally tractable estimation algorithm. A natural way to test CI is to add some function f of the previous period control variables to the current period value function: $v_\theta(x_t, d_t) + \alpha f(d_{t-1})$. Under the null hypothesis that CI is valid, the decision taken in period $t-1$ will have no effect on the decision made in period t once we condition on x_t since $\{v_\theta(x_t, d), d \in D(x_t)\}$ constitutes a set of "sufficient statistics" for the agent's decision in period t. Thus, $\alpha = 0$ under the null hypothesis that CI holds. However under the alternative hypothesis that CI doesn't hold, ε_t and ε_{t-1} will be serially correlated, even conditional, on x_t, so that d_{t-1} will generally be useful for predicting the agent's choice d_t. Thus, $\alpha \neq 0$ under the alternative hypothesis. The Wald, Likelihood Ratio or LM statistics can be used to test the hypothesis that $\alpha = 0$. For example, the Wald statistic is simply $A\hat{\alpha}^2/\hat{\sigma}^2(\hat{\alpha})$, where $\hat{\sigma}^2(\hat{\alpha})$ is the asymptotic variance of $\hat{\alpha}$.

Example 3

The chi-square goodness-of-fit statistic provides an overall test of the null hypothesis that an econometric model is correctly specified (i.e. that the parametric model coincides with the "true" data generating process). In the case of a DDP model, this amounts to a test of the joint hypotheses that (1) agents are rational, i.e. they act "as if" their behavior is governed by an optimal decision rule from some DDP model, and (2) the particular parametric specification of this model is correct. However, the analysis of the identification problem in Section 3.5 reveals that the hypothesis of rationality per se imposes essentially no testable restrictions: the empirical content of the theory arises from additional restrictions on the primitives (u, p, β) of the DDP model. In this sense, testing the theory is tantamount to testing the econometrician's assumptions about the parametric functional form of (u, p, β). Although there are other omnibus specification tests [such as White's (1982) "information matrix test"] the chi-square goodness-of-fit test is far more useful in diagnosing the source of specification errors. There are two versions of this statistic, one for models without covariates [Pollard (1979)], and one for models with covariates [Andrews (1988, 1989)].[34] The former is useful for testing complete realizations of

[34] Strictly speaking, Pollard's results are only applicable if the full likelihood includes the probability density of the initial state x_0. Otherwise the full likelihood (3.25) can be analyzed as a conditional likelihood using Andrews's analysis of chi-square tests of parametric models with covariates.

the controlled process $\{x_t, d_t\}$ using the full Markov transition probability $P(d_{t+1} | x_{t+1}, \theta)\pi(dx_{t+1} | x_t, d_t, \theta)$ derived in Theorem 3.3, whereas the version with covariates is useful for testing the conditional choice probability $P(d_t | x_t, \theta)$ and the transition probability $\pi(dx_{t+1} | x_t, d_t, \theta)$. Both formulations are based on a partition of the relevant space of realizations of x_t and d_t. In the case of the full likelihood function, the relevant space is $X^T \times D^T$, and in the case of P or π it is $X \times D$ or $X \times X \times D$, respectively. We partition this space into a fixed number M of mutually exclusive cells. The cells can be randomly chosen, or chosen based on some data-dependent procedure provided (1) the total number of cells M is fixed for all sample sizes, (2) the elements of the partition are members of a Vapnik–Cervonenkis (VC) class, and (3) the partition converges in probability to a fixed, non-stochastic partition whose elements are also members of a VC class.[35] If we let Ω denote the $M \times 1$ vector of elements of this partition $\Omega = (\Omega_1, \ldots, \Omega_M)$, we can define a vector $\lambda(\Omega, \theta)$ of differences between sample frequencies of Ω and the predicted probabilities of the DDP model with parameter θ. In a chi-square test of the specification of the conditional choice probability P, the ith element of $\lambda(\Omega, \theta)$ is given by

$$\lambda_i(\Omega, \theta) = \frac{1}{A} \sum_{a=1}^{A} I\{(d_t^a, x_t^a) \in \Omega_i\} - \frac{1}{A} \sum_{a=1}^{A} \sum_{d \in D(x_t^a)} I\{(d, x_t^a) \in \Omega_i\} P(d | x_t^a, \theta). \tag{3.41}$$

The first term in (3.41) is the sample proportion of (x_t, d_t) pairs falling into partition element Ω_i whereas the second term is the DDP model's prediction of the probability that $(x_t, d_t) \in \Omega_i$. Note that since the law of motion for x_t is not specified, we simply average over all sample points x_t^a. An analogous formula holds for the case of chi-square tests of the full likelihood function: in that case $\lambda(\Omega_i, \theta)$ is the difference between the sample fraction of paths $\{x_t^a, d_t^a\}$ falling in partition element Ω_i less the probability that these realizations fall in Ω_i, computed by integrating with respect to the probability measure on $X^T \times D^T$ generated by the controlled transition probability $P(d_{t+1} | x_{t+1}, \theta)\pi(dx_{t+1} | x_t, d_t, \theta)$. The chi-square test statistic is given by

$$\chi^2(\Omega, \hat{\theta}) = \lambda(\Omega, \hat{\theta})' \hat{\Sigma}^+ \lambda(\Omega, \hat{\theta}), \tag{3.42}$$

where $\hat{\Sigma}^+$ is a generalized inverse of the asymptotic covariance matrix $\hat{\Sigma}$ (which is generally singular). Andrews (1989) showed that under the null hypothesis of correct specification, $\lambda(\Omega, \hat{\theta})$ converges in distribution to $N(0, \Sigma)$, which implies that $\chi^2(\Omega, \hat{\theta})$ converges in distribution to a chi-square random variable whose degrees of freedom equal the rank of Σ. Andrews provides formulas for $\hat{\Sigma}$ that take relatively simple forms when $\hat{\theta}$ is an asymptotically efficient estimator such as in the case of the full information maximum likelihood estimator $\hat{\theta}^f$. A natural strategy is to start with a chi-square test of the specification of the full likelihood (3.25). If this is rejected, one can then do a chi-square test of π to see if the rejection is due to a misspecification of agents' beliefs. If this is not rejected, then we have an

[35] See Pollard (1984) for a definition of a VC class.

indication that the rejection of the model is due to a misspecification of preferences $\{\beta, u\}$ or that the CI assumption is not valid. A further test of the CI assumption using one of the holy trinity specification tests described in Example 2 can be used to determine if CI is the source of the problem. If not, an investigation of the cells which have the largest prediction errors (3.41) may provide valuable insights on the source of specification errors in u or β.

3.3. *Alternative estimation methods: Finite-horizon DDP problems*

Although the previous section showed that the AS–CI assumption leads to a simple and general estimation theory for DDP's, it would be desirable to develop estimation methods that can relax these assumptions. This section considers estimation methods for DDP models with unobservables that are serially correlated and enter the utility function in a non-separable fashion. Unfortunately there is no general estimation theory for this class of models at present. Instead there are a variety of problem-specific specifications and estimation methods, most of which are designed for finite-horizon binary choice problems. As we will see, there are substantial theoretical and computational obstacles to developing an estimation theory for a more general class of DDP models that relaxes AS–CI. This section presents two examples that illustrate the successful approaches.[36]

Example 1

Wolpin (1984) pioneered the nested numerical solution method for a class of binary choice models where unobservables enter the utility function $u(x, \varepsilon, d)$ in a non-additive fashion. In a binary choice model, one does not need a 2-dimensional vector of unobservables to yield a saturated choice model: in Wolpin's specification ε is uni-dimensional and interacts with the observed state variable x_t, yet the choice probability satisfies $P_\theta(1|x) > 0$ for all x and is continuously differentiable in θ. Wolpin's application concerned a Malaysian family's choice of whether or not to conceive a child in period t: $d_t = 1$ if a child is born in year t, $d_t = 0$, otherwise.[37] Wolpin used the following quadratic family utility function,

$$u_\theta(x_t, \varepsilon_t, d_t) = (\theta_1 + \varepsilon_t)n_t - \theta_2 n_t^2 + \theta_3 c_t - \theta c_t^2, \tag{3.43}$$

where $x_t = (n_t, c_t)$ and n_t is the number of children in the family and c_t is total family consumption (treated as an exogenous state variable rather than as a control variable). Assuming that $\{\varepsilon_t\}$ is an IID Gaussian process with mean 0 and variance σ^2 and that $\{x_t\}$ is a Markov process independent of $\{\varepsilon_t\}$, the family's optimal

[36] For additional examples, see the survey by Eckstein and Wolpin (1989a).

[37] Wolpin's model thus assumed that there is no uncertainty about fertility and that contraception is 100% effective.

decision rule $d_t = \delta_t(x_t, \varepsilon_t)$ can be computed by backward induction starting in the final period T at which the family is fertile. It is not difficult to see that (3.43) implies δ_t is given by the threshold rule

$$\delta_t(x_t, \varepsilon_t) = \begin{cases} 1 & \text{if} \quad \varepsilon_t > \eta_t(x_t, \theta) \\ 0 & \text{if} \quad \varepsilon_t \leqslant \eta_t(x_t, \theta). \end{cases} \tag{3.44}$$

The cutoffs $\{\eta_t(x, \theta)\}$ define the value of ε_t such that the family is indifferent as to whether or not to have another child: for any given value of θ and for each possible t and x, the cutoffs can be computed by backward induction. The assumption that $\{\varepsilon_t\}$ is IID $N(0, \sigma^2)$ then implies that the conditional choice probability is given by

$$P_t(d_t | x_t, \theta) = \begin{cases} 1 - \Phi[\eta_t(x_t, \theta)/\sigma] & \text{if} \quad d = 1 \\ \Phi[\eta_t(x_t, \theta)/\sigma] & \text{if} \quad d = 0, \end{cases} \tag{3.45}$$

where Φ is the standard normal CDF. One can estimate θ maximum likelihood using the partial, full or Newton-step maximum likelihood estimators given in equations (3.25) to (3.27) of Section 3.2. Using the implicit function theorem one can show that the cutoffs are smooth functions of θ. From (3.45) it is clear that this implies that the likelihood function is a smooth function of θ, allowing Wolpin to establish that his maximum likelihood estimator has standard asymptotic properties.

Example 2

Although Example 1 shows how one can incorporate measurement errors and unobservables that enter the utility function non-separably, it relies on the independence of the $\{\varepsilon_t\}$ process, a strong form of CI. Pakes (1986) developed a method for estimating binary DDP's with serially correlated observables that enter u additively. Pakes developed an optimal stopping model of whether or not to renew a patent. He used Monte Carlo simulations to "integrate out" the serially correlated $\{\varepsilon_t\}$ process in order to evaluate the likelihood function, avoiding a potentially intractable numerical integration problem. Let t denote the number of years a patent has been in force, and let ε_t denote the cash flow accruing to the patent in year t. In Europe, patent holders must pay an annual renewal fee c_t, an increasing function of t. The patent holder must decide whether or not to pay the cost c_t and renew the patent, or let it lapse in which case the patent is permanently cancelled and the idea is assumed to have 0 net value thereafter. In Pakes's data set one cannot observe patent holders' earnings ε_t, so the only observable state variable is $x_t = t$, the number of years the patent is kept in force. Assuming that $\{\varepsilon_t\}$ is a first order Markov process with transition probability $q_t(\mathrm{d}\varepsilon_t | \varepsilon_{t-1})$, Bellman's equation is given by

$$V_T(\varepsilon) = \max\{0, \varepsilon - c_T\},$$

$$V_t(\varepsilon) = \max\left\{0, \varepsilon - c_t + \beta \int V_{t+1}(\varepsilon') q_{t+1}(\mathrm{d}\varepsilon' | \varepsilon)\right\}, \tag{3.46}$$

where T is the statutory limit on patent life. Under fairly general conditions on the transition probabilities $\{q_t\}$ (analogous to the weak continuity conditions of Section 2), one can establish that the optimal decision rule, $d = \delta_t(\varepsilon)$, is given by a threshold rule of the form (3.44), but in this case the cutoffs $\eta_t(\theta)$ only depend on t and the parameter θ. Pakes used a particular family $\{q_t\}$ to simplify the recursive calculation of the cutoffs:

$$q_{t+1}(\mathrm{d}\varepsilon_{t+1}|\varepsilon_t) = \begin{cases} \exp\{-\theta_1\varepsilon_t\} & \text{if} \quad \varepsilon_{t+1} = 0 \\ 0 & \text{if} \quad 0 \leqslant \varepsilon_{t+1} \leqslant \theta_2\varepsilon_t \\ \dfrac{1}{(\theta_4^{t-1}\theta_5)}\exp\left\{\dfrac{-(\theta_3 + \varepsilon_{t+1})}{(\theta_4^{t-1}\theta_5)}\right\} & \text{if} \quad \varepsilon_{t+1} > \theta_2\varepsilon_2. \end{cases} \tag{3.47}$$

The interpretation of (3.47) is that with probability $\exp\{-\theta_1\varepsilon_t\}$ the patent is discovered to be worthless (in which case $\varepsilon_{t+k} = 0$ for all $k \geqslant 1$), or alternatively, the patent is believed to have value next period given by $\varepsilon_{t+1} = \max\{\theta_2\varepsilon_t, z\}$, where z is an IID exponential random variable whose density is given by the third term in (3.47). The likelihood function for this problem requires computation of the probability $\lambda_t(\theta)$ that the patent lapses in year t:

$$\begin{aligned} \lambda_\theta(t) &= \Pr\{\delta_t(\varepsilon_t) = 0, \delta_{t-1}(\varepsilon_{t-1}) = 1, \ldots, \delta_1(\varepsilon_1) = 1\} \\ &= \int_{\varepsilon_t} \cdots \int_{\varepsilon_0} \left[I\{\varepsilon_t < \eta_t(\theta)\} \prod_{s=1}^{t-1} I\{\varepsilon_s \geqslant \eta_s(\theta)\} \right] \prod_{s=1}^{t} q_t(\mathrm{d}\varepsilon_s|\varepsilon_{s-1}, \theta) q_0(\mathrm{d}\varepsilon_0|\theta), \end{aligned} \tag{3.48}$$

where q_0 is the initial distribution of returns at the time the patent was applied for, assumed to be log-normal with parameters (θ_6, θ_7). To establish the asymptotic properties of the maximum likelihood estimator $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_7)$, we need to show that $\lambda_\theta(t)$ is a smooth function of θ. It is straightforward to show that the cutoffs $\eta_t(\theta)$ are smooth functions of θ. Pakes showed that this fact, together with the smoothing that occurs when integrating over the indicator functions in (3.48), yields a likelihood function which is continuously differentiable in θ and has standard asymptotic properties. However, in practice the numerical integrations required in (3.48) become intractable for t larger than 3 or 4, whereas many patents are held for 20 years or more. To overcome the problem Pakes used Monte Carlo integration, calculating a consistent estimate $\hat{\lambda}_t(\theta)$ by simulating a large number of realizations of the process $\{\tilde{\varepsilon}_t\}$ and tabulating the fraction of realizations that lead to drop-outs at year t, i.e. $\tilde{\varepsilon}_t < \eta_t(\theta)$, $\tilde{\varepsilon}_s \geqslant \eta_s(\theta)$, $s = 1, \ldots, t-1$. This requires a "nested simulation algorithm" consisting of an outer optimization algorithm that searches for a value of θ to maximize the likelihood function, and an inner simulation/DP algorithm that

1. solves the DDP problem for a given value of θ, using backward induction to calculate the cutoffs $\{\eta_t(\theta)\}$,

2. simulates NSIM draws of the process $\{\tilde{\varepsilon}_t\}$ to calculate a consistent estimate of $\hat{\lambda}_t(\theta)$ and the full likelihood function $\hat{L}^f(\theta)$

Note that the realizations $\{\tilde{\varepsilon}_t\}$ change each time we update θ. The main burden of Pakes's, simulation estimator arose from having to repeatedly re-simulate NSIM realizations of $\{\varepsilon_t\}$: the recursive calculation of the cutoffs $\{\eta_t(\theta)\}$ took little time in comparison.[38] An additional burden is imposed by the need to calculate numerical derivatives of the likelihood function: each iteration requires 8 separate solutions and simulations of the DDP problem, once to compute $\hat{L}^f(\theta)$ and 7 additional times to compute its partial derivatives with respect to $(\theta_1, \ldots, \theta_7)$.

McFadden (1989) and Pakes and Pollard (1989) subsequently developed a new class of simulation estimators that dramatically lowers the value of NSIM needed to obtain consistent and asymptotically normal parameter estimates. In fact, consistent, asymptotically normal estimates of θ can be obtained with NSIM as small as 1, whereas consistency of Pakes's original simulation estimators required NSIM to tend to infinity with sample size. The new approach uses minimum distance rather than maximum likelihood to estimate θ. In the case of the patent renewal problem, the estimator is defined by

$$\begin{aligned}
&\hat{\theta} = \operatorname{argmin} H_A(\theta)' W_A H_A(\theta), \\
&H_A(\theta) = [n_1/A - \hat{\lambda}_1(\theta), \ldots, n_T/A - \hat{\lambda}_T(\theta)]', \\
&A = \sum_{t=1}^{T} n_t \\
&\hat{\lambda}_t(\theta) = \frac{1}{\text{NSIM}} \sum_{j=1}^{\text{NSIM}} \left[I\{\tilde{\varepsilon}_{tj} < \eta_t(\theta)\} \prod_{s=1}^{t-1} I\{\tilde{\varepsilon}_{sj} \geqslant \eta_s(\theta)\} \right],
\end{aligned} \tag{3.49}$$

where n_t is the number of patent holders who dropped out (failed to renew) after t years, A is the total number of agents (patent holders) in the sample, W_A is a $T \times T$ positive definite weighting matrix, and $\{\tilde{\varepsilon}_{jt}\}$ denotes the jth simulation of the Markov process when the transition probabilities q_t in (3.47) are evaluated at θ. In order to satisfy the uniformity conditions necessary to prove consistency and asymptotic normality of the estimator, it is crucial to note that one *does not* draw independent realizations $\{\tilde{\varepsilon}_{tj}\}$ for distinct values of θ. Instead, at the beginning of estimation we simulate NSIM random vectors $(\tilde{\xi}_1, \ldots, \tilde{\xi}_{\text{NSIM}})$ which are held fixed throughout the entire estimation process, and the jth draw $\tilde{\xi}_j$ is used to construct $\{\tilde{\varepsilon}_{jt}\}$ for each trial value of θ. In the patent renewal problem $\tilde{\xi}_j$ is given by

$$\tilde{\xi}_j = (\tilde{z}, \tilde{\omega}_1, \ldots, \tilde{\omega}_T, \tilde{\gamma}_1, \ldots, \tilde{\gamma}_T), \tag{3.50}$$

[38] Pakes varied NSIM depending on how close the outer algorithm was to converging, setting NSIM equal to a few thousand in the initial stages of the optimization when one is far from convergence and gradually increasing NSIM to 20000 in the final stages of estimation in order to get more precise estimates of the likelihood function and its derivatives.

where $\tilde{z}$ is standard normal, and the $\tilde{\omega}_t$ and $\tilde{\gamma}_t$ are IID draws from an exponential distribution with unit mean. Using ξ_j we can construct a realization of $\{\tilde{\varepsilon}_{jt}\}$ for any value of θ using the recursion

$$\begin{aligned}
&\tilde{\varepsilon}_{j0} = \exp\{\theta_6 + \theta_7 \tilde{z}\},\\
&\tilde{\varepsilon}_{jt} = I\{\theta_1 \tilde{\varepsilon}_{jt-1} \geqslant \tilde{\omega}_t\} \max[\theta_2 \tilde{\varepsilon}_{jt-1}, \theta_4^t \theta_5 \tilde{\gamma}_t - \theta_3].
\end{aligned} \tag{3.51}$$

This simulation process insures that no extraneous simulation error is introduced in the process of minimizing over θ in (3.49), insuring that the uniformity conditions given in Theorems 3.1 and 3.2 of Pakes and Pollard hold. These conditions imply the following asymptotic distribution for $\hat{\theta}$:

$$\begin{aligned}
&\sqrt{A}[\hat{\theta} - \theta^*] \to N(0, \Omega),\\
&\Omega = (1 + \mathrm{NSIM}^{-1})(\Lambda' W \Lambda)^{-1} \Lambda' W \Gamma W' \Lambda (\Lambda' W \Lambda)^{-1},\\
&\Lambda = \partial \lambda(\theta^*)/\partial \theta,\\
&\Gamma = \mathrm{diag}[\lambda(\theta^*)] - \lambda(\theta^*)\lambda(\theta^*)',\\
&\lambda(\theta) = [\lambda_1(\theta), \ldots, \lambda_T(\theta)]'.
\end{aligned} \tag{3.52}$$

By Aitken's theorem [Theil (1971)], the most efficient choice for the weighting matrix is $W = \mathrm{diag}[\lambda(\theta^*)]$. Notice that when one uses the optimal weighting matrix, W_A, the relative efficiency of the simulation estimator increases rapidly to the efficiency of maximum likelihood [which requires exact numerical integration to compute $\lambda(\theta)$]. The standard errors of the simulation estimator are only twice as large as maximum likelihood when NSIM equals one, and are only 10% greater when NSIM equals 10.

While the "frequency simulation estimator" (3.49) substantially reduces the value of NSIM required to estimate θ, it has the drawback that the objective function is a step function in θ, so a non-differentiable search algorithm must be used which typically requires many more function evaluations to find $\hat{\theta}$ than gradient optimization methods. Thus, the frequency simulator will be feasible for problems where we can rapidly solve the DDP problem for the cutoffs $\{\eta_t(\theta)\}$.[39] An important question is whether simulation methods can be used to avoid solving the DDP problem itself. Note that the main burden of DDP estimation methods involves calculation of the value function $V_\theta(s)$, which is itself a conditional expectation with respect to the controlled stochastic process $\{s_t\}$. Can we also use Monte Carlo integration to compute an approximate value function $\hat{V}_\theta(s)$ rather than computing exactly by backward induction? The next section discusses how this might be done in the context of infinite-horizon DDP models.

[39] Recent progress in developing "smooth simulation estimators" [Stern (1989), McFadden (1990)] may help overcome this problem.

3.4. Alternative estimation methods: Infinite-horizon DDP's

Hotz et al. (1993) introduced a simulation estimator for DDP problems that avoids the computational burden of nested numerical solution of the Bellman equation. Unlike the simulation estimators described in Section 3.3, the Hotz et al. estimator is a smooth function of θ. The simulation estimator is based on the following result of Hotz and Miller (1993).

Lemma 3.1

Suppose that $q(d\varepsilon|x)$ has a density with finite first moments and support equal to $R^{|D(x)|}$. Then for each $x \in X$, the choice probabilities given in equation (3.18) define a 1:1 mapping between the space of normalized value functions $\{v \in R^{|D(x)|} | v(1) = 0\}$ and the $|D(x)|$-dimensional simplex $\Delta^{|D(x)|}$.

In the special case where q has an extreme-value distribution, the inverse mapping has a closed-form solution: it is simply the log-odds transformation. The idea behind the simulation estimator is to use non-parametric estimation to obtain consistent estimates of the conditional choice probabilities $\hat{P}(d|x)$ and then invert these estimates to obtain consistent estimates of the normalized value function, $\hat{v}(x,d) - \hat{v}(x,1)$. If we also have estimates of agents' beliefs $\hat{p}$ we can simulate (a consistent estimate of) the controlled stochastic process $\{x_t, \varepsilon_t, d_t\}$. For any given values of x and θ we can use the simulated realizations of $\{x_t, \varepsilon_t, d_t\}$ to construct a simulation estimate of the normalized value function, $\tilde{v}_\theta(x,d) - \tilde{v}_\theta(x,1)$. At the true parameter θ^*, $\tilde{v}_{\theta^*}(x,d) - \tilde{v}_{\theta^*}(x,1)$ is an (asymptotically) unbiased estimate of $v_{\theta^*}(x,d) - v_{\theta^*}(x,1)$. Since the latter quantity can be consistently estimated by inverting non-parametric estimates of the choice probability, the simulation estimator can be roughly described as the value of θ that minimizes the distance between the simulated value function $\tilde{v}_\theta(x,d) - \tilde{v}(x,1)$ and $\hat{v}(x,d) - \hat{v}(x,1)$, averaged over the various (x,d) pairs in the sample.

More precisely, the simulated value (SV) function estimator consists of 5 steps. For simplicity I assume that the observed state variable x has finite support and that there are no unknown parameters in the transition probability π, and that q is an IID extreme-value distribution.

Step 1. Invert non-parametric estimates of $P(d|x)$ to compute consistent estimates of the normalized value functions $\Delta v(x,d) = v(x,d) - v(x,1)$:

$$\Delta\hat{v}(x,d) = \log\left\{\frac{\hat{P}(d|x)}{\hat{P}(1|x)}\right\}. \tag{3.53}$$

Step 2. Use the consistent estimate of $\Delta\hat{v}$ to uncover a consistent estimate of the decision rule δ:

$$\hat{\delta}(x,\varepsilon) = \underset{d \in D(x)}{\operatorname{argmax}}\,[\Delta\hat{v}(x,d) + \varepsilon(d)]. \tag{3.54}$$

Step 3. Using $\hat{\delta}, q$ and π simulate realizations of the controlled stochastic process $\{x_t, \varepsilon_t\}$. Given an initial condition (x_0, d_0), this consists of the following steps:

1. Given (x_{t-1}, d_{t-1}) draw x_t from the transition probability $\pi(\mathrm{d}x_t | x_{t-1}, d_{t-1})$,
2. Given x_t draw ε_t from the density $q(\varepsilon_t | x_t)$,
3. Given (x_t, ε_t) compute $d_t = \hat{\delta}(x_t, \varepsilon_t)$,
4. If $t > S(A)$, stop, otherwise set $t = t + 1$ and return to step 1.[40]

Repeat step 2 using each of the sample points (x_t^a, d_t^a), $a = 1, \ldots, A$, $t = 1, \ldots, T_a$ as initial conditions.

Step 4. Using the simulations $\{\tilde{x}_t, \tilde{d}_t\}$ from step 2, compute the *simulated value function* $\tilde{v}_\theta(x_0, d_0)$:

$$\tilde{v}_\theta(x_0, d_0) = \sum_{t=0}^{S(A)} \beta^t \{u_\theta[\tilde{x}_t, \hat{\delta}(\tilde{x}_t, \tilde{\varepsilon}_t)] + \tilde{\varepsilon}_t[\hat{\delta}(\tilde{x}_t, \tilde{\varepsilon}_t)]\}, \tag{3.55}$$

where (x_0, d_0) is the initial condition from step 2. Repeat step 3 using each of the sample points (x_t^a, d_t^a) as initial conditions.

Step 5. Using the normalized simulated value function $\Delta\tilde{v}_\theta(x_t^a, d_t^a)$ and corresponding non-parametric estimates $\Delta\hat{v}(x_t^a, d_t^a)$ as "data", compute the parameter estimate θ as the solution to

$$\hat{\theta}_A = \operatorname{argmin} H_A(\theta)' W_A H_A(\theta),$$

$$H_A(q) = \frac{1}{A} \sum_{a=1}^{A} \sum_{t=1}^{T_a} [\Delta\hat{v}(x_t^a, d_t^a) - \Delta\tilde{v}_\theta(x_t^a, d_t^a)] Z_t^a, \tag{3.56}$$

where W_A is a $K \times K$ positive definite weighting matrix and K is the dimension of the instrument vector, Z_t^a.

The SV estimator has several attractive features: (1) it avoids the need for repeated calculation of the fixed point (3.22), (2) the simulation error in $\Delta\tilde{v}_\theta(x_t^a, d_t^a)$ averages out over sample observations, so that one can estimate the underlying parameters using as few as one simulated path of $\{\tilde{x}_t, \tilde{\varepsilon}_t\}$ per (x_t^a, d_t^a) observation, (3) since each term $\Delta\tilde{v}_\theta(x, d)$ is a smooth function of θ, the objective function (3.56) is a smooth function of θ allowing use of gradient optimization algorithms to estimate $\hat{\theta}_A$. Note that the simulation in step 2 needs to be done only once at the start of the estimation, so the main computational burden is the repeated evaluation of (3.55) each time θ is updated.

Although the SV estimator is consistent, its main drawback is its sensitivity to

[40] Here $S(A)$ denotes any stopping rule (possibly stochastic) with the property that with probability 1, $S(A) \to \infty$ as $A \to \infty$.

the non-parametric estimates of $P(d|x)$ in finite samples. One can see from the log-odds transformation (3.53) that if one of the estimated choice probabilities entering the odds ratio is 0, the normalized value function will equal plus or minus ∞. For states x with few observations, the natural non-parametric estimator for $P(d|x)$ – sample frequencies – can frequently be 0 if the true value of $P(d|x)$ is non-zero but small. In general, the inversion process can amplify estimation errors in the non-parametric estimates of $P(d|x)$, and these errors can result in biased estimates of θ. A related problem is that the SV estimator requires estimates of the normalized value function $\Delta v(x,d)$ for all possible (x,d) pairs, but many data sets will frequently have data concentrated only in a small subset of (x,d) pairs. Such cases may require parametric specifications of $P(d|x)$ to be able to extrapolate estimates of $P(d|x)$ to the set of (x,d) pairs where there are no observations. In their Monte Carlo study of he SV estimator, Hotz et al. found that smoothed estimates of $P(d|x)$ (such as produced by kernel estimators, for example) can reduce bias, but in general the SV estimator depends on having lots of data to obtain accurate non-parametric estimates of P.[41]

3.5. *The identification problem*

I conclude with a brief discussion of the identification problem for MDP's and DDP's. Without strong restrictions on the primitives (u,p,β) the structural model is "non-parametrically unidentified": there are infinitely many distinct primitives that are consistent with any given decision rule δ. Indeed, we show that the hypothesis that agents behave according to Bellman's principle of optimality imposes no testable restrictions in the sense that we can always find primitives $\{\beta,u,p\}$ that "rationalize" an arbitrary decision rule δ. This result stands in contrast to the case of static choice models where we know that the hypothesis of optimization per se does imply testable restrictions.[42] The absence of restrictions in the dynamic case may seem surprising given that the structure of the MDP problem already imposes a number of strong restrictions such as time additive preferences and constant intertemporal discount factors, as well as the expected utility hypothesis itself. While Bellman's principle does not place restrictions on historical choice behavior, it can yield restrictions in choice experiments where we have at least partial control over agents' preferences or beliefs.[43] For example, by varying agents' beliefs from p to p', an experiment implies a new optimal decision rule $\delta(\beta^*,u^*,p')$, where β^* and u^* are the agent's "true" discount factor and utility function. This experiment imposes the

[41] See also Bansal et al. (1993) for another interesting non-parametric simulation estimator that might also be effective for estimating large scale DDP problems.

[42] These restrictions include the symmetry and negative-semidefiniteness of the Slutsky matrix [Hurwicz and Uzawa (1971)], the generalized axiom of revealed preference [Varian (1984)], and in the case of discrete choice, restrictions on conditional choice probability [Block and Marschak (1960)].

[43] An example is the experiment which revealed the "Allais paradox" mentioned in Footnote 9.

following restriction on candidate (β, u) combinations: they must lie in the set R defined by

$$R = \{(\beta, u) | \delta(\beta, u, p) = \delta(\beta^*, u^*, p)\} \cap \{(\beta, u) | \delta(\beta, u, p') = \delta(\beta^*, u^*, p')\}. \tag{3.57}$$

Although experimental methods offer a valuable source of additional testable restrictions that can help narrow down the equivalence class of competing structural explanations of observed behavior, it should be clear that even extensive experimentation will be unable to uniquely identify the "true" structure of agents' decision-making processes.

As is standard in analyses of the identification problem, we will assume that the econometrician has access to an unlimited number of observations. This is justified since our results in this section are primarily negative: if the primitives $\{\beta, u, p\}$ cannot be identified with infinite amounts of data, then they certainly can't be identified in finite samples. We consider general MDP's without unobservables, since the existence of unobservables can only make the identification problem worse. In order to formulate the identification problem à la Cowles Commission, we need to translate the concepts of reduced-form and structure to the context of a nonlinear MDP model.

Definition 3.2

The *reduced-form* of an MDP is the agent's optimal decision rule, δ.

Definition 3.3

The *structure* of an MDP is the mapping: $\Lambda: \{\beta, u, p\} \to \delta$ defined by

$$\delta(s) = \underset{d \in D(s)}{\operatorname{argmax}} \, [v(s, d)], \tag{3.58}$$

where v is the unique fixed point to

$$v(s, d) = u(s, d) + \beta \int \max_{d' \in D(s')} [v(s', d')] p(\mathrm{d}s' | s, d). \tag{3.59}$$

The rationale for identifying δ as the reduced-form of the MDP is that it embodies all observable implications of the theory and can be consistently estimated by non-parametric regression given sufficient number of observations $\{s_t, d_t\}$.[44] We can

[44] Since the econometrician fully observes all components of (s, d), the non-parametric regression is degenerate in the sense that the model $d = \delta(s)$ has an "error term" that is identically 0. Nevertheless, a variety of non-parametric regression methods will be able to consistently estimate δ under weak regularity conditions.

use the reduced-form estimate of δ to define an equivalence relation over the space of primitives.

Definition 3.4

Primitives (u, p, β) and $(\bar{u}, \bar{p}, \bar{\beta})$ are *observationally equivalent* if

$$\Lambda(\beta, u, p) = \Lambda(\bar{\beta}, \bar{u}, \bar{p}). \tag{3.60}$$

Thus $\Lambda^{-1}(\delta)$ is the equivalence class of primitives consistent with decision rule δ. Expected-utility theory implies that $\Lambda(\beta, u, p) = \Lambda(\beta, au + b, p)$ for any constants a and b satisfying $a > 0$, so at best we will only be able to identify an agent's preferences u modulo cardinal equivalence, i.e. up to a positive linear transformation of u.

Definition 3.5

The stationary MDP problem (3.58) and (3.59) is *non-parametrically identified* if given any reduced-form δ in the range of Λ, and any primitives (u, p, β) and $(\bar{u}, \bar{p}, \bar{\beta})$ in $\Lambda^{-1}(\delta)$ we have

$$\begin{aligned} \beta &= \bar{\beta}, \\ p &= \bar{p}, \\ u &= a\bar{u} + b, \end{aligned} \tag{3.61}$$

for some constant a and b satisfying $a > 0$.

Lemma 3.2

The MDP problem (3.58) and (3.59) is non-parametrically unidentified.

The proof of this result is quite simple. Given any δ in the range of Λ, let $(\beta, u, p) \in \Lambda^{-1}(\delta)$. Define a new set of primitives $(\bar{u}, \bar{p}, \bar{\beta})$ by

$$\begin{aligned} \bar{\beta} &= \beta, \\ \bar{p}(\mathrm{d}s' \mid s, d) &= p(\mathrm{d}s' \mid s, d), \\ \bar{u}(s, d) &= u(s, d) + f(s) - \beta \int f(s') p(\mathrm{d}s' \mid s, d), \end{aligned} \tag{3.62}$$

where f is an arbitrary measurable function of s. Then $\bar{u}$ is clearly not cardinally equivalent to u unless f is a constant. To see that both (u, p, β) and $(\bar{u}, \bar{p}, \bar{\beta})$ are observationally equivalent, note that if $v(s, d)$ is the value function corresponding to primitives (u, p, β) then we conjecture that $\bar{v}(s, d) = v(s, d) + f(s)$ is the value function corresponding to $(\bar{u}, \bar{p}, \bar{\beta})$:

$$
\begin{aligned}
\bar{v}(s,d) &= \bar{u}(s,d) + \beta \int \max_{d' \in D(s')} [\bar{v}(s',d')] p(\mathrm{d}s' \mid s,d) \\
&= u(s,d) + f(s) - \beta \int f(s') p(\mathrm{d}s' \mid s,d) + \beta \int \max_{d' \in D(s')} [v(s',d') + f(s')] p(\mathrm{d}s' \mid s,d) \\
&= u(s,d) + f(s) + \beta \int \max_{d' \in D(s')} [v(s',d')] p(\mathrm{d}s' \mid s,d) \\
&\equiv v(s,d) + f(s).
\end{aligned} \tag{3.63}
$$

Since v is the unique fixed point to (3.59), it follows that $v + f$ is the unique fixed point to (3.63), so our conjecture $\bar{v} = v + f$ is indeed the unique fixed point to Bellman's equation with primitives $(\bar{u}, \bar{p}, \bar{\beta})$. Clearly $\{v(s,d), d \in D(s)\}$ and $\{v(s,d) + f(s), d \in D(s)\}$ yield the same decision rule δ. It follows that $(\beta, u + f - \beta Ef, p)$ is observationally equivalent to (β, u, p), but $u + f - \beta Ef$ is not cardinally equivalent to u.

We now ask whether Bellman's principle places any restrictions on the decision rule δ. In the case of MDP's Blackwell's theorem (Theorem 2.3) does provide two restrictions: (1) δ is Markovian, (2) δ is deterministic. In practice, it is extremely difficult to test these restrictions empirically. Presumably we could test the first restriction by seeing whether agents' decisions depend on lagged states s_{t-k} for $k = 1, 2, \ldots$. However given that we have not placed any a priori bounds on the dimensionality of S, the well-known trick of "expanding the state space" [Bertsekas (1987, Chapter 4)] can be used to transform an Nth order Markov process into a 1st order Markov process. The second restriction might be tested by looking for agents who make different choices in the same state: $\delta(s_t) \neq \delta(s_{t+k})$, for some state $s_t = s_{t+k} = s$. However, this behavior can be rationalized by a model where the agent is indifferent between several alternatives available in state s and simply chooses one at random. The following lemma shows that Bellman's principle implies no other restrictions beyond the two essentially untestable restrictions of Blackwell's theorem:

Lemma 3.3

Given an arbitrary measurable mapping $\delta: S \to D$ there exist primitives (u, p, β) such that $\delta = \Lambda(\beta, u, p)$.

The proof of this result is straightforward. Given an arbitrary discount factor $\beta \in (0,1)$ and transition probability p, define u by

$$
u(s,d) = I\{d = \delta(s)\} - \beta. \tag{3.64}
$$

Then it is easy to see that $v(s,d) = I\{d = \delta(s)\}$ is the unique solution to Bellman's equation (3.59), so that δ is the optimal decision rule implied by (u, p, β).

If we are unwilling to place any restrictions on (u, p, β), then Lemma 3.2 shows that the resulting MDP model is non-parametrically unidentified and Lemma 3.3 shows that it has no testable implications, in the sense that we can "rationalize" any decision rule δ. However, it is clear that from an economic standpoint, many of the utility functions $u + f - \beta E f$ will generally be implausible, as is the utility function $u(s,d) = I\{d = \delta(s)\}$. These results provide a simple illustration of the need for auxiliary identifying restrictions to eliminate primitives that we deem "unreasonable" while at the same time pointing out the futility of direct non-parametric estimation of (u, p, β). The proofs of the lemmas also indicate that in order to obtain testable implications, we will need to impose very strong identifying restrictions on u. To see this, suppose that the agents' discount factor β is known a priori, and suppose that we invoke the hypothesis of rational expectations: i.e. agents' subjective beliefs about the evolution of the state variables coincide with the objectively measurable population probability measure. This implies that we can use observed realizations of the controlled process $\{s_t, d_t\}$ to consistently estimate the $p(\mathrm{d}s_{t+1} | s_t, d_t)$, which means that we can treat both β and p as known a priori. Note, however, that the proofs of Lemmas 3.2 and 3.3 are unchanged whether or not we know β and p, so it is clear that we must impose identifying restrictions directly on u itself. The usual way to impose restrictions is to assume that u is a smooth function of a vector of unknown parameters θ. However, in the case where S and D are finite sets, this is insufficient to uniquely identify the model: there is an equivalence class of θ with non-empty interior consistent with any decision rule δ_θ.

This latter identification problem is an artifact of the degeneracy of MDP models without unobservables. Rust (1996) provides a parallel identification analysis for the subclass of DDP models satisfying the AS–CI assumptions presented in Section 4.3, and shows that while AS–CI does impose testable restrictions, the primitives (u, q, π, β) are still non-parametrically unidentified in the absence of further identifying restrictions. However, when identifying restrictions take the form of smooth parametric specifications $(u_\theta, q_\theta, \pi_\theta, \beta_\theta)$, the presence of unobservables succeeds in smoothing out the problem, and the results of Section 4.3 imply that the likelihood is a smooth function of θ. Results from differentiable topology can then be used to show that the resulting parametric model is generically identified.

While the results in this section suggest that, in principle, one can always "rig" a DDP model to rationalize any given set of observations, we should emphasize that there is no guarantee that we can do it in a way that is theoretically plausible. Given the increasing size and complexity of current data sets, it can be a considerable challenge to find a plausible DDP model that is consistent with the available data, let alone one that is capable of making accurate out-of-sample forecasts of policy changes. However, in the cases where we do succeed in specifying a plausible structural model that fits the data well, we need to exercise a certain amount of caution in using the model for policy forecasting and welfare analyses, etc. Keep in mind that a model will be credible only until such time as it is rejected by new, out-of-sample observations, or faces competition from alternative models that are also consistent with available data and are equally plausible. The identification analysis suggests that we can't rule out the existence of alternative plausible, observationally equivalent models that generate completely different policy forecasts or welfare impacts. Put simply, since structural models can be falsified but never proven to be true, their predictions should always be treated as tentative and subject to continual verification.

4. Empirical applications

I conclude this chapter with brief reviews of two empirical applications of DDP models: Rust's (1987) model of optimal replacement of bus engines and the Lumsdaine et al. (1992) model of optimal retirement from a firm.

4.1. *Optimal replacement of bus engines*

One of the simplest applications of the specific class of DDP models given in Definition 3.1 is Rust's (1987) application to bus engine replacement. In contrast to the macroeconomic studies that investigate aggregate replacement investment, this model goes to the other extreme and examines the replacement investment decision at the level of an individual agent. In this case the agent, Harold Zurcher, is the maintenance manager of the Madison Metropolitan Bus Company. One of the problems he faces is to decide how long to operate a bus before replacing its engine with a new or completely overhauled bus engine. We can represent Zurcher's problem as a DDP with state variable x_t equal to the cumulative mileage on a bus since last engine replacement, and control variable d_t which equals 1 if Zurcher decides to replace the bus engine, and 0 otherwise. Rust assumed that Zurcher behaves as a cost minimizer, so his utility function is given by

$$u(x_t, d_t, \theta_1, \theta_2) = \begin{cases} -\theta_1 - c(0, \theta_2) - \varepsilon_t(1) & \text{if} \quad d_t = 1 \\ -c(x_t, \theta_2) - \varepsilon_t(0) & \text{if} \quad d_t = 0, \end{cases} \tag{4.1}$$

where θ_1 represents the labor and parts cost of installing a replacement bus engine and $c(x, \theta_1)$ represents the expected monthly operating and maintenance costs of a bus with x accumulated miles since last replacement. Implicit in the specification (4.1) is the assumption that when a bus engine is replaced, it is "as good as new", so the state of the system regenerates to $x_t = 0$ when $d_t = 1$. This regeneration property is also captured in the transition probability for x_t:

$$\pi(x_{t+1} | x_t, d_t) = \begin{cases} g(x_{t+1} - 0) & \text{if} \quad d_t = 1 \\ g(x_{t+1} - x_t) & \text{if} \quad d_t = 0, \end{cases} \tag{4.2}$$

where g is a probability density function. The renewal property given in equations (4.1) and (4.2) defines a *regenerative optimal stopping problem*, and under an optimal decision rule $d_t = \delta^*(x_t, \varepsilon_t, \theta)$ the mileage process $\{x_t\}$ is a *regenerative random walk*. Using data on 8156 bus/month observations over the period 1975 to 1986, Rust estimated θ and g using the maximum likelihood estimator (3.25). Figures 2 and 3 present the estimated value and replacement hazard functions assuming a linear cost function, $c(x, \theta_2) = \theta_2 x$ and two different values of β. Comparing the estimated hazard function $P(1|x, \hat{\theta})$ to the non-parametrically estimated hazard, both the dynamic ($\beta = 0.9999$) and myopic ($\beta = 0$) models appear to fit the data equally well. In particular, both models predict that the probability of a replacement is essentially 0 for x less than 100 000 miles. However likelihood ratio and chi-square tests both strongly reject the myopic model in favor of the dynamic model: the data imply that the concave value function for $\beta = 0.9999$ fits the data better than the linear value function $\theta_2 x$ when $\beta = 0$. The precise value of β could not be identified: the likelihood was virtually flat for $\beta \geqslant 0.98$, although with a very slight upward slope as $\beta \to 1$. The latter finding, together with the *final value theorem* [Bhattacharya and Majumdar (1989)] may indicate that Zurcher is minimizing long-run average costs rather than discounted costs:

$$\lim_{\beta \to 1} \max_{\delta} E_\delta \left\{ (1-\beta) \sum_{t=0}^{\infty} \beta^t u(x_t, d_t) \right\} = \lim_{T \to \infty} \max_{\delta} \frac{1}{T} E_\delta \left\{ \sum_{t=1}^{T} u(x_t, d_t) \right\}. \tag{4.3}$$

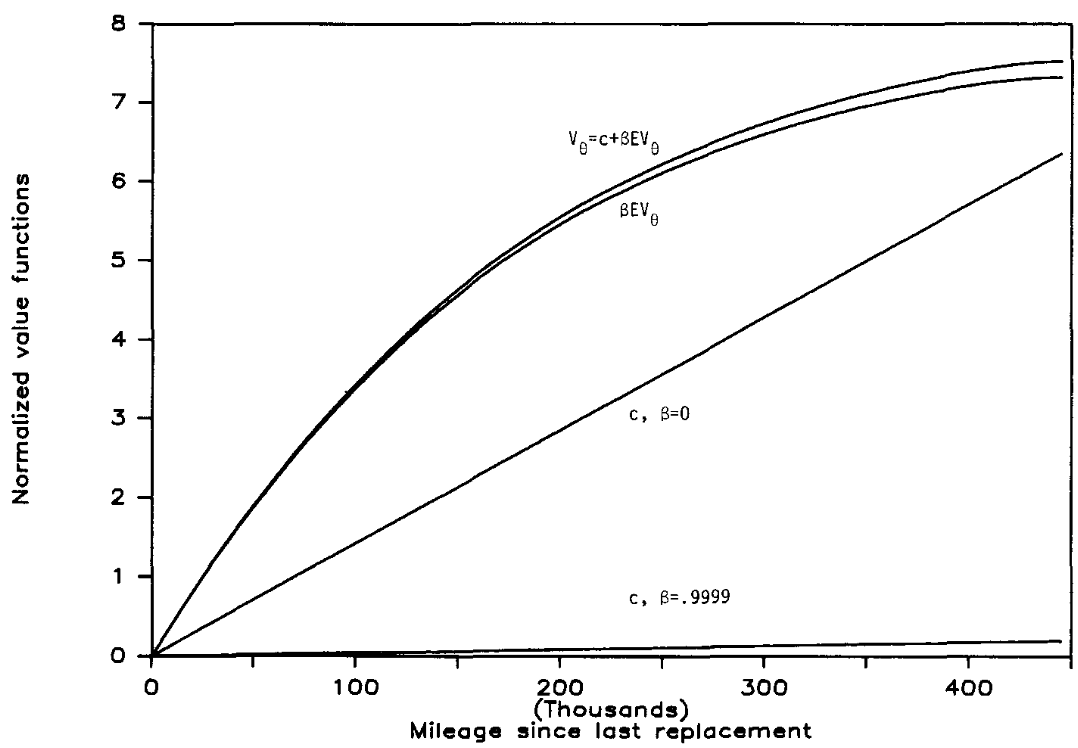

Figure 2. Estimated value functions

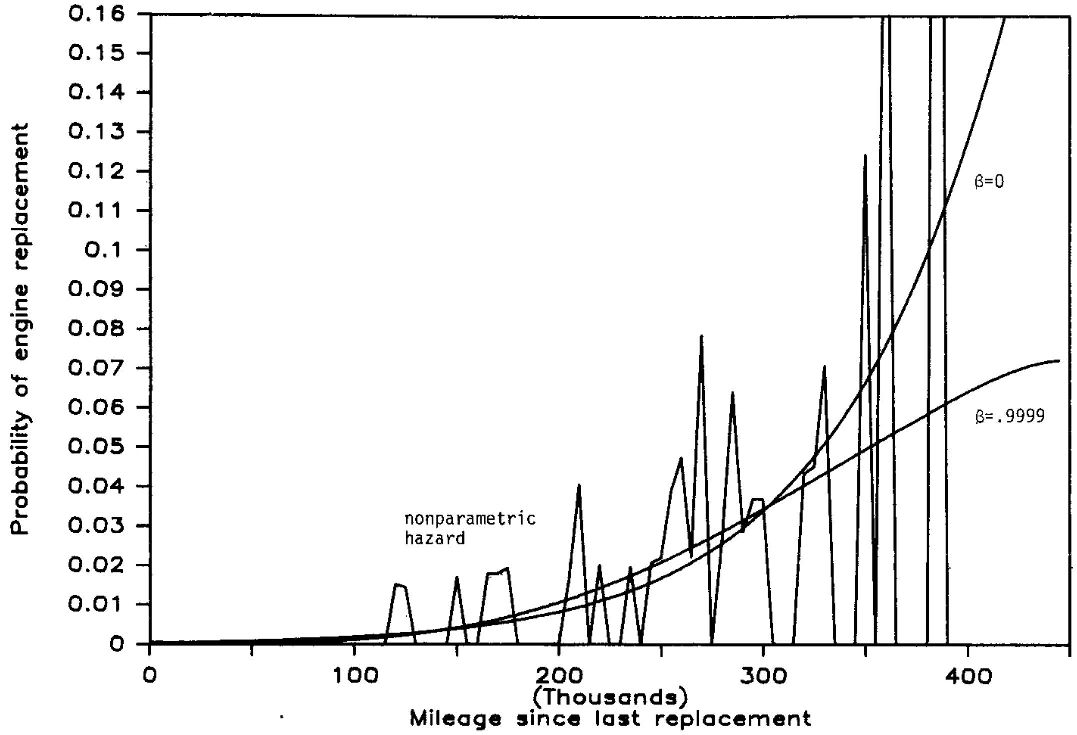

Figure 3. Estimated hazard functions

The estimated model implies that expected monthly maintenance costs increase by \$1.87 for every additional 5000 miles. Thus, a bus with 300 000 miles costs an average of \$112 per month more to maintain than a bus with a newly replaced engine. Rust found evidence of heterogeneity across bus types, since the estimated value of θ_2 for the newer 1979 model GMC buses is nearly twice as high as for the 1974 models. This finding resolves a puzzling aspect of the raw data: engine replacements for the 1979 model buses occur on average 57 000 *earlier* than the 1974 model buses despite the fact that the engine replacement cost for the 1979 models is 25% higher. One of the nice features of estimating the preferences at the level of a single individual is that one can evaluate the accuracy of the structural model by simply asking the individual whether the estimated utility function is reasonable.[45] In this case conversations with Zurcher revealed that implied cost estimates of the structural model corresponded closely with his perceptions of operating costs, including the finding that monthly maintenance expenditures for the 1979 model GMC buses were nearly twice as high as for the 1974 models.

[45] Another nice feature is that we completely avoid the problem of *unobserved heterogeneity* that can confound attempts to estimate dynamic models using panel data. Heckman (1981a, b) provides a good discussion of this problem in the context of models for discrete panel data.

Figure 4 illustrates the value of structural models for use in policy forecasting. In this case, we are interested in forecasting Zurcher's demand for replacement bus engines as their price varies. Figure 4 contains two "demand curves" computed for models estimated with $\beta = 0.9999$ and $\beta = 0$, respectively. The maximum likelihood procedure insures that both models generate the same predictions at the actual replacement price of $4343. A reduced-form approach to forecasting bus engine demand would bypass the difficulties of structural estimation and simply regress the number of engine replacements in a given period as a function of replacement costs during the period. However, since the cost of replacement bus engines has not varied much in the past, the reduced-form approach will be incapable of generating reliable predictions of replacement demand. In terms of Figure 4, all the data would be clustered in a small ball about the intersection of the two curves: obviously many different demand functions will be able to fit the data equally well. By parameterizing our prior knowledge about the replacement problem Zurcher is facing, and by efficiently using the additional information contained in the $\{x_t^a, d_t^a\}$ sequences, the structural model is able to generate very precise estimates of the replacement demand function.

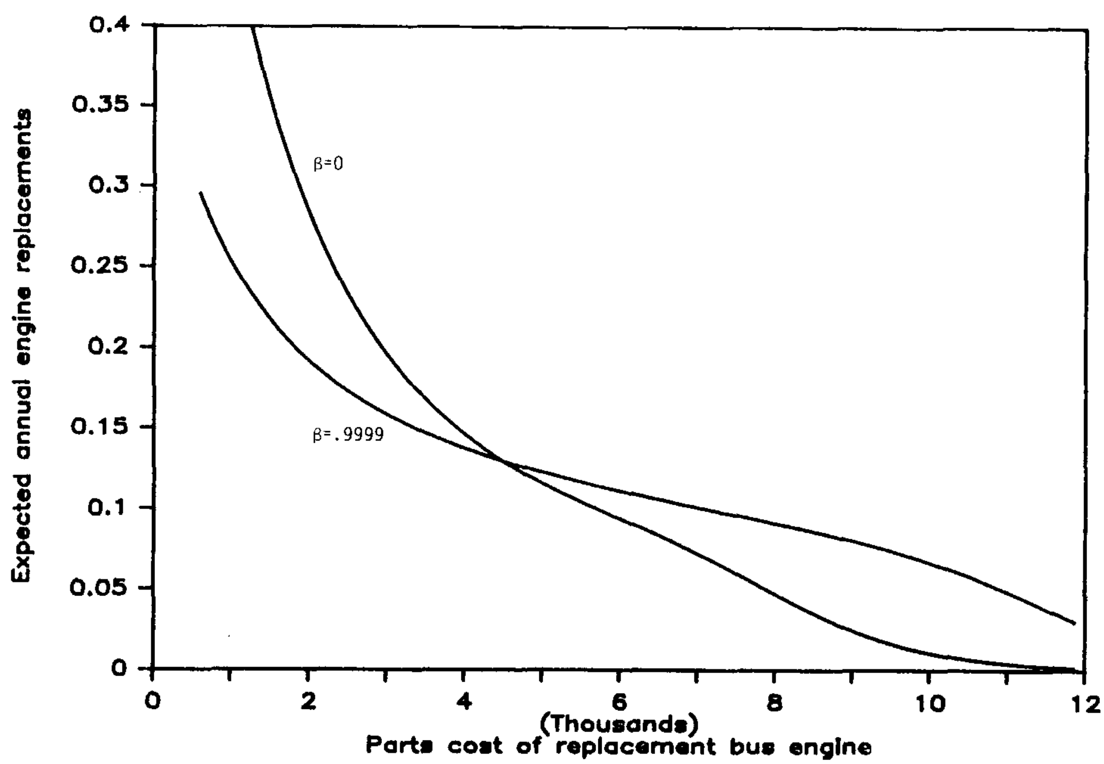

Figure 4. Expected replacement demand function

4.2. *Optimal retirement from a firm*

Lumsdaine, Stock and Wise (1992) (LSW) used a data set that provides a unique "natural experiment" that allows them to compare the policy forecasts of four competing structural and reduced-form models. The data consist of observations of departure rates of older office employees at a Fortune 500 company. These workers were covered by a "defined benefit" pension plan that provided substantial incentives to remain with the firm until age 55 and then substantial incentives to leave the firm before age 65. In 1982 non-managerial employees who were over 55 and vested in the pension plan were offered a temporary 1 year "window plan". Under the window plan employees who retired in 1982 were offered a bonus equivalent to 3 to 12 months' salary, depending on the years of service with the firm. Needless to say, the firm experienced a substantial increase in departure rates in 1982.

Using data prior to 1982, LSW fit four alternative econometric models and used the fitted models to make out-of-sample forecasts of departure rates under the 1982 window plan. One of the models was a reduced-form probit model with various explanatory variables and the remaining three were different types of dynamic structural models.[46] Two of the structural models were finite-horizon DDP models with a binary decision variable: continue working ($d = 1$) or quit ($d = 0$). Since quitting is treated as an absorbing state and workers were subject to mandatory retirement at age 70, the DDP model reduces to a simple finite-horizon optimal stopping problem. The observed state variable x_t is the benefit (wage or pension) obtained in year t and the unobserved state variable ε_t is assumed to be an IID extreme-value process in the first specification and an IID Gaussian process in the other specification. LSW used the following specification for workers' utility functions:

$$u(x_t, \varepsilon_t, d_t, \theta) = \begin{cases} x_t^{\theta_1} + \mu_1 + \varepsilon_t(1) & \text{if} \quad d_t = 1 \\ (\mu_2 \theta_2 x_t)^{\theta_1} + \varepsilon_t(0) & \text{if} \quad d_t = 0. \end{cases} \tag{4.4}$$

The two components of $\mu = (\mu_1, \mu_2)$ represent time-invariant worker-specific heterogeneity. In the extreme-value specification for $\{\varepsilon_t\}$, μ_1 is assumed to be identically 0 and μ_2 is assumed to have a log-normal population distribution with mean 1 and scale parameter θ_3. In the Gaussian specification for $\{\varepsilon_t\}$, μ_2 is identically 1 and μ_1 is assumed to have Gaussian population distribution with mean 0 and standard deviation θ_4. Although the model does not directly include "leisure" in the utility

[46] LSW used different sets of explanatory variables in the reduced-form model including calculated "option values" of continued work (the expected present value of benefits earned by retiring at the "optimal" age, i.e. the age at which the total present value of benefits, wages plus pensions, is maximized). Other specifications used the levels and present values of Social Security and pension benefits as well as changes in the present value of these benefits ("pension accruals"), predicted earnings in the next year of employment, and age.

function, it is implicitly included via the parameter θ_2. Thus, we expect that $\theta_2 > 1$ since the additional leisure time should imply that a dollar of pension income is worth more than a dollar of wage income.

The third structural model is the "option value model" developed in previous work by Stock and Wise (1990). The option value model predicts that the worker will leave the firm at the first year t in which the expected presented discounted value of benefits from departing at t exceeds the maximum of the expected values of departing at any future date. This rule differs from an optimal stopping rule generated from the solution to a DP problem by interchanging the "expectation" and "maximization" operators. This results in a temporally inconsistent decision rule in which the worker ignores the fact that as new information arrives he will be continually revising his estimate of the optimal departure date t^*.[47]

The parameter estimates for the three structural models are presented in Table 1. There are significant differences in the parameter estimates in the three models. The

Table 1
Parameter estimates for the option value and the dynamic programing models.

Parameter	Option Value Models		Dynamic Programming Models					
	(1)	(2)	Extreme Value			Normal		
			(1)	(2)	(3)	(4)	(5)	(6)
θ_1	1.00*	0.612	1.00*	1.018	1.187	1.00*	1.187	1.109
		(0.072)		(0.045)	(0.215)		(0.110)	(0.275)
θ_2	1.902	1.477	1.864	1.881	1.411	2.592	2.975	2.974
	(0.192)	(0.445)	(0.144)	(0.185)	(0.307)	(0.100)	(0.039)	(0.374)
β	0.855	0.895	0.618	0.620	0.583	0.899	0.916	0.920
	(0.046)	(0.083)	(0.048)	(0.063)	(0.105)	(0.017)	(0.013)	(0.023)
θ_3	0.168	0.109	0.306	0.302	0.392	0.224	0.202	0.168
	(0.016)	(0.046)	(0.037)	(0.036)	(0.090)	(0.021)	(0.022)	(0.023)
θ_4			0.00*	0.00*	0.407	0.00*	0.00*	0.183
					(0.138)			(0.243)
Summary Statistics								
$-\ln \mathscr{L}$	294.59	280.32	279.60	279.57	277.25	277.24	276.49	276.17
χ^2 sample	36.5	53.5	38.9	38.2	36.2	45.0	40.7	41.5
χ^2 window	43.9	37.5	32.4	33.5	33.4	29.0	25.0	24.3

Notes: Estimation is by maximum likelihood. All monetary values are in $100,000 (1980 dollars).
*Parameter value imposed.

[47] Thus, it is somewhat of a misnomer to call it an "option value" model since it ignores the option value of new information that is explicitly accounted for in a dynamic programming formulation. Stern (1995) shows that in many problems the computationally simpler procedure of interchanging the expectation and maximization operators yields a very poor approximation to the optimal decision rule computed by DP methods.

Gaussian DP specification predicts a much higher implicit value for leisure (θ_2) than the other two models, and the extreme-value specification yields a much lower estimate of the discount factor β. The estimated standard deviation σ of the ε_t's is also much higher in the extreme-value specification than the Gaussian. Allowing for unobserved heterogeneity did not significantly improve the fit of the Gaussian DP model, although it did have a marginal impact on the extreme-value model.[48]

Figures 5 to 7 summarize the ability of the four models to fit the historical data.

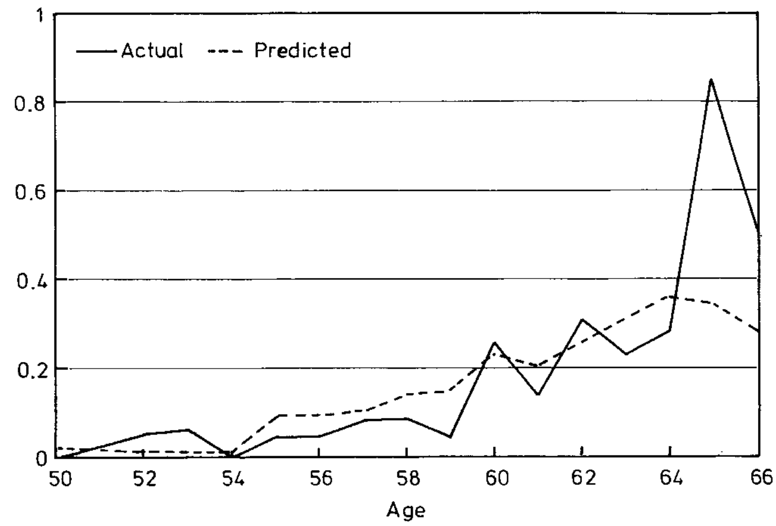

Figure 5. Predicted versus actual 1980 departure rates and implicit cumulative departures, option value model

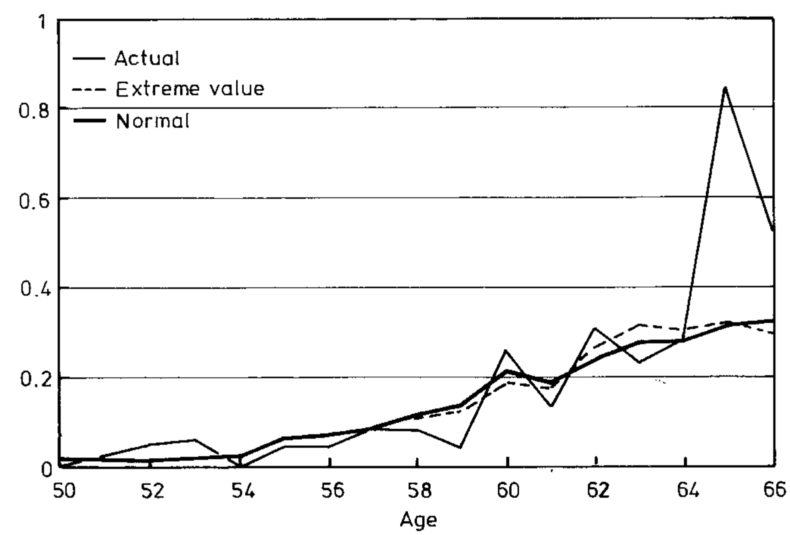

Figure 6. Predicted versus actual 1980 departure rates and implicit cumulative departures, dynamic programming model

[48]The log-likelihoods in the Gaussian and extreme-value specifications improved to -276.17 and -277.25, respectively.

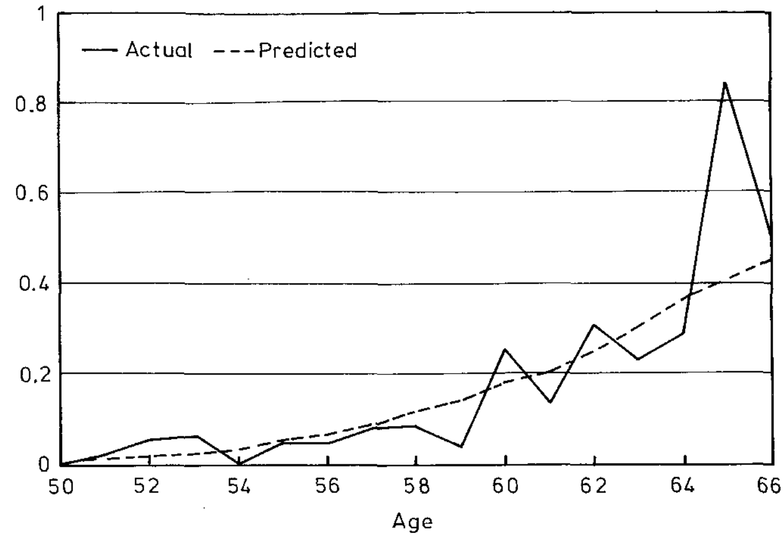

Figure 7. Predicted versus actual departure rates and implicit cumulative departures probit model

Figure 5 compares the actual departure rates (solid line) to the departure rates predicted by the option value model (dashed line). Figure 6 presents a similar comparison for the two DDP models, and Figure 7 presents the results for the best-fitting probit model. As one can see from the graphs all four models provide a relatively good fit of actual departure rates except that they all miss the pronounced peak in departure rates at age 65. Table 1 presents the χ^2 goodness-of-fit statistics for each of the models. The four models are generally comparable, although the option model fits slightly worse and the probit model fits slightly better than the two DDP models. The superior fit of the probit model is probably due to the inclusion of age trends that are excluded from the other models.

Figures 8 to 10 summarize the ability of the models to track the shift in departure rates induced by the 1982 window plan. All forecasts were based on the estimated utility function parameters using data prior to 1982. Using these parameters, predictions were generated from all four models after incorporating the extra bonus provisions of the window plan. As is evident from Figures 8–10, the structural models were generally able to accurately predict the large increase in departure rates induced by the window plan, although once again none of the models was able to capture the peak in departure rates at age 65. On the other hand, the reduced-form probit model predicted that the window plan had essentially no effect on departure rates. Other reduced-form specifications greatly overpredicted departure rates under the window plan. The χ^2 goodness-of-fit statistics presented in Table 1 show that all of the structural models do a significantly better job of predicting the impact of the window plan than any of the reduced-form models.[49] LSW concluded that:

[49]The smallest χ^2 value for any of the reduced-form models under the window plan was 57.3, the largest was 623.3.

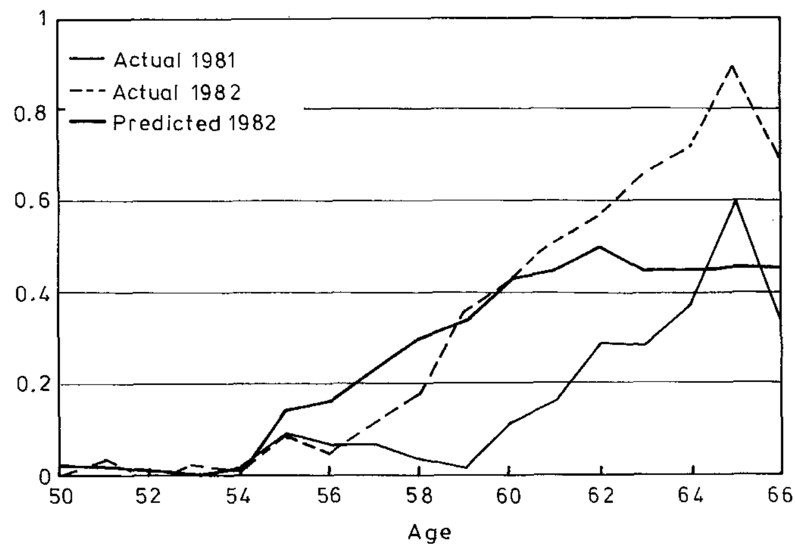

Figure 8. Predicted versus actual departure rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: option value model

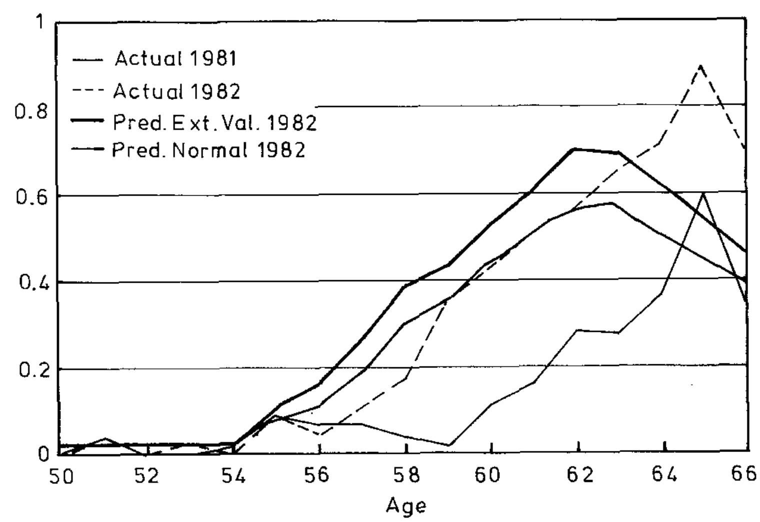

Figure 9. Predicted versus actual departure rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: dynamic programming models

The option value and the dynamic programming models fit the data equally well, with a slight advantage to the normal dynamic programming model. Both models correctly predicted the very large increase in retirement under the window plan, with some advantage in the fit to the option value model. In short, this evidence suggests that the option value and dynamic programming models are considerably more successful than the less complex probit model in approximating the rules

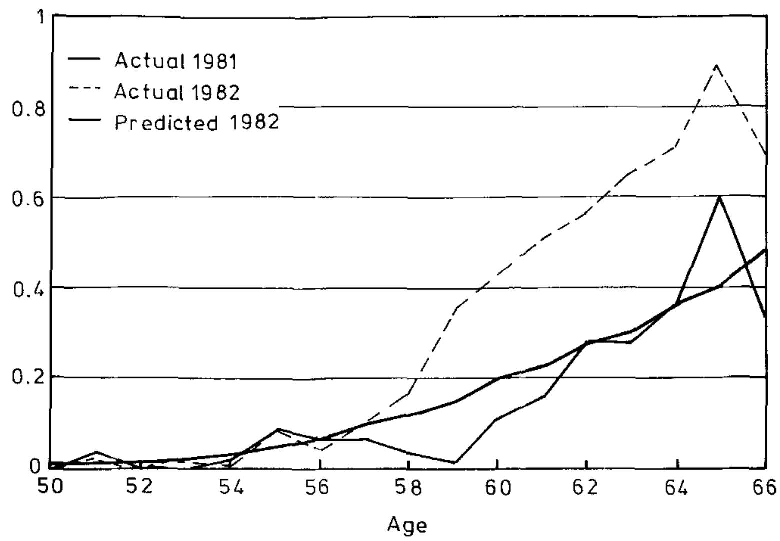

Figure 10. Predicted versus actual departures rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: probit model

individuals use to make retirement decisions, but that the more complex dynamic programming rule approximates behavior no better than the simpler option value rule. More definitive conclusions will have to await accumulated evidence based on additional comparisons using different data sets and with respect to different pension plan provisions. (p. 31).

References

Ahn, M.Y. (1993a) Duration Dependence, Search Effort, and Labor Market Outcomes: A Structural Model of Labor Market History, manuscript, Duke University.

Ahn, M.Y. (1993b) "Econometric Analysis of Sequential Discrete Choice Behavior with Duration Dependence", manuscript, Duke University.

Ahn, H. and Manski, C. (1993) "Distribution Theory for the Analysis of Binary Choice Under Uncertainty with Nonparametric Estimation of Expectations", *Journal of Econometrics*, 56(3), 270–291.

Amemiya, T. (1976) "On A Two-Step Estimation of a Multivariate Logit Model", *Journal of Econometrics*, 8, 13–21.

Andrews, D.W.K. (1988a) "Chi-Square Diagnostic Tests for Econometric Models", *Journal of Econometrics*, 37, 135–156.

Andrews, D.W.K. (1988b) "Chi-Square Tests for Econometric Models: Theory", *Econometrica*, 56, 1414–1453.

Arrow, K.J., T. Harris and J. Marschak (1951) "Optimal Inventory Policy", *Econometrica*, 19(3), 250–272.

Bansal, R., A.R. Gallant, R. Hussey and G. Tauchen (1993) "Nonparametric Estimation of Structural Models for High Frequency Currency Market Data", *Journal of Econometrics*, forthcoming.

Basawa, I.V. and B.L.S. Prakasa Rao (1980) *Statistical Inference for Stochastic Processes*. New York: Academic Press.

Bellman, R. (1957) *Dynamic Programming*. Princeton University Press: Princeton.

Berkovec, J. and S. Stern (1991) "Job Exit Behavior of Older Men", *Econometrica*, 59(1), 189–210.

Berndt, E., B. Hall, R. Hall and J. Hausman (1974) "Estimation and Inference in Nonlinear Structural Models", *Annals of Economic and Social Measurement*, 3, 653–665.
Bertsekas, D. (1987) *Dynamic Programming Deterministic and Stochastic Models*, Prentice Hall: New York.
Bertsekas, D. and D. Castañon (1989) "Adaptive Aggregation Methods for Infinite Horizon Dynamic Programming", *IEEE Transactions on Automatic Control*, 34(6), 589–598.
Bhattacharya, R.N. and M. Majumdar (1989) "Controlled Semi-Markov Models – The Discounted Case", *Journal of Statistical Planning and Inference*, 21, 365–381.
Billingsley, P. (1961) *Statistical Inference for Markov Processes*. University of Chicago Press: Chicago.
Blackwell, D. (1962) "Discrete Dynamic Programming", *Annals of Mathematical Statistics*, 33, 719–726.
Blackwell, D. (1965) "Discounted Dynamic Programming", *Annals of Mathematical Statistics*, 36, 226–235.
Blackwell, D. (1967) "Positive Dynamic Programming", *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1, 415–418.
Block, H. and J. Marschak (1960) "Random Orderings and Stochastic Theories of Response", in: I. Olkin, ed., *Contributions to Probability and Statistics*, Stanford University Press: Stanford.
Boldrin, M. and L. Montrucchio (1986) "On the Indeterminacy of Capital Accumulation Paths", *Journal of Economic Theory*, 40(1), 26–39.
Brock, W.A. (1982) "Asset Prices in a Production Economy", in: J.J. McCall, ed., *The Economics of Information and Uncertainty*, Chicago: University of Chicago Press.
Brock, W.A. and L.J. Mirman (1972) "Optimal Economic Growth Under Uncertainty: The Discounted Case", *Journal of Economic Theory*, 4, 479–513.
Chamberlain, G. (1984) "Panel Data", in: Z. Griliches and M.D. Intrilligator, eds., *Handbook of Econometrics Volume 2*. North-Holland: Amsterdam. 1247–1318.
Chew, S.H. and L.G. Epstein (1989) "The Structure of Preferences and Attitudes Towards the Timing and Resolution of Uncertainty", *International Economic Review*, 30(1), 103–118.
Christensen, B.J. and N.M. Kiefer (1991a) "The Exact Likelihood Function for an Empirical Job Search Model", *Econometric Theory*.
Christensen, B.J. and N.M. Kiefer (1991b) Measurement Error in the Prototypical Job Search Model, manuscript, Cornell University.
Cox, D.R. (1975) "Partial Likelihood", *Biometrika*, 62(2), 269–276.
Cox, D.R. and D.V. Hinkley (1974) *Theoretical Statistics*. Chapman and Hall: London.
Dagsvik, J. (1983) "Discrete Dynamic Choice: An Extension of the Choice Models of Luce and Thurstone", *Journal of Mathematical Psychology*, 27, 1–43.
Dagsvik, J. (1991) A Note on the Theoretical Content of the GEU Model for Discrete Choice, manuscript, Central Bureau of Statistics, Oslo.
Daly, A.J. and S. Zachary, (1978) "Improved Multiple Choice Models", in: D.A. Hensher and Q. Dalvi, eds., *Determinants of Travel Choice*, 335–357, Teakfield, Hampshire.
Das, M. (1992) "A Micro Econometric Model of Capital Utilization and Retirement: The Case of the Cement Industry", *Review of Economic Studies*, 59(2), 287–298.
Daula, T. and R. Moffitt (1991) "Estimating a Dynamic Programming Model of Army Reenlistment Behavior", *Military Compensation and Personnel Retention.*
Debreu, G. (1960) "Review of R.D. Luce *Individual Choice Behavior*", *American Economic Review*, 50, 186–188.
Denardo, E.V. (1967) "Contraction Mappings Underlying the Theory of Dynamic Programming", *SIAM Review*, 9, 165–177.
Dongarra, J.J. and T. Hewitt (1986) "Implementing Dense Linear Algebra Algorithms Using Multitasking on the Cray X-MP-4", *SIAM Journal on Scientific and Statistical Computing*, 7(1), 347–350.
Eckstein, Z. and K. Wolpin (1989a) "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models", *Journal of Human Resources*, 24(4), 562–598.
Eckstein, Z. and K. Wolpin (1989b) "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience", *Review of Economic Studies*, 56, 375–390.
Eckstein, Z. and K. Wolpin (1990) "Estimating a Market Equilibrium Search Model from Panel Data on Individuals", *Econometrica*, 58(4), 783–808.
Engle, R. (1984) "Wald, Likelihood Ratio and Lagrange Multiplier Tests in Econometrics", in: Z. Griliches and M.D. Intrilligator, eds., *Handbook of Econometrics Vol 2*. North-Holland: Amsterdam.

Epstein, L.G. and S.E. Zin (1989) "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework", *Econometrica*, 57(4), 937–970.
Flinn, C. and J.J. Heckman (1982) "New Methods for Analyzing Laborforce Dynamics", *Journal of Econometrics*, 18, 115–168.
Gihman, I.I. and A.V. Skorohod (1979) *Controlled Stochastic Processes*, Springer–Verlag New York.
Gotz, G.A. and J.J. McCall (1980) "Estimation in Sequential Decisionmaking Models: A Methodological Note", *Economics Letters*, 6, 131–136.
Gotz, G.A. and J.J. McCall (1984) A Dynamic Retention Model for Air Force Officers, Report R–3028–AF, The RAND Corporation, Santa Monica, California.
Haavelmö, T. (1944) "The Probability Approach in Econometrics", *Econometrica Supplement*, 12, 1–115.
Hakansson, N. (1970) "Optimal Investment and Consumption Strategies Under Risk for a Class of Utility Functions", *Econometrica*, 38, 587–607.
Hansen, L.P. (1982) "Large Sample Properties of Method of Moments Estimators", *Econometrica*, 50, 1029–1054.
Hansen, L.P. (1994) in: R. Engle and D. McFadden, eds., *Handbook of Econometrics Vol. 4*. North-Holland: Amsterdam.
Hansen, L.P. and T.J. Sargent (1980a) "Formulating and Estimating Dynamic Linear Rational Expectations Models", *Journal of Economic Dynamics and Control*, 2(1), 7–46.
Hansen, L.P. and T.J. Sargent (1980b) "Linear Rational Expectations Models for Dynamically Interrelated Variables", in: R.E. Lucas, Jr. and T.J. Sargent, eds., *Rational Expectations and Econometric Practice*, Minneapolis: University of Minnesota Press.
Hansen, L.P. and T.J. Sargent (1993) Recursive Models of Dynamic Economies, manuscript, Hoover Institution.
Hansen, L.P. and K. Singleton (1982) "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models", *Econometrica*, 50, 1269–1281.
Hansen, L.P. and K. Singleton (1983) "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns", *Journal of Political Economy*, 91(2), 249–265.
Heckman, J.J. (1981a) "Statistical Models of Discrete Panel Data", in: C.F. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data*, MIT Press: Cambridge, Massachusetts.
Heckman, J.J. (1981b) "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating Discrete-Time, Discrete-Data Stochastic Processes", in: C.F. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data*, MIT Press: Cambridge, Massachusetts.
Heckman, J.J. (1991) Randomization and Social Policy Evaluation, NBER Technical Working paper 107.
Heckman, J.J. (1994) "Alternative Approaches to the Evaluation of Social Programs: Econometric and Experimental Methods", in: J. Laffont and C. Sims, eds., *Advances in Econometrics: Sixth World Congress*, Econometric Society Monographs, Cambridge University Press.
Hotz, V.J. and R.A. Miller (1993) "Conditional Choice Probabilities and the Estimation of Dynamic Models", *Review of Economic Studies*, forthcoming.
Hotz, V.J., R.A. Miller, S. Sanders and J. Smith (1993) "A Simulation Estimator for Dynamic Models of Discrete Choice", *Review of Economic Studies*, 60, 397–429.
Howard, R. (1960) *Dynamic Programming and Markov Processes*. J. Wiley: New York.
Howard, R. (1971) *Dynamic Probabilistic Systems: Volume 2 – Semi-Markov and Decision Processes*. J. Wiley: New York.
Hurwicz, L. and H. Uzawa (1971) "On the Integrability of Demand Functions", in: J. Chipman et al., eds., *Preferences, Utility, and Demand*. New York: Harcourt, Brace and Jovanovich.
Judd, K. (1994) Numerical Methods in Economics, manuscript, Hoover Institution.
Keane, M. and K. Wolpin (1994) The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation: Monte Carlo Evidence, manuscript, University of Minnesota.
Kennet, M. (1994) "A Structural Model of Aircraft Engine Maintenance", *Journal of Applied Econometrics*, forthcoming.
Kreps, D. and E. Porteus (1978) "Temporal Resolution of Uncertainty and Dynamic Choice Theory", *Econometrica*, 46, 185–200.
Kushner, H.J. (1990) "Numerical Methods for Stochastic Control Problems in Continuous Time", *SIAM Journal on Control and Optimization*, 28(5), 999–1048.
Kydland, F. and E.C. Prescott (1982) "Time to Build and Aggregate Fluctuations", *Econometrica*, 50, 1345–1370.

Lancaster, T. (1990) *The Econometric Analysis of Transition Data.* Cambridge University Press.
Leland, H. (1974) "Optimal Growth in a Stochastic Environment", *Review of Economic Studies*, 41, 75–86.
Levhari, D. and T. Srinivasan (1969) "Optimal Savings Under Uncertainty", *Review of Economic Studies*, 36, 153–163.
Long, J.B. and C. Plosser (1983) "Real Business Cycles", *Journal of Political Economy*, 91(1), 39–69.
Lucas, R.E. Jr. (1976) "Econometric Policy Evaluation: A Critique", in: K. Brunner and A.K. Meltzer, eds., *The Phillips Curve and Labour Markets.* Carnegie–Rochester Conference on Public Policy, North-Holland: Amsterdam.
Lucas, R.E. Jr. (1978) "Asset Prices in an Exchange Economy", *Econometrica*, 46, 1426–1446.
Lucas, R.E. Jr. and C.E. Prescott (1971) "Investment Under Uncertainty", *Econometrica*, 39(5), 659–681.
Lumsdaine, R., J. Stock and D. Wise (1992) "Three Models of Retirement: Computational Complexity vs. Predictive Validity", in: D. Wise, ed., *Topics in the Economics of Aging.* Chicago: University of Chicago Press.
Machina, M.J. (1982) "Expected Utility without the Independence Axiom", *Econometrica*, 50–2, 277–324.
Machina, M.J. (1987) "Choice Under Uncertainty: Problems Solved and Unsolved", *Journal of Economic Perspectives*, 1(1), 121–154.
Mantel, R. (1974) "On the Characterization of Excess Demand", *Journal of Economic Theory*, 7, 348–353.
Marcet, A. (1994) "Simulation Analysis of Stochastic Dynamic Models: Applications to Theory and Estimation", in: C. Sims and J. Laffont, eds., *Advances in Econometrics: Proceedings of the 1990 Meetings of the Econometric Society.*
Marschak, T. (1953) "Economic Measurements for Policy and Prediction", in: W.C. Hood and T.J. Koopmans, eds., *Studies in Econometric Method.* Wiley: New York.
McFadden, D. (1973) "Conditional Logit Analysis of Qualitative Choice Behavior", in: P. Zarembka, ed., *Frontiers of Econometrics.* Academic Press: New York.
McFadden, D. (1981) "Econometric Models of Probabilistic Choice", in: C.F. Manski and D. McFadden, eds. *Structural Analysis of Discrete Data.* MIT Press: Cambridge, Massachusetts.
McFadden, D. (1984) "Econometric Analysis of Qualitative Response Models", in: Z. Griliches and M.D. Intriligator, eds., *Handbook of Econometrics Vol. 2.* North-Holland: Amsterdam. 1395–1457.
McFadden, D. (1989) "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration", *Econometrica*, 57(5), 995–1026.
Merton, R.C. (1969) "Lifetime Portfolio Selection Under Uncertainty: The Continuous-time Case", *Review of Economics and Statistics*, 51, 247–257.
Miller, R. (1984) "Job Matching and Occupational Choice", *Journal of Political Economy*, 92(6), 1086–1120.
Pakes, A. (1986) "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks", *Econometrica*, 54, 755–785.
Pakes, A. (1994) "Estimation of Dynamic Structural Models: Problems and Prospects Part II: Mixed Continuous–Discrete Models and Market Interactions", in: C. Sims and J.J. Laffont, eds., *Proceedings of the 6th World Congress of the Econometric Society, Barcelona, Spain.* Cambridge University Press.
Pakes, A. and D. Pollard (1989) "Simulation and the Asymptotics of Optimization Estimators", *Econometrica*, 57(5), 1027–1057.
Phelps, E. (1962) "Accumulation of Risky Capital", *Econometrica*, **30**, 729–743.
Pollard, D. (1979) "General Chi-Square Goodness of Fit Tests with Data-Dependent Cells", *Z. Wahrscheinlichkeitstheorie verw. Gebeite*, 50, 317–331.
Polland, D. (1984) *Convergence of Stochastic Processes.* Springer Verlag.
Puterman, M.L. (1990) "Markov Decision Processes", in: D.P. Heyman and M.J. Sobel, eds., *Handbooks in Operations Research and Management Science Volume 2.* North-Holland/Elsevier: Amsterdam.
Rust, J. (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher", *Econometrica*, 55(5), 999–1033.
Rust, J. (1988a) "Statistical Models of Discrete Choice Processes", *Transportation Research*, 22B(2), 125–158.
Rust, J. (1988b) "Maximum Likelihood Estimation of Discrete Control Processes", *SIAM Journal on Control and Optimization*, 26(5), 1006–1023.
Rust, J. (1989) "A Dynamic Programming Model of Retirement Behavior", in: D. Wise, ed., *The Economics of Aging.* University of Chicago Press: Chicago. 359–398.
Rust, J. (1992) Do People Behave According to Bellman's Principal of Optimality?, Hoover Institution Working Paper, E–92–10.

Rust, J. (1993) How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets, manuscript, University of Wisconsin.
Rust, J. (1994) "Estimation of Dynamic Structural Models: Problems and Prospects Part I: Discrete Decision Processes", in: C. Sims and J.J. Laffont, eds., *Proceedings of the 6th World Congress of the Econometric Society, Barcelona, Spain.* Cambridge University Press.
Rust, J. (1995a) "Numerical Dynamic Programming in Economics", in H. Amman, D. Kendrick and J. Rust, eds., *Handbook of Computational Economics*, North-Holland, forthcoming.
Rust, J. (1995b) Using Randomization to Break the Curse of Dimensionality, manuscript, University of Wisconsin.
Rust, J. (1996) Stochastic Decision Processes: Theory, Computation, and Estimation, manuscript, University of Wisconsin.
Samuelson, P.A. (1969) "Lifetime Portfolio Selection by Dynamic Stochastic Programming", *Review of Economics and Statistics*, 51, 239–246.
Sargent, T.J. (1978) "Estimation of Dynamic Labor Demand Schedules Under Rational Expectations", *Journal of Political Economy*, 86(6), 1009–1044.
Sargent, T.J. (1981) "Interpreting Economic Time Series", *Journal of Political Economy*, 89(2), 213–248.
Smith, A.A. Jr. (1991) Solving Stochastic Dynamic Programming Problems Using Rules of Thumb, Discussion Paper 818, Institute for Economic Research, Queen's University, Ontario, Canada.
Sonnenschein, H. (1973) "Do Walras Law and Continuity Characterize the Class of Community Excess Demand Functions?", *Journal of Economic Theory*, 6, 345–354.
Stern, S. (1992) "A Method for Smoothing Simulated Moments of Discrete Probabilities in Multinomial Probit Models", *Econometrica*, 60(4), 943–952.
Stern, S. (1995) "Approximate Solutions to Stochastic Dynamic Programming Problems", *Econometric Theory*, forthcoming.
Stock, J. and D. Wise (1990) "Pensions, The Option Value of Work and Retirement", *Econometrica*, 58(5), 1151–1180.
Stokey, N.L., R.E. Lucas, Jr. and E.C. Prescott (1989) *Recursive Methods in Economic Dynamics.* Harvard University Press: Cambridge, Massachusetts.
Theil, H. (1971) *Principles of Econometrics.* Wiley: New York.
van Dijk, N.M. (1984) Controlled Markov Processes: Time Discretization, CWI Tract 11, Center for Mathematics and Computer Science, Amsterdam.
Verim, H. (1982) "The Nonparametric Approach to Demand Analysis", *Econometrica*, 52(3), 945–972.
White, H. (1982) "Maximum Likelihood Estimation of Misspecified Models", *Econometrica*, 50, 1–26.
Williams, H.C. (1977) "On the Formation of Travel Demand Models and Economic Evaluation of User Benefit", *Environment and Planning*, A-9, 285–344.
Wolpin, K. (1984) "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality", *Journal of Political Economy*, 92(5), 852–874.
Wolpin, K. (1987) "Estimating a Structural Search Model: The Transition from Schooling to Work", *Econometrica*, 55, 801–818.