DO PEOPLE BEHAVE ACCORDING TO BELLMAN'S PRINCIPLE OF OPTIMALITY?

John Rust

Working Papers in Economics E-92-10

The Hoover Institution
Stanford University
May 1992

This is a preliminary draft, written in March 1992, while the author was a fellow at the Hoover Institution. Parts of this article have been borrowed from my manuscript *Stochastic Decision Processes: Theory, Computation and Estimation* written as part of the Lief Johansen lectures presented at the University of Oslo in the fall of 1991. Research by Phelan and Rust described in section 5.6 was supported by NSF grants SES 8721199 and SES 911348. I benefitted from discussions and suggestions from a number of people including Colin Camerer, John Cochrane, Robert Hall, Jim Heckman, Kenneth Singleton, and Tom MaCurdy. Comments or criticisms are very much appreciated.

Hoover Institution
Working Paper No. E-92-10
May 1992

DO PEOPLE BEHAVE ACCORDING TO BELLMAN'S PRINCIPLE OF OPTIMALITY?

Abstract

This paper investigates whether the main mathematical procedure for solving sequential decision problems under uncertainty, dynamic programming (DP), provides a good model of the way humans actually solve such problems. The theory implies that behavior is determined by a *decision rule* that satisfies *Bellman's principle of optimality.* In infinite-horizon Markovian decision problems (MDP), Bellman's principle implies *Blackwell's Theorem:* an optimal decision rule is stationary and Markovian. In section 3 I show that without further restrictions on an MDP problem, Blackwell's Theorem provides a necessary and sufficient characterization of optimal decision rules, implying that Bellman's principle *per se* has no empirically testable content. However if researchers are willing to impose *a priori identifying restrictions* on the functional forms of agents' preferences and beliefs, the theory does lead to strong testable restrictions. In section 4 I review the class of *continuous decision processes* (CDP) and show that many of these models (estimated using aggregate time-series data under the representative consumer paradigm) are strongly rejected by the data. In section 5 I review the class of *discrete decision processes* (DDP) and show that the predictions of these models (estimated using micro panel data and incorporating richer forms of population heterogeneity) seem to be much more consistent with observed behavior. I conclude in section 6 with some speculations about possible reasons for these empirical findings, highlighting the potential importance of AI-type methods in yielding more realistic and workable models of larger aspects of human decision-making, and discussing the role of experimental versus econometric methods as a testing grounds for these theories.

Keywords: Dynamic Programming, Bellman's Principle of Optimality, Markovian Decision Problems, Continuous Decision Problems, Discrete Decision Problems, Maximum Likelihood and Generalized Method of Moments Estimation and Testing

JEL Classification Numbers: C10, C60, D80, D90

John Rust
Department of Economics
University of Wisconsin
Madison, Wisconsin 53706
JRUST@MACC.WISC.EDU

1. Introduction

Although the idea of solving sequential decision problems (SDP's) by "backward induction" can be traced back at least to the work of Wald (1947), Richard Bellman (1957) has been credited with the systematic application of the method which he called *dynamic programming* (DP). One can use DP to solve very general SDP's, but due a problem Bellman called the "curse of dimensionality" most practical applications of the method are restricted to the subclass of Markovian decision problems (MDP's). However, as I show in section 3, from an empirical standpoint there is essentially no loss of generality in restricting attention to this subclass. In an MDP there are two kinds of variables: a *state variable* s_t and a *control variable* d_t. A decisionmaker or *agent* is represented by *primitives* (u, p, β) where $u(s_t, d_t)$ is a utility function representing the agent's preferences, $p(s_{t+1}|s_t, d_t)$ 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, expected utility maximizers: they behave according to an *optimal decision rule* $d_t = \delta^*(s_t)$ that maximizes $E_\delta\left\{\sum_{t=0}^{T} \beta^t u(s_t, d_t)\right\}$ where E_δ is the expectation with respect to the stochastic process $\{s_t, d_t\}$ induced by the decision rule δ.

Bellman showed that the the optimal decision rule to an MDP problem satisfies the *principle of optimality:* at any time t, and in any state s_t, δ^* has the property that it must also be optimal for the continuation process treating the current state as starting point. This principle allows one to compute the optimal decision rule by backward induction starting at the terminal period T. Integral to the process of backward induction is the *value function* V which summarizes the expected future utility at any node of the decision tree, *assuming that an optimal policy will be followed in the future.* Howard (1971) has aptly described the value function as a "portable genius" since it is akin to a "shadow price" which enables one to decentralize a complicated multiperiod decision problem into a sequence of simpler static decision problems. Subsequent work by Blackwell (1962), Denardo (1967) and others showed that the logic of backward induction could be extended to infinite-horizon MDP's, putting the theory of MDP's on a solid probabilistic foundation. This work showed that the value function in stationary, infinite-horizon, discounted Markovian decision problems is the unique solution to a recursive functional equation known as *Bellman's equation,* which turns out to be mathematically equivalent to a fixed point of a contraction mapping.

One of the first economic applications of MDP's was the pioneering work on optimal inventory policy by Arrow, Harris and Marschak (1951). They derived and solved a functional equation describing the expected discounted cost of inventory holdings which we now recognize as Bellman's equation. Subsequent applications of MDP's include models of optimal intertemporal consumption, savings and portfolio selection under uncertainty,[1] studies of investment under

[1] See Phelps (1962), Hakansson (1970), Levhari and Srinivasan (1969), Merton (1969) and Samuelson (1969).

uncertainty,[2] models of optimal growth under uncertainty,[3] models of search, job-matching and turnover,[4] studies of pricing of financial assets and durable goods in production and exchange economies,[5] and models of equilibrium business cycles.[6] 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, economists began to ask whether MDP's might provide good empirical models of how real-world agents actually behave. In a 1978 issue of the *Journal of Political Economy* two seminal papers by Robert Hall and Thomas Sargent lead to significant changes in the way economists interpret data and formulate their empirical models. Hall developed a simple MDP model of optimal consumption and savings and showed that under certain assumptions the marginal utility of consumption is a martingale. Hall's empirical findings showed that the U.S. aggregate consumption expenditures appeared to be consistent with the martingale hypothesis. Sargent developed an MDP model of a firm's decision of how many workers to hire in a world with randomly changing prices and productivities. Existing neoclassical models implied that firms would hire labor to the point where the real wage equals the marginal product of labor. However Sargent developed a more realistic model that accounts for the fact that firms must incurs costs of hiring, training and integrating new workers. These "adjustment costs" imply that the firm's optimal dynamic labor demand schedule will generally differ from the static neoclassical demand curve. Instead of restricting attention to general, qualitative implications of the MDP theory, Sargent developed a maximum likelihood estimation method that allowed him to infer the underlying primitives of the MDP model, which in this case include the firm's production function and the parameters of its adjustment cost function.

Hall's study can be viewed as one of the first econometric tests of Bellman's principle of optimality, i.e. whether peoples' behavior is consistent with an optimal decision rule δ^* to an MDP. Sargent's study can be viewed as the first example of *structural estimation* of an MDP. Most economic data sets are of the form $\{d_t^a, s_t^a\}$ where s_t^a is the state and d_t^a is the decision of an agent a at time t. Existing *reduced-form* estimation methods can be viewed as uncovering agents' decision 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.[7] If people behave according to Bellman's principle,

---

[2] See Lucas and Prescott (1971).

[3] See Brock and Mirman (1972) and Leland (1974).

[4] See Lippman and McCall (1976) and Jovanovic (1979).

[5] See Lucas (1978), Brock (1982), and Rust (1985).

[6] See Kydland and Prescott (1982) and Long and Plosser (1983).

[7] For an overview of this literature, see Billingsley (1961), Chamberlain (1984), Heckman (1981), Lancaster (1990) and Prakasa Rao (1986).

then $\{d_t^a, s_t^a\}$ is a realization of a *controlled stochastic process.* While reduced-form estimation methods can be used to test general qualitative implications of Bellman's principal (such as Hall's test that the marginal utility of consumption is a martingale), an entirely new econometric methodology was required in order to do structural estimation. Structural estimation of MDP's can be regarded as an "inverse stochastic control problem" which not only uncovers the form of the stochastic process generating $\{d_t^a, s_t^a\}$, but also the primitives (u, p, β) which determine it. Typically parametric methods are used to infer (u, p, β), which are specified up to an unknown parameter vector θ.[8]

This paper provides an overview of the rapidly growing literature on structural estimation and testing of MDP's. Given the huge volume of work on this subject over the last decade, I do not pretend to offer anything more than a selective survey of a few of the most promising estimation methodologies and a sampling of empirical findings. Since estimation methodologies are covered in much more detail in recent surveys by Eckstein and Wolpin (1989) and Rust (1992), the focus of this paper is empirical: have MDP's yielded good models of real-world decisionmakers? I begin in section 2 with a brief review of dynamic programming, Bellman's principle, and *Blackwell's Theorem.* The latter theorem establishes the implications of Bellman's principle in infinite horizon MDP's: δ^* is deterministic and Markovian. In section 3 I show that without further restrictions on the primitives (u, p, β), these two properties are both necessary and sufficient characterizations of δ^*. It follows that Bellman's principle *per se* has no empirical content, in the sense that it does not impose any testable restrictions on a given body of historical data $\{x_t^a, d_t^a\}$. On the other hand, as will be clear from the remaining sections of the paper, parametric restrictions on (u, p, β) almost always lead to strong testable implications. Thus, the rhetorical question stated in the title to this paper is ill-posed: the objective of this paper might be more accurately stated as: "have researchers succeeded in discovering theoretically plausible specifications for (u, p, β) such that the implied MDP models provide accurate descriptions of peoples' behavior?" Section 4 surveys empirical findings for the subclass of *continous decision processes* (CDP's), i.e. MDP's where the decision variable d_t can assume a continuum of values. I survey empirical findings for CDP models of the decision of how much to invest in various financial assets, how much to save and consume over the life-cycle, how much a representative consumer should work over the business cycle, and what level of inventories a firm should hold in the face of randomly fluctuating sales. Section 5 summarizes empirical findings for the subclass of *discrete decision processes* (DDP's), i.e. MDP's where the decision variable d_t can assume only a countable number of values. I survey empirical findings for DDP models of the decision of an older person whether or not to retire, an unemployed person's decision of whether or not to accept a job offer or continue to search, a couple's decision of whether or not to have a child and which contraceptive method to use, and a firm's decision of whether or not to build a new plant and whether or not to repair or scrap a broken down machine, and a woman's decision whether to work in a job or work at home.

---

[8] Under certain assumptions, such as the hypothesis of *rational expectations*, agents' beliefs p can be estimated non-parametrically.

Overall, my view of the literature is that there currently is mixed support for the hypothesis that people behave "as if" they were following an optimal decision rule to an MDP: there have been notable successes, and some notable failures (i.e. rejections). A general observation is that CDP models using macro data have not fared as well as DDP models using micro data. However it is not clear whether the rejections reflect a rejection of particular parametric specifications for (u, p, β), data problems (measurement error, aggregation problems), or whether there is a deeper sense in which people don't act rationally. Section 6 provides some concluding remarks and some observations about possible new directions for future research, with particular emphasis on the use of experimental methods to provide controlled tests of the implications of Bellman's principle.

2. Bellman's Principle: A Review

The logic of backward induction underlying Bellman's principle is very general, and can be used to solve a much wider class of problems than the time-separable Markovian decision processes considered in sections 4 and 5. To appreciate the simplifications of the time-separable Markovian structure, I start by describing the method in the general case, showing that Bellman's principle is equivalent to a "subgame perfect equilibrium" of a "game against Nature."

Definition 2.1: A (discrete-time) *stochastic decision process* consists of the following objects:

- A *time index* $t \in \{0, 1, 2, \ldots, T\}$, $T \leq \infty$
- A *state space* S
- A *decision space* D
- A family of *constraint sets* $\{D_t(s_t, H_{t-1}) \subseteq D\}$
- A family of *transition probabilities* $\{p_t(ds_t | H_{t-1})\}$
- A *utility functional* $U(\mathbf{s}, \mathbf{d})$ *where* $\mathbf{s} = (s_0, s_1, s_2, \ldots) \in S^T$ and $\mathbf{d} = (d_0, d_1, d_2, \ldots) \in D^T$.

In definition 2.1 H_{t-1} denotes the *history* of realized states and decisions up to and including time $t-1$: $H_{t-1} = (s_0, d_0, s_1, d_1, \ldots, s_{t-1}, d_{t-1})$. At each time t, the agent's decisions are constrained to be an element of a choice set $D_t(H_{t-1}, s_t)$, a non-empty subset of the decision space D. The constraint sets are allowed to depend upon time t, the current state s_t, and the realized history of the process H_{t-1} up to time $t-1$. Once the agent selects a decision $d_t \in D_t(H_{t-1}, s_t)$ the history is updated to $H_t = (H_{t-1}, s_t, d_t)$ and a new state s_{t+1} is randomly drawn from the transition probability $p_{t+1}(ds_{t+1} | H_t) = p_{t+1}(ds_{t+1} | H_{t-1}, s_t, d_t)$.

The agent's optimization problem can be stated formally as follows:

$$\delta^* = \underset{\delta}{argmax}\, E_\delta\{U(\tilde{\mathbf{s}}, \tilde{\mathbf{d}})\}. \tag{2.1}$$

This is equivalent to a static optimization problem, except that rather than selecting a fixed sequence of *decisions* $\mathbf{d} = (d_1, \ldots, d_T)$, the agent can generally do better by selecting a sequence of *decision rules* $\delta^* = (\delta_1^*, \ldots, \delta_T^*)$, where each δ_t^* specifies the agent's best decision as a function of his information at time t: $d_t = \delta_t^*(H_{t-1}, s_t)$. The notation E_δ to denotes the fact that the expectation in (2.1) is taken with respect to the controlled stochastic process $\{s_t, d_t\}$ induced by δ. Although (2.1) is formally equivalent to a static optimization problem, it involves searching over an infinite-dimensional space of possible decision rules. Bellman's contribution was to show that one can reformulate the problem in way that leads to a much more straightforward and computationally tractable solution procedure.

To understand Bellman's approach, it is helpful to think of a stochastic control problem as a "game against Nature". The extensive form of this game can be represented by a *decision tree* such as is illustrated in figure 2.1. The tree consists of a series of nodes or *information sets* connected by arcs representing decisions by the agent and "Nature", respectively. Nodes marked "A" represent decision points of the agent, nodes marked "N" represent "decisions" by Nature, and the terminal nodes marked "P" represent the final payoffs to the agent. Each payoff corresponds to a separate history or path through the tree. For example the payoff $U(1,1,3,2)$ corresponds to history $H_1 = (1,1,3,2)$ starting from initial state $s_0 = 1$ after which the agent chooses $d_0 = 1$, Nature "chooses" $s_1 = 3$, and the agent chooses $d_1 = 2$. Note that each of Nature's nodes can be identified with a history H_t and each of the agent's nodes can be identified with the sub-history (H_{t-1}, s_t). A *sub-game* can be identified as the sub-tree extending from any node. For example Figure 2.1 indicates a sub-game extending from the node coresponding to history $H_0 = (1,1)$. The "game" consists of alternating moves by Nature and the agent until the terminal nodes are reached and payoffs are collected. Unlike a standard two-player game, Nature does not try to behave optimally, but rather makes a random choice s_t at node H_{t-1} according to the pre-specified transition probability $p_t(ds_t|H_{t-1})$. In figure 2.1, these probabilities are indicated along the arcs extending from each of Nature's decision nodes. The agent, on the other hand, does try to behave optimally. Using the game tree, we can state Bellman's principle in the following way:

Optimality Principle: *at each of the agent's decision nodes (H_{t-1}, s_t), the optimal decision $d_t = \delta_t^*(H_{t-1}, s_t)$ maximizes the conditional expectation of utility in the sub-game extending from that node, assuming that optimal actions will be selected at all future nodes in the sub-game.*

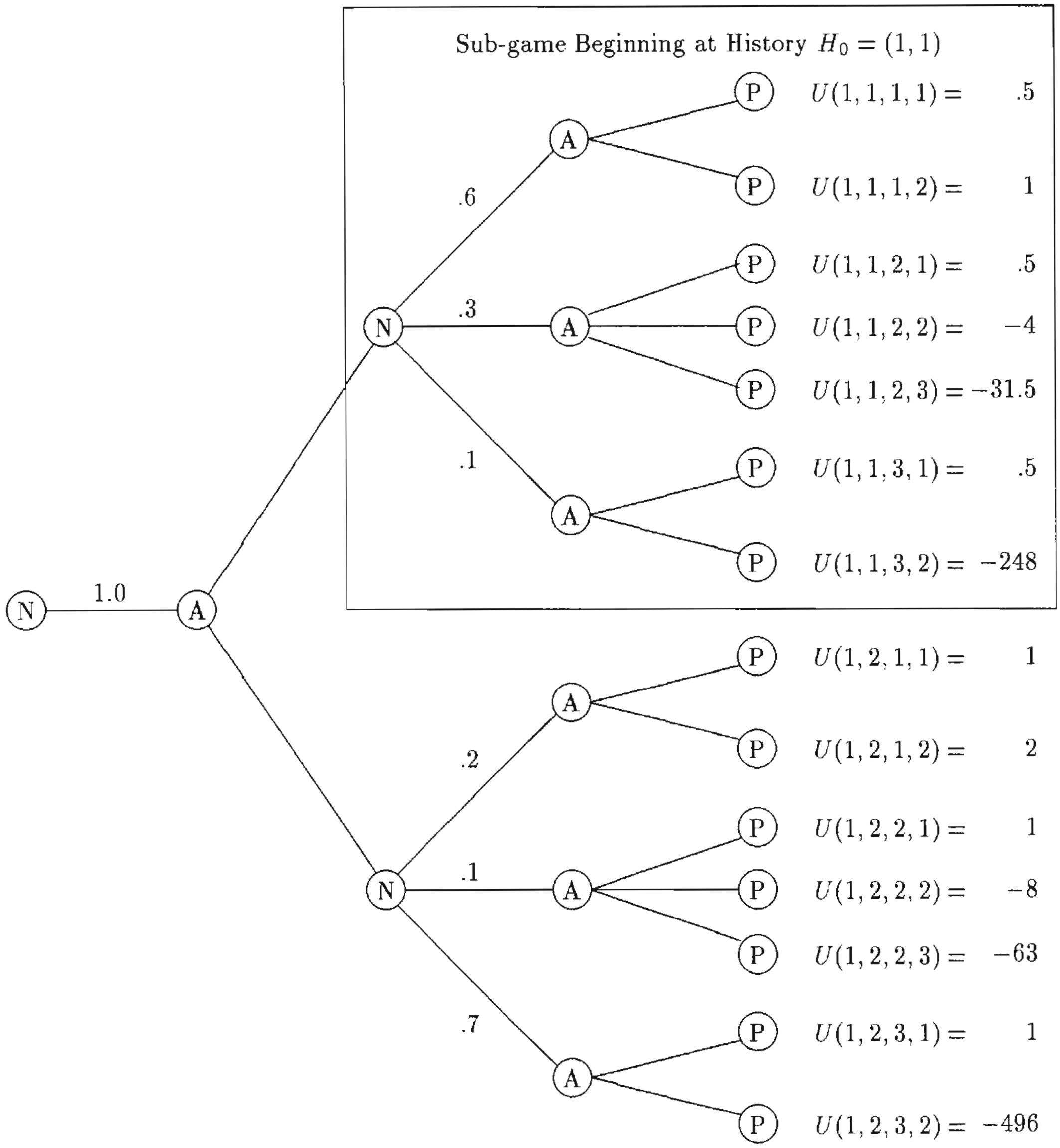


Figure 2.1: Stochastic Control as a Game Against Nature

Bellman's principle implies that we can solve for the optimal decision rule by backward induction starting at the terminal nodes of the decision tree. Integral to the backward induction process is the *value function,* which serves the role of a "shadow price" that provides the correct valuation of the future consequences of current decisions. In the final time period T the decision problem is static so no value function is necessary: given any realized history H_{T-1} and current state s_T the agent selects an action d_T from the constraint set $D_T(H_{T-1}, s_T)$ to maximize utility $U(H_{T-1}, s_T, d_T)$. This maximization problem is repeated for all terminal nodes, i.e. for all possible histories H_{T-1} and all possible states s_T. Let the function $\delta_T^*(H_{T-1}, s_T)$ denote the agent's utility-maximizing decision at node

(H_{T-1}, s_T). By the principal of optimality, $\delta_T^*(H_{T-1}, s_T)$ is an optimal decision rule at time T. Given that the agent uses this decision rule, we can use Nature's probabilistic "decision rule" $p_T(ds_T|H_{T-1})$ to compute the agent's expected utility conditional on any history H_{T-1}:

$$V_{T-1}^*(H_{T-1}) = \int U(H_{T-1}, s_T, \delta_T^*(H_{T-1}, s_T)) p_T(ds_T|H_{T-1}). \tag{2.2}$$

Thus, the value function $V_{T-1}^*(H_{T-1})$ summarizes the agent's expected utility of following an optimal strategy in the sub-game extending from node H_{T-1}. By induction, if we have computed the value function and decision rule through period $t+1$, then we can compute the same two objects in period t. The agent's optimal decision rule in period t is defined by:[9]

$$\delta_t^*(H_{t-1}, s_t) = \underset{d_t \in D_t(H_{t-1}, s_t)}{argmax} \; V_t^*(H_{t-1}, s_t, d_t). \tag{2.3}$$

The value function in period t is defined by:[10]

$$V_t^*(H_t) = \int V_{t+1}^*(H_t, s_{t+1}, \delta_{t+1}^*(H_t, s_{t+1})) p_{t+1}(ds_{t+1}|H_t). \tag{2.5}$$

Backward recursion stops at time $t = 0$ resulting in decision rule $\delta^* = (\delta_0^*, \ldots, \delta_T^*)$, the solution to the stochastic control problem. The value function at time 0, $V_0^*(s_0)$, represents the expected utility of following the optimal strategy conditional on starting the process in state s_0. Figure 2.2 summarizes the first few steps of the backward induction process.

As an example, consider solving the example problem in figure 2.1 by dynamic programming. In this case the utility functional is given by $U(\mathbf{s}, \mathbf{d}) = s_0 d_0 [d_1^{s_1} - .5 d_1^{s_1^2}]$, whose values at each terminal node are displayed in figure 2.1. In the terminal period $t = 1$, calculation of the optimal decision rule at each of the A-nodes merely involves selecting the alternative yielding the highest utility. We then assign a value to each A-node equal to the highest utility, i.e. $U_T(H_{T-1}, s_T, \delta_T^*(H_{T-1}, s_T))$. Using these values, we can then assign values to the two preceeding N-nodes using the conditional probabilities shown in figure 2.2: $V_{T-1}^*(H_{T-1}) = .8$ if $H_{T-1} = (1,1)$ and $V_{T-1}^*(H_{T-1}) = 1.2$ if $H_{T-1} = (1,2)$. It follows immediately that the optimal decision rule at time 0 is given by $\delta_0^*(s_0) = 2$, and the implied value is $V_0^*(s_0) = 1.2$. Note that under the optimal strategy, the upper subgame in figure 2.1 is never reached. Nevertheless it is necessary to fully solve this sub-game in order to assign the right value to N-node $H_0 = (1,1)$ so that the correct decision can be made at time 0. From an economist's perspective, the value function is akin to assigning shadow prices to each node of the decision tree, allowing a complex multi-period decision problem to be "decentralized" into a sequence of static decision problems.

[9] Note that in defining (2.3) and (2.4) I have used the identity $H_{t+1} = (H_t, s_{t+1}, d_{t+1})$.

[10] Note that if there are multiple actions $d_t \in D_t(H_{t-1}, s_t)$ that attain the maximum in (2.3) the optimal decision rule is not uniquely determined. Any (measurable) selection rule can be used to select a maximizing element, including randomized strategies. It is easy to verify that any such rule yields the same formula for the value function in (2.4).

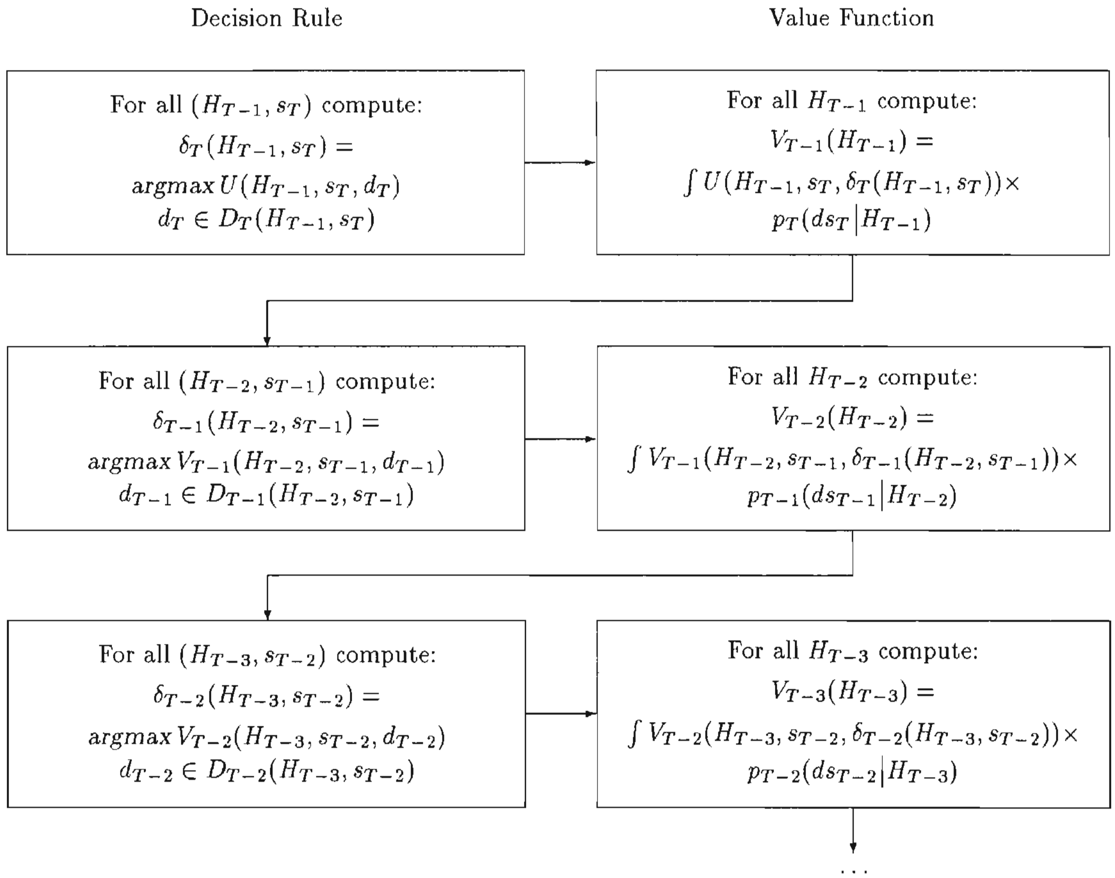


Figure 2.2: Computing Optimal Decision Rules by Backward Recursion

In theory, dynamic programming is a very powerful and general technique, allowing us to compute optimal decision rules to general sequential decision problems with non-stationary, history-dependent transition probabilities and time non-separable utility functions. However there is a practical limitation to the method which Bellman termed "the curse of dimensionality". Consider the case of a finite horizon problem with T periods, where the state space has a finite number of elements S and the decision space and each of the constraint sets D_t have D elements. After T periods there are $(SD)^T$ possible histories. For example, in a reasonably "small" problem with $S = 5$, $D = 3$, and $T = 10$, there are over $5.76E^{11}$ possible histories. To perform backward induction, the utility function must be evaluated at each of the $(SD)^T$ possible histories and a total of $S^T D^{T-1}$ maximizations performed. Then $(SD)^{T-1}$ conditional expectations must be calculated for each history H_{T-1}. Thus, the overall work involved of a backward recursion from period T to period 0 is proportional to the number of histories, or $(SD)^T$, which goes up exponentially fast as the horizon T increases. Addition of extra state or decision variables also increases the amount of work by exponential factors. For example, if we add a second state and control variable which assume S and D values, respectively, then the amount of work required to solve the expanded problem is proportional to $(SD)^{2T}$. It is clear that a large part of the problem arises from the need to keep track of all possible histories. One way to substantially reduce the amount of work required is to focus on problems where there are much lower-dimensional "sufficient statistics" for H_t.

Definition 2.2: A (discrete-time) *Markov decision process* consists of the following objects:

- A *time index* $t \in \{0, 1, 2, \ldots, T\},\ T \leq \infty$
- A *state space* S
- A *decision space* D
- A family of *constraint sets* $\{D_t(s_t) \subseteq D\}$
- A family of *transition probabilities* $\{p_{t+1}(\cdot | s_t, d_t)\}$
- A discount factor $\beta \in (0, 1)$ and family of *single period utility functions* $\{u_t(s_t, d_t)\}$ such that the utility functional U has the additively separable decomposition:

$$U(\mathbf{s}, \mathbf{d}) = \sum_{t=0}^{T} \beta^t u_t(s_t, d_t) \tag{2.7}$$

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, $d_t = \delta_t^*(s_t, H_{t-1})$, however in an MDP it is easy to see that the Markovian structure of p and the additive separability of U imply that it is not necessary 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: $d_t = \delta_t^*(s_t)$. For example, starting in period T we have:

$$\delta_T^*(H_{T-1}, s_T) = \underset{d_T \in D_T(s_T)}{argmax}\ U(H_{T-1}, s_T, d_T). \tag{2.8}$$

Rewriting U in terms of definition (2.7) we have:

$$\begin{aligned} U(H_{T-1}, s_T, d_T) &= \sum_{t=0}^{T} \beta^t u_t(s_t, d_t) \\ &= \sum_{t=0}^{T-1} \beta^t u_t(s_t, d_t) + \beta^T u_T(s_T, d_T). \end{aligned} \tag{2.9}$$

From (2.9) it is clear that previous history H_{T-1} does not affect the optimal decision of d_T in (2.8) since d_T appears only in the final term $u_T(s_T, d_T)$ in the right hand side of the last equation in (2.9). 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 which has the property that it depends on the past history of the process only via the current state s_t is called *Markovian.*

The Markov property implies that we can define the value functions V_t^* prospectively as the expected discounted utility (under an optimal policy) from time t onwards. In the terminal period V_T^* and δ_T^* are defined by:

$$\delta_T^*(s_T) = \underset{d_T \in D_T(s_T)}{argmax} \left[u_T(s_T, d_T)\right], \tag{2.10}$$

$$V_T^*(s_T) = \max_{d_T \in D_T(s_T)} \left[u_T(s_T, d_T)\right]. \tag{2.11}$$

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

$$\delta_t^*(s_t) = \underset{d_t \in D_t(s_t)}{argmax} \left[u_t(s_t, d_t) + \beta \int V_{t+1}^*(s_{t+1}, \delta_{t+1}^*(s_{t+1}))p_{t+1}(ds_{t+1}|s_t, d_t)\right], \tag{2.12}$$

$$V_t^*(s_t) = \max_{d_t \in D_t(s_t)} \left[u_t(s_t, d_t) + \beta \int V_{t+1}^*(s_{t+1}, \delta_{t+1}^*(s_{t+1}))p_{t+1}(ds_{t+1}|s_t, d_t)\right]. \tag{2.13}$$

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

$$V_0^*(s_0) = \max_{\delta} E_\delta\{U(\tilde{\mathbf{s}}, \tilde{\mathbf{d}})\}. \tag{2.14}$$

In infinite horizon MDP's there is no last period, so how does one apply dynamic programming? The general idea is that if U satisfies certain boundedness and continuity conditions, then one can alway construct an approximate finite-horizon utility function U_N and associated optimal decision rule δ_N^* which converge to U and δ^* as $N \to \infty$. Proving this in the general case is technically difficult and takes us away from the main points of this survey.[11] Instead we will illustrate the result in the special case of a *stationary* MDP where $u_t = u$ and $p_t = p$ for all t. In this case, it is easy to see that 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 $t \geq 0$ and all $s \in S$, $\delta_t^*(s) = \delta^*(s)$ and $V_t^*(s) = V^*(s)$. In analogy with equation (2.12), δ^* should satisfy:

$$\delta^*(s) = \underset{d \in D(s)}{argmax} \left[u(s, d) + \beta \int V^*(s')p(ds'|s, d)\right], \tag{2.15}$$

where V^* is defined recursively by the solution to *Bellman's equation:*

$$V^*(s) = \max_{d \in D(s)} \left[u(s, d) + \beta \int V^*(s')p(ds'|s, d)\right]. \tag{2.16}$$

---

[11] For proofs of the general case see Gihman and Skorohod (1979), Kreps (1977a,b, 1978), or Rust, (1992).

It is not difficult to show that under certain regularity conditions (boundedness of u), Bellman's equation (2.16) is mathematically equivalent to a fixed point of a contraction mapping, which implies that it exists and is unique. The following theorem, known as *Blackwell's Theorem,* shows that the stationary decision rule δ^* computed from Bellman's equation in (2.15) is indeed an optimal decision rule to the original optimization problem:

Theorem 2.1 (Blackwell, Bhattacharya and Majumdar): *Given an infinite-horizon, stationary MDP satisfying the regularity conditions in Bhattacharya and Majumdar (1989) we have:*

1. *A unique solution V^* to Bellman's equation (2.16) exists, and it coincides with the agent's expected disounted utility under an optimal policy.*

2. *An optimal decision rule δ^* exists and coincides with the stationary, non-randomized, Markovian optimal control δ^* given by the solution to (2.15).*

3. *δ^* can be approximated arbitrarily closely by the solution δ^*_N (computed by backward induction) to an N-period problem with utility function $U_N(\mathbf{s},\mathbf{d}) = \sum_{t=0}^{N} \beta^t u(s_t, d_t)$ such that:*

$$\lim_{N\to\infty} E_{\delta^*_N}\{U_N(\tilde{\mathbf{s}},\tilde{\mathbf{d}})\} = \lim_{N\to\infty} \sup_{\delta} E_{\delta}\{U_N(\tilde{\mathbf{s}},\tilde{\mathbf{d}})\} = \sup_{\delta} E_{\delta}\{U(\tilde{\mathbf{s}},\tilde{\mathbf{d}})\}. \tag{2.17}$$

3. The Empirical Content of Bellman's Principle

Since we are interested in MDP's from a positive rather than normative perspective, we now ask whether Bellman's principle leads to any empricially testable implications. In addition, given our interest in structural estimation, we want to know under what conditions agents' primitives are *identified.* We can pose the latter question as a dynamic form of "revealed preference": given infinite observations $\{s_t, d_t\}$ on the states and decisions of an agent, under what conditions can we use this data to uniquely recover (u, p, β)? Unfortunately, without strong *a priori* restrictions on (u, p, β), the answer to both these questions is negative: the MDP model is "non-parametrically unidentifed" in the sense that there is an equivalence class containing infinitely many distinct primitives (u, p, β) that are consistent with the observations $\{s_t, d_t\}$, and Bellman's principle *per se* has no empirical content in the sense that we can always find primitives that "rationalize" any set of observations. These results justify my intorductory comment that from an empirical standpoint, restricting attention to the class of MDP's involves no loss of generality.

Note that these negative findings stand in contrast to the case of static choice models where we know that the hypothesis of optimization *per se* does imply testable restrictions.[12] The absence of restrictions in the dynamic case

[12] 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-Marschak 1960, and McFadden and Richter, 1970).

may seem surprising given that the structure of the MDP problem already imposes a number of strong restrictions such as time additive preferences, constant intertemporal discount factors, as well as the expected utility hypothesis itself. However many economists will probably not be surprised by this result since similar results have appeared in other areas such as the literature on choice under uncertainty, game theory, and general equilibrium theory.[13]

To simplify notation, we establish these results in the context of stationary infinite-horizon MDP's, although the argument carries over almost without modification to the nonstationary, finite horizon case. In order to formulate the identification problem *á la* Cowles Commission, we need to translate the concepts of *reduced-form* and *structure* to the context of a nonlinear MDP model.

Definition 3.1: The *reduced-form* of an MDP model is the agent's optimal decision rule, δ.

Definition 3.2: The *structure* of an MDP model is the mapping: $\Lambda(u,p,\beta)=\delta$ defined by:

$$\delta(s)=\underset{d\in D(s)}{argmax}[v(s,d)], \tag{3.1}$$

where v is the unique fixed point to:

$$v(s,d)=u(s,d)+\beta\int \underset{d'\in D(s')}{max}[v(s',d')]p(ds'|s,d). \tag{3.2}$$

The rationale for identifying δ as the reduced-form of the MDP is that it embodies all observable implications of the theory[14] and can be consistently estimated by non-parametric regression given sufficient number of observations $\{s_t,d_t\}$.[15] We can use the reduced-form δ to define an equivalence relation over the space of primitives:

Definition 3.3: Primitives (u,p,β) and $(\overline{u},\overline{p},\overline{\beta})$ are *observationally equivalent* if:

$$\Lambda(u,p,\beta)=\Lambda(\overline{u},\overline{p},\overline{\beta}). \tag{3.3}$$

---

[13] For example it is well-known that in "subjective expected utility" models with state dependent preferences, one cannot separately identify the agent's subjective beliefs p from preferences u (see Kreps, (1988) Proposition 7.4 or Myerson (1991) Theorem 1.2). In game theoretic models Ledyard (1986) has shown that any undominated strategy profile can be rationalized as a Bayesian Nash equilibrium outcome of a game of incomplete information given enough freedom in selecting agents' beliefs (priors) and utility functions. Heckman and MaCurdy (1987) showed that a similar result obtains in the case of dynamic general equilibrium models. Mantel (1973) and Sonnenschein (1974) and others have shown that at an aggregate level, static demand theory implies no testable restriction beyond Walras Law. Finally, even in the simplest context of deterministic optimal growth model with a single capital stock k_t, Boldrin and Montrucchio (1986) showed that any continuous law of motion $k_{t+1}=f(k_t)$ is optimal for some specification of preferences and technology provided the discount factor β is not too close to 1.

[14] We use δ^* as the reduced-form rather than the transition probability $Pr\{s_{t+1},d_{t+1}|s_t,d_t\}=I\{d_{t+1}=\delta^*(s_{t+1})\}\ p(s_{t+1}|s_t,d_t)$ since if p represents the agent's subjective beliefs, there is no objective way to estimate it except through its observable implications via δ^*. However if we assume that agents have "rational expectations" then p can be estimated non-parametrically (Basawa and Prakasa Rao, 1983), and all remaining observable implications of the theory are determined by δ^*.

[15] Since the econometrician observes all components of (s,d) the non-parametric regression used to estimate δ^* is degenerate in the sense that the model $d=\delta^*(s)$ has an "error term" that is identically 0. Nevertheless, a variety of nonparametric regression methods will be able to consistently estimate δ^* under very general regularity conditions. See, for example, Häerdle (1990).

Thus $\Lambda^{-1}(\delta)$ is the equivalence class of primitives consistent with decision rule δ. Expected utility theory implies that $\Lambda(u,p,\beta) = \Lambda(au+b,p,\beta)$ 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 transformations of u.

Definition 3.4: The stationary MDP problem (3.1) and (3.2) is *non-parametrically identified* if given any reduced-form δ in the range of Λ, and any primitives (u,p,β) and $(\overline{u},\overline{p},\overline{\beta})$ in $\Lambda^{-1}(\delta)$ we have:

$$\begin{aligned} &\beta = \overline{\beta} \\ &p = \overline{p} \\ &u = a\overline{u} + b, \quad \textit{for some constants } a \textit{ and } b \textit{ satisfying } a > 0. \end{aligned} \tag{3.4}$$

Lemma 3.1: *The MDP problem (3.1) and (3.2) is non-parametrically unidentified.*

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

$$\begin{aligned} &\overline{\beta} = \beta \\ &\overline{p}(ds'|s,d) = p(ds'|s,d) \\ &\overline{u}(s,d) = u(s,d) + f(s) - \beta \int f(s')p(ds'|s,d), \end{aligned} \tag{3.5}$$

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

$$\begin{aligned} \overline{v}(s,d) &= \overline{u}(s,d) + \beta \int \max_{d' \in D(s')} [\overline{v}(s',d')]p(ds'|s,d) \\ &= u(s,d) + f(s) - \beta \int f(s')p(ds'|s,d) + \beta \int \max_{d' \in D(s')} [v(s',d') + f(s')]p(ds'|s,d) \\ &= u(s,d) + f(s) + \beta \int \max_{d' \in D(s')} [v(s',d')]p(ds'|s,d) \\ &= v(s,d) + f(s). \end{aligned} \tag{3.6}$$

Since v is the unique fixed point to (3.2), it follows that $v + f$ is the unique fixed point to (3.6), so our conjecture $\overline{v} = v + f$ is indeed the unique fixed point to (3.2) equation with primitives $\{\overline{\beta},\overline{u},\overline{p}\}$. From (3.1) it is clear that $\{v(s,d), d \in D(s)\}$ and $\{v(s,d) + f(s), d \in D(s)\}$ yield the same decision rule δ. It follows that $(u + f - \beta Ef, p, \beta)$ is observationally equivalent to (u,p,β), 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 infinite-horizon MDP's, Blackwell's Theorem does provide two general restrictions: δ^* is Markovian and 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, chapter 4, Whittle, section 11.4) can be used to transform an N^{th} order Markov process into a 1^{st} order Markov process. For example, by defining a new state x_t by $x_t = (H_{t-1}, s_t)$, we see that the general SDP given in definition 2.1 is actually a special case of an MDP where the objective is to maximize the utility of the terminal state and decision, $U(x_T, d_T) = U(H_{T-1}, s_T, d_T) = U(\mathbf{s}, \mathbf{d})$. The second restriction might be tested by looking for agents who make different choices in some state s in two different time periods: $\delta(s_t) \neq \delta(s_{t+k})$, 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.[16] The following lemma shows that Bellman's principle implies no other restrictions beyond the two essentially untestable restrictions of Blackwell's Theorem:

Lemma 3.2: *Given an arbitrary measurable mapping $\delta : S \to D$ there exists primitives (u, p, β) such that $\delta = \Lambda(u, p, \beta)$.*

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

$$u(s, d) = I\{d = \delta(s)\} - \beta. \tag{3.7}$$

Then it is easy to see that $v(s, d) = I\{d = \delta(s)\}$ is the unique solution to Bellman's equation (3.2), 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 3.1 shows that the resulting MDP model is non-parametrically unidentified and Lemma 3.2 shows that it has no testable implications, in the sense that we can "rig" an MDP model to "rationalize" any decision rule δ. The identification problem was clearly foreseen by Sargent in his 1978 *JPE* article:

> "The empirical results are moderately comforting to the view that the employment-real-wage observations lie along a demand schedule for employment. It is important to emphasize that this view has content (i.e. imposes overidentifying restrictions) because I have *a priori* imposed restrictions on the orders of the adjustment-cost processes and on the Markov processes governing disturbances. At a general level without such restrictions, it is doubtful whether the equilibrium view has content." (p. 1042)

Rather than end this survey right here, I suggest two ways around the problem. First, Bellman's principal does lead to empirically testable implications in the case of *laboratory experiments* where we have at least partial control over an agent's preferences or beliefs. A famous example of such an experiment is the *Allais paradox* which usually succeeds in rejecting the hypothesis that subjects are expected utility maximizers.[17] Second, it is clear that from an

[16] In addition, if the data does not record all relevant components of s, the agent's actions may appear to be stochastic from the standpoint of the econometrician even though it is deterministic from the standpoint of the agent. I treat this case in detail in section 5.

[17] See Machina (1987) for a survey of other examples of laboratory experiments of choice under uncertainty. Machina (1982) identifies the *independence axiom* as a potential source of many of the rejections.

economic standpoint, many of the utility functions $u + f - \beta Ef$ will be completely unreasonable, as is the utility function $u(s,d) = I\{d = \delta(s)\}$. By imposing additional *identifying restrictions* on (u, p, β) we can usually succeed in identifying a unique set of primitives that are consistent with the data and are plausible on *a priori* grounds. At the present time, almost all identifying restrictions are imposed by assuming that (u, p, β) belongs to a parametric family (i.e. u is quadratic, Cobb-Douglas, CES, etc.). The use of laboratory experiments as an additional source of identifying restrictions for structural models is still in its infancy. Heckman (1991) has begun pioneering work on integration of experimental results and *a priori* identifying restrictions in structural estimation problems. However as I discuss in the conclusion, it seems unlikely that even extensive experimentation will be sufficient by itself to identify the structure of agents' decisionmaking processes.[18]

The point of this section is to show that there are strict limits on what we can learn from the data alone: the empirical content of any MDP theory is entirely a result of *a priori* identifying restrictions on (u, p, β). Although the results point out the futility of non-parametric estimation of (u, p, β), I definitely do not view them as suggesting that structural estimation and testing of parametric models is a futile exercise, and given the huge and rapidly growing empirical literature on structural estimation, it is clear that most economists do not view it as a futile exercise either. The proofs of the lemmas do indicate that in order to obtain testable implications, we can get by with weaker identifying restrictions on p than on u and β. To see this, suppose 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 identifying restriction allows us to do structural estimation by *semi-parametric* methods: i.e. we can use nonparametric methods to consistently estimate $p(ds_{t+1}|s_t, d_t)$ using observed realizations of the controlled process $\{s_t, d_t\}$, and parametric methods to estimate (u, β). However looking at the proofs of Lemmas 3.1 and 3.2 we can see that it's not possible to go non-parametric on (u, β) since the identification problem persists even if we assume that p is known *a priori*. Thus we must impose strong identifying restrictions on u and β: the usual way this is done is to assume that u and β are smooth functions of a vector of unknown parameters θ. As will be evident from sections 4 and 5 this is sufficient to produce strong, empirically testable restrictions on the controlled process $\{s_t, d_t\}$.[19]

---

[18] Experiments have been predominantly used as a means of rejecting candidate structural models, rather than identifying their primitives. A classic example is the study by LaLonde (1986) and LaLonde and Maynard (1987) who found that forecasts of earnings differentials due to job training programs produced by "selection regression" models were substantially different from the results of randomized experiments. However in my opinion they go too far in interpreting their findings as "further evidence that the current skepticism surrounding the results of nonexperimental evaluations is justified." Heckman (1991) has shown that experimental methods have many limitations that structural forecasting methods avoid, so that our best hopes for policy evaluation is via an integration of the methods rather than a rejection of structural forecasts. Note that in experienced hands (Heckman and Dabos, 1987), the forecasts of the structural selectivity model are found to be much closer to the experimental findings.

[19] Identification of β is particularly problematic and in many cases its value is specificied *a priori*. For example in life-cycle models discussed in sections 4 and 5, it is very difficult to discriminate between a model with a high β without bequests from a model with a low β that allows for bequests: both models imply a slow rate of asset decumulation that we observe in the data. If we assume a particular parametric functional form for the bequest function B we may be able to identify β "by virtue of functional form", but the estimates are likely to be very sensitive to the specification of B.

4. Continuous Decision Processes: Implications and Evidence

In CDP's the decision variable can assume a continuum of values. For these problems it turns out that it is much easier to estimate and test implications of Bellman's principle rather than explicitly solve the CDP problem for δ^* by dynamic programming. A key implication is that a first order condition derived from the maximization of Bellman's equation, the so-called *Euler equation,* should be 0 when evaluated at $d_t = \delta^*(s_t)$. However this implication will generally not hold for arbitrary MDP's with continuous choice sets: we need to insure that the optimal decision is an interior point of the choice set and that d_t enters the probability distribution $p(ds_{t+1}|s_t, d_t)$ in a particular way. I now define a class of MDP's for which Euler equations are well-defined.

Definition 4.1: *A **Continuous Decision Process** is an MDP satisfying the following restrictions:*[20]

- *The decision space D is a subset of R^M.*
- *The state space is the product space $S = Y \times Z$, where Y is a closed subset of R^J and Z is a closed subset of R^K.*
- *$D(s) = D(y, z)$ is an upper hemicontinuous correspondence which is increasing in its first argument: $y \leq y' \Rightarrow D(y, z) \subset D(y', z)$.*
- *$D(\cdot, \cdot)$ is convex in its first argument: i.e. for all $z \in Z$ and all $y, y' \in D$ if $d \in D(y, z)$ and $d' \in D(y', z)$ then $\theta d + (1-\theta)d' \in D(\theta y + (1-\theta)y', z)$ for all $\theta \in [0,1]$.*
- *The transition probability $p(ds_{t+1}|s_t, d_t)$ factors as:*

$$p(ds_{t+1}|s_t, d_t) = p(dy_{t+1}, dz_{t+1}|y_t, z_t, d_t) = I\{dy_{t+1} = r(y_t, d_t, z_{t+1})\}q(dz_{t+1}|z_t). \tag{4.1}$$

  where $q(\cdot|z)$ is a weakly continuous in z, and r is continuously differentiable in (y, d).
- *For each (y,z,d) there exists an $M \times J$ matrix $h(y, d)$ satisfying:*

$$\frac{\partial r(y, d, z)}{\partial y} = \frac{\partial r(y, d, z)}{\partial d} h(y, d). \tag{4.2}$$

- *The utility function $u(s, d) = u(y, z, d)$ is a strictly concave function of its first and third arguments for each $z \in Z$ and strictly increasing in its first argument for each $(z, d) \in Z \times D$.*

The interpretation of this class of problems is that the state variable $s_t = (y_t, z_t)$ consists of two components: an "endogenous" state variable y_t and an "exogenous" state variable z_t. The exogenous state variable is so named

[20] This is a generalization of a class of problems analyzed in chapters 9 and 10 of Stokey and Lucas, 1989.

because its law of motion $q(dz_{t+1}|z_t)$ is unaffected by the agent's decision d_t. The exogenous state variable y_t is affected by the agent's decisions, but in a particular way: with probability 1 we have: $y_{t+1} = r(y_t, d_t, z_{t+1})$. [21] Bellman's equation for this class of problems is given by:

$$V^*(s) = V^*(y,z) = \max_{d \in D(y,z)} \big[u(y,z,d) + \beta \int V^*\big(r(y,d,z'),z'\big)q(dz'|z)\big]. \tag{4.3}$$

Computing the first order condition for the optimal d in (4.3) and invoking the "envelope theorem", it is not difficult to derive the Euler equation for this problem:[22]

$$0 = \frac{\partial u(y,z,d)}{\partial d} + \beta \int \left[\frac{\partial u(y',z',d')}{\partial y} - \frac{\partial u(y',z',d')}{\partial d} h(y',d')\right] \frac{\partial r(y',d,z')}{\partial d} q(dz'|z), \tag{4.4}$$

where $d = \delta^*(y,z)$, $y' = r(y, \delta^*(y,z), z')$, and $d' = \delta^*(y', z')$. The Euler equation is simply a first order necessary condition characterizing the optimal decision rule $\delta^*(y,z)$. It says that the change in expected discounted expected utility from a small change in d will be zero when $d = \delta^*(y,z)$.

The remarkable feature of the Euler equation is that the impact on expected discounted utility can be evaluated only in terms of the value of marginal utility in period t and the expected discounted value of marginal utility in period $t+1$: it is not necessary to consider the impact on marginal utility on all future periods $t+2, t+3, \ldots$. This idea forms the basis for a very effective estimation and testing procedure pioneered by Hansen and Singleton (1982). Suppose the functions u, r, h, and q depend on a vector of unknown parameters θ. Then when evaluated at the agent's true parameter vector θ^* the Euler equation (4.4) can be written schematically as:

$$E\{g(y_t, z_t, d_t, \tilde{y}_{t+1}, \tilde{z}_{t+1}, \tilde{d}_{t+1}, \theta^*)|H_t\} = E\{g(x_t, \tilde{x}_{t+1}, \theta^*)|H_t\} = 0, \tag{4.5}$$

where $x_t = (y_t, z_t, d_t)$. Equation (4.5) states that under an optimal decision rule δ^*, the conditional expectation of the random variable $g(x_t, \tilde{x}_{t+1}, \theta)$ should be zero when evaluated at the true parameter $\theta = \theta^*$. The law of iterated expectations implies that the unconditional expectation of g is also zero:

$$E\{g(x_t, x_{t+1}, \theta^*)\} = E\big\{E\{g(x_t, x_{t+1}, \theta^*)|H_t\}\big\} = 0. \tag{4.6}$$

The "generalized method of moments" (GMM) (Hansen, 1982) can be used to estimate θ^* by finding a value $\hat{\theta}$ that sets the sample average $\sum_{t=1}^T g(x_t, x_{t+1}, \hat{\theta})/T$ as close to zero as possible. In practice, we multiply g by a vector

---

[21] The Stokey-Lucas (1989) model is a special case when $y_{t+1} = d_t$ with probability 1.

[22] For derivation and proof, see Rust (1992).

of *instrumental variables* w_t, noticing that the orthogonality condition (4.5) will continue to hold provided $w_t \in H_t$. Thus, the GMM estimator $\hat{\theta}$ is defined by:

$$\begin{aligned}\hat{\theta} &= \underset{\theta}{argmin}\, F_T(\theta)' W_T F_T(\theta), \\ F_T(\theta) &= \frac{1}{T}\sum_{t=1}^{T} f(x_t, x_{t+1}, y_t, \theta), \\ f(x_t, x_{t+1}, y_t, \theta) &= g(x_t, x_{t+1}, \theta) \otimes w_t,\end{aligned} \tag{4.7}$$

where W_T is a symmetric, positive definite weighting matrix. A particularly convenient feature of the GMM estimator is that in cases where are more orthogonality conditions (components of f) than parameters in θ, the Euler equation leads to a natural test of *overidentifying restrictions* that $F_T(\hat{\theta})$ should be approximately zero. The test statistic is simply $TF_T(\hat{\theta})'W_T F_T(\hat{\theta})$ which has an asymptotic Chi-square distribution with $O - N$ degrees of freedom where O is the number of orthogonality conditions, and N is the dimension of θ. Due to the simplicity of the Euler equation approach, it has been the dominant approach to estimation and testing of CDP's. However since it imposes a necessary condition for optimization, the method does not allow one to examine all implications of the theory. For this reason there has also been a substantial amount of work which attempts to numerically solve MDP models, estimating (or calibrating) its parameters so that the implied Markov process $\{s_t, d_t\}$ is able to "best fit" the data.[23]

4.1 Empirical Application: Models of Optimal Consumption and Portfolio Selection

The first application of GMM estimation and testing of the Euler equation was the model of optimal consumption and portfolio selection of Hansen and Singleton (1982, 1983). In this case the decision variable $d_t = (c_t, \mathbf{q}_t)$ where c_t denotes consumption expenditures and $\mathbf{q}_t = (q_{1t}, \ldots, q_{Nt})$ denotes the number of shares of N securities held at time t. The consumer's budget constaint is:

$$c_t + \sum_{n=1}^{N} p_{nt} q_{nt} \le \sum_{n=1}^{N} (p_{nt} + d_{nt}) q_{nt-1} + w_t \tag{4.8}$$

where p_{nt} is the price of security N at time t, e_{nt} are dividends paid per share of security n at time t, and w_t is wage earnings. The vector $z_t = (w_t, \mathbf{p}_t, \mathbf{e}_t)$ is assumed to evolve according to an exogenous Markov process with transition probability $q(dz_{t+1}|z_t)$. The stochastic Euler equation (4.4) for this problem can be written as:

$$0 = p_{jt} u'(c_t) - \beta E\{(\tilde{p}_{jt+1} + \tilde{d}_{jt+1}) u'(\tilde{c}_{t+1}) \big| H_t\}. \tag{4.9}$$

[23] It is also possible that the unconditional expectation of (4.6) is close to zero even though the conditional expectation (4.5) may be non-zero (for a possible example of this problem, see section 4.5). One effectively needs an infinite number of moment conditions to insure that the conditional expectation (4.5) is zero for all histories H_t. Bierens (1990) has developed a consistent test statistic for the hypothesis that the conditional moment restrictions hold for each H_t, but to my knowledge there is no corresponding estimation procedure that insures that this restriction holds with probability 1.

Hansen and Singleton assumed that preferences u satisfy constant relative risk aversion:

$$u(c) = \frac{c^{\gamma} - 1}{\gamma}, \tag{4.10}$$

in which case the unconditional expectation of the Euler equation reduces to:

$$E\left\{\beta\left(\frac{c_{t+1}}{c_t}\right)^{\alpha}\mathbf{R}_{t+1} - 1\right\} = 0, \tag{4.11}$$

where $\alpha = \gamma - 1$ and $\mathbf{R}_{t+1} = (R_{1t+1}, \ldots, R_{Nt+1})$, where $R_{jt+1} = (p_{jt+1} + d_{jt+1})/p_{jt}$ is the realized rate of return on security j. Using monthly observations on aggregate U.S. consumption expenditures on nondurables and services c_t and returns for 3 portfolios of securities in the chemicals, transportation, and retail trade sectors, they estimated the 2×1 parameter vector $\theta = (\beta, \alpha)$ by the method of moments procedure described above. The instruments y_t included a constant, and L lagged values of consumption and the three stock returns. The following table presents their estimation results for the cases $L = 1$ and $L = 4$:[24]

L	$\hat{\alpha}$	$\hat{\sigma}(\hat{\alpha})$	$\hat{\beta}$	$\hat{\sigma}(\hat{\beta})$	χ^2	DF	Prob
1	$-.9993$	.2632	.9941	.0028	19.591	13	.8941
4	$-.4600$	.1388	.9961	.0024	82.735	49	.9982

Table 4.1: GMM Estimation Results from Hansen and Singleton, 1982

From table 4.1 we can see that the estimates of α are very sensitive to the number of instruments chosen, and the large values of the Chi-squared test of overidentifying restrictions in the case $L = 4$ provide strong evidence against the orthogonality conditions (4.11). These findings have been confirmed in a numerous replications. In Singleton's (1990) survey of the literature, he concludes that there are "striking differences in the point estimates of (β, γ) across bills and stocks", noting a general pattern that "whenver $\hat{\beta}$ exceeds unity, $\hat{\gamma}$ is less than unity and vice versa" (p. 610). Grossman, Melino and Shiller (1987) argued that a continuous-time MDP is the appropriate model of consumption and portfolio decisions, and that substantial biases can result if one uses time-averaged data without accounting for the effects of the time-averaging in the estimation procedure. However even after doing this, they found that their maximum likelihood estimates of γ displayed substantial instability, ranging from -1 to -184 depending on the sample period.[25] Specification tests of their model lead to overwhelming rejections.

---

[24] The final column of table 5.1, Prob, denotes the probability that a χ^2 random variable with degrees of freedom in column DF of the table is less than the reported value of the test statistic in the column χ^2.

[25] The GMM estimates also displayed even greater variability, ranging from -1 to -948.

Despite numerous studies and the immense amount of creative talent devoted to this area, Singleton (1990) concluded that "co-movements in consumptions and various asset returns are not well described by a wide variety of represenative agent models of asset price determination." (p. 622). It is still not clear whether the rejections are due to misspecification of preferences (4.10), measurement error in aggregate consumption c_t, or a rejection of the implicit assumption that c_t can be viewed 'as if' it resulted from decisions of a single "representative consumer".[26]

4.2 Empirical Application: Representative Agent Asset Pricing and the Equity Premium Puzzle

Mehra and Prescott (1985) provide an example of an alternative methodology for empirical evaluation of CDP models: numerically solve for the optimal decision rule δ^* and see how well the controlled Markov process $\{s_t, d_t\}$ accords with the data. Since the CDP model typcially involves one or more unknown parameters, one estimates or "calibrates" these parameters so that the CDP model "best fits" the data according to some metric. Mehra and Prescott used a modified version of the Lucas (1978) asset pricing model to compute a representative agent's optimal decision rule for holding a risky security (stock) and a riskless security (Treasury bill). They modelled stock growth rates as an exogenous two-state Markov chain with transition probability q, and used the CRR specification for utility (4.10). They calibrated their model by choosing the unknown parameters of q "so that the average growth rate of per capita consumption, the standard deviation of the growth rate of per capita consumption and the first order serial correlation of this growth rate, all with respect to the model's stationary distribution, matched the sample values for the U.S. economy between 1889-1978." Given these values, Mehra and Prescott chose the remaining parameters of preferences (γ, β) so that "the model's averaged risk-free rate and equity risk premium match those observed for the U.S. economy for this ninety-year period." Unfortunately they found that no combination of (γ, β) succeeded in matching the data: "The observed real return of 0.80 percent and equity premium of 6 percent is clearly inconsistent with the predictions of the model. The largest premium obtainable with the model is 0.35 percent, which is not close to the observed value." Their finding has since become known as the *equity premium puzzle*.

Subsequent research into the question has demonstrated that it is possible to rig MDP models to "resolve" the equity premium puzzle. One way to resolve the puzzle is to make agents extremely risk-averse. Cochrane and Hansen (1992) show that within the CRR family, a value of γ around -210 is sufficient to resolve the puzzle, in the sense that the resulting model satisfies the non-parametric bounds on returns derived by Hansen and Jagannanthan (1991). Reitz (1989) assumed that people believe that with small probability there could be a market crash, generating a massive decline in consumption. He shows that these beliefs together with strong risk aversion (i.e. a γ of about

[26] Singleton's survey outlined several directions in which current research in this area is proceeding: 1) allowing for incomplete markets, 2) allowing for fixed costs and nonconvexities, 3) developing heterogeneous agent models with heterogeneous beliefs and 4) modelling motives for transacting in fiat money and pricing of goods in monetary economies.

-10) succeed in "solving" the equity premium puzzle.[27] Kocherlakota (1990a) used an N-state approximation of the Markov transition probability q based on a much more accurate approximation procedure developed by Tauchen and Hussey (1991). Using a calibration approach similar to Prescott and Mehra (1985), Kocherlakota found that he was able to "resolve" the equity premium puzzle with the parameters $\beta = 1.139$ and $\gamma = -12.7$. Although Kocherlakota (1990b) argues that "discount factors larger than one are not nonsensical if consumption is growing over time" (p. 291), most economists would probably agree with Epstein and Zin (1990) that "such a specification is widely viewed with misgivings" given that "reasonableness of the utility function is an integral part of the puzzle posed by Mehra and Prescott" (p. 401).

The latest work in this area has actually departed from the time-separable MDP framework.[28] Weil (1989) and Epstein and Zin (1990a,b) use a specification of preferences which are time non-separable, but recursive, and generalize the expected utility hypothesis by allowing other functionals of random future utility than just the expectation. Neither study finds that these more general preferences are able to resolve the equity puzzle: "Our findings indicate that we can account for a low risk-free rate and and an average equity premium of roughly 2%. This is in contrast to the historical average risk premium of 6.2% and the largest premium obtainable by Prescott and Mehra of 0.35%. Thus our utility specification can only partially resolve the puzzle." (Epstein and Zin, p. 406). However Constantinides (1990) found that he could resolve the equity premium puzzle in a continuous-time model with time non-separable preferences that reflect a form of "habit persistence". Using a CRR utility with moderate risk aversion ($\gamma = -1.2$) his model is able to generate "the requisite high variablity in the marginal rate of substitution in consumption with relatively low variability in the consumption growth rate" due to the fact that habit persistence leads to "a subsistence level of consumption (which must be about 80 percent of the normal consumption rate to explain the mean equity premium). A small drop in consumption generates a large drop in consumption net of the subsistence level and a large drop in the marginal rate of substitution that makes it possible to match the observed equity premium with low risk aversion." (p. 535). However the book is not closed on the issue. Subsequent studies by Heaton (1991) and Cochrane and Hansen (1992) have questioned the reasonableness of the habit-persistence specification. The latter study found that significantly lower values of γ (in range $\gamma \in [-22, -7]$) are necessary to satisfy the Hansen-Jagannathan (1991) asset-pricing bounds, and concluded "one might argue that these preferences have traded extreme risk aversion for extreme habit persistence, since utility falls to $-\infty$ at 50% to 60% of last quarter's consumption." (p. 34). Heaton

[27] Mehra and Prescott (1988) find this assumption problematic: "In Reitz's examples, the smallest annual decline in consumption is 25 percent and the largest over 98 percent. Declines of this magnitude have not been experienced in the United States. During the last 100 years, a period that includes the Great Depression, consumption has fallen by more than 5 percent in a year only four times. And the largest of those four declines was only 8.8 percent." (p. 134).

[28] Although these problems are beyond the scope of this survey, I decided to mention their work since it is relevant to comments on the future direction of the literature in section 6.

(1991) finds that the habit persistence model implies an unrealistically high level of intertemporal variation in the implied single period risk-free interest rate.

4.3 Empirical Application: Testing the Permanent Income/Life-Cycle Hypothesis

A distinct branch of literature has focused on testing of the life-cycle/permanent income hypothesis. This literature is conceptually very closely related to the literature reviewed in sections 4.1 and 4.2, but focuses on the intertemporal allocation of consumption rather than the allocation of wealth between risky and risk-free securities. Phelps (1962) was one of the first to formulate the life-cycle consumption/savings problem as an MDP problem. He assumed that consumers could invest in a single risky security with return $\tilde{R}$ which is *IID* with distribution F. In this case the control d_t is consumption, which we denote by c_t, and the state variable s_t is the consumer's current wealth, w_t. Bellman's equation for this problem is given by:

$$V^*(w) = \max_{0 \le c \le w} [u(c) + \beta \int V^*(R(w-c))\, F(dR)]. \tag{4.12}$$

If $u(c) = log(c)$ we can explicitly solve (4.12) for δ^*:

$$\delta^*(w) = (1-\beta)w. \tag{4.13}$$

Thus, for logarithmic utility we get a strong form of Friedman's (1957) *permanent income hypothesis:* optimal consumption equals permanent income $(1-\beta)w$, independent of the value of *transitory income* $\tilde{R}w$.

Instead of explicitly solving the CDP model for the optimal consumption rule δ^*, Hall's (1978) test of the life-cycle/permanent income hypothesis was based on the Euler equation to the problem (4.12). Hall's model generalizes Phelp's model by allowing stochastically evolving labor income $\{y_t\}$, but assumes that R is the (non-stochastic) risk-free rate of return on bonds. In this case the Euler equation takes the form:

$$0 = u'(c_t) - \beta R E\{u'(c_{t+1})|H_t\}. \tag{4.14}$$

Hall observed that (4.14) implies that c_t is a "sufficient statistic" for c_{t+1}. For example, in the special case where $R = 1/\beta$, (4.14) says that $\{u'(c_t)\}$ is a martingale. Hall tested (4.14) by running the regression:

$$c_{t+1} = \alpha_1 + \alpha_2 c_t + \alpha_3' H_t + \xi_t. \tag{4.15}$$

If u is quadratic, then it is not difficult to see that (4.14) implies that the regression in (4.15) will hold with $\alpha_3 = 0$: i.e. no other information in consumers' information set H_t will be useful predictors of c_{t+1} once we know c_t. Using elements c_{t-1}, y_t and y_{t-1} of H_t (where y_t denotes income), Hall was unable to reject the hypothesis that $\alpha_3 = 0$ using U.S. aggregate time series data from 1948 to 1977.

Flavin (1981) re-investigated Hall's findings, estimating a structural model with quadratic utility and an explicit AR representation for the stochastic process of income $\{y_t\}$:

$$y_t = \rho_0 + \rho_1 y_{t-1} + \cdots + \rho_n y_{t-n} + \epsilon_t, \tag{4.16}$$

where $\{\epsilon_t\}$ is a white-noise disturbance process. Flavin showed that disturbance ϵ_t is proportional to the consumers' re-evaluation of their permanent income at time t. If the permanent income hypothesis holds, we have:

$$\begin{aligned} \Delta c_t &\equiv c_t - c_{t-1} = k\epsilon_t \\ k &= \frac{(R-1)}{1 - \rho_1 R^{-1} - \rho_2 R^{-2} - \cdots - \rho_n R^{-n}} \end{aligned} \tag{4.17}$$

Thus, (4.17) says that changes in consumption are driven entirely by "surprises" in labor income which alter consumers' view of their permanent income. Flavin tested the permanent income hypothesis by estimating the income process (4.16) together with the regression:

$$\Delta c_t = k(y_t - \rho_0 - \rho_1 y_{t-1} - \cdots - \rho_n y_{t-n}) + \phi_0 + \phi_1 \Delta y_{t-1} + \cdots \phi_m \Delta y_{t-m} + u_t, \tag{4.18}$$

testing the hypothesis that the coefficients on lagged income changes $(\phi_0, \ldots, \phi_m)$ are zero. Flavin found that the estimated ϕ's were significantly different from zero, leading her to conclude

> "The empirical results indicate that the observed sensitivity of consumption to current income is greater than is warranted by the permanent income-life cycle hypothesis, even when the role of current income in signalling changes in permanent income is taken into account. The restrictions implied by the permanent income-rational expectations hypothesis can be rejected statistically at very high confidence levels. Further, the estimates of the marginal propensity to consume out of current income are quite large, indicating that the failure of the permanent income hypothesis is quantitatively, as well as statistically, significant." (p. 976-977).

Flavin's finding that consumption responds to predictable innovations in income has since become known as the problem of *excess sensitivity,* and was confirmed in in subsequent studies including Hall and Mishkin (1982) using the PSID[29] In his survey of the literature Deaton (1985) concluded: "Surprise consumption functions estimated on quarterly U.S. aggregate time series show clear evidence of excess sensitivity of consumption to predictable events." (p. 146).

However recent research has questioned whether there really is an an excess sensitivity puzzle. Mariger and Shaw's (1988) reinvestigation of the Hall-Mishkin results suggests that the excess sensitivity might be a phenomenon of the early 1970's, since they find little little evidence of excess sensitivity in subsequent waves of the PSID. Zeldes (1989) points out that the excess sensitivity finding may be a rejection of the certainty equivalence principle implied

[29] Panel Study on Income Dynamics. Note that the PSID only has data on food consumption.

by the quadratic utility specification used by Hall, Flavin, and Deaton. Zeldes numerically computed the optimal consumption rule δ^* in a model with CRR preferences (4.10) and concluded:

> "The results indicate that rational individuals with constant relative risk aversion utility will optimally exhibit "excess" sensitivity to transitory income, save "too" much, and have expected growths of consumption that are "too" high, relative to the simple permanent income hypothesis benchmark, even in the absence of borrowing constraints. This suggests that we should rethink our presumption that the certainty equivalent model is the appropriate benchmark, especially at low levels of financial wealth. (p. 295-296).

Attanasio and Browning (1991) argue that the rejections of the life-cycle hypothesis in models estimated using aggregate time series data "may plausibly be attributed to aggregation bias." (p. 2). Using detailed micro data from the U.K. Family Expenditure Survey, they find that "although very simple forms of the consumption function seem to display an excess sensitivity to expected change in income this completely disappears when we allow for the effects of demographics and labor supply." (p. 3-4). Runkle (1991) obtains similar results using the PSID data, concluding: "Aggregate studies assume a representative agent who conditions expectations on *aggregate* variables. But households may not find aggregate data useful in predicting their future economic conditions. In this study I assumed only that each household knew its own past economic condition. With this assumption, I could not reject the permanent income hypothesis." (p. 91).

Another possible reason for the excess sensitivity finding is measurement error. Altonji and Siow (1986) and Runkle (1991) found significant measurement error in income and consumption variables in the PSID. They re-estimated (4.15) using an instrumental variables estimator which is consistent in the presence of measurement error. Contrary to the findings of Hall and Mishkin (1982), they found that no variables in H_t were significant (not even lagged income changes).

A final explanation for the excess sensitivity finding is borrowing constraints. In his survey of this literature, Hayashi (1985) concludes that "The available evidence indicates that for a significant fraction of households in the population consumption is affected in a way predicted by credit rationing and borrowing constraints." (p. 91). Liquidity constraints invalidate the simple Euler equation (4.14) equating current marginal utility to expected discounted marginal utility. Instead under an optimal decision rule $c_t = \delta^*(w_t + y_t)$, marginal utility is the unique solution to the following functional equation:

$$u'(c_t) = max[u'(w_t + y_t), \beta RE\{u'(c_{t+1})|H_t\}]. \tag{4.19}$$

Deaton and Laroque (1991) prove that (4.19) is a contraction mapping which provides a computationally efficient alternative to solving Bellman's equation for the optimal decision rule δ^*. Deaton (1991) computed numerical solutions to (4.19) and showed that when the income process is *IID* or moderately autocorrelated, liquidity constraints create a precautionary demand for saving, with consumers holding a "buffer stock" of assets w_t to smooth out consumption in periods of low income. However in the limit as income tends to a random walk it is optimal to set consumption equal to income. Since U.S. aggregate labor income can be closely approximated by a random walk, Deaton concluded that "a liquidity constrained representative consumer cannot generate aggregate U.S. saving behavior if that agent receives aggregate labor income" (p. 1221). However using a "bottom up" approach which aggregated a heterogeneous collection of constrained and unconstrained individuals, Deaton found that a CDP model is capable of "reconciling the actual (or at least possible) orthogonality condition failures in the micro data (Hall and Mishkin) with those in the macro data (Flavin)" resulting in a "coherent account of a number of disparate phenomena in both the micro and macro data" (p. 1246). Zeldes (1989) tested for the presence of borrowing constraints by splitting his PSID sample into poor and wealthy subsamples. He found that the Euler equation restriction (4.14) could not be rejected for the wealthy subsample, but was decisively rejected for the sample of poor households. Zeldes also found that for the poor households, the Euler equation deviated from 0 in the predicted direction, i.e. the marginal utility of current consumption exceeded the expected discounted marginal utility of future consumption, which is consistent with the hypothesis that these households face binding borrowing constraints. Zeldes's estimates of (4.15) also shows that lagged income is useful for predicting consumption growth for the poor households, which is suggests that the poor, liquidity-constrained households might be responsible for Hall and Mishkin's (1982) finding that "The observed covariation of income and consumption is compatible with the pure life cycle/permanent income hypothesis for 80 per cent of consumption and simple proportionality of consumption and income for the remaining 20 per cent". (p. 461). However the jury is still out on this issue. Runkle (1991) and Keane and Runkle (1991) estimated specifications similar to Zeldes's model using an alternative estimator. Runkle found no evidence of liquidity constraints for rich or poor households and Runkle and Keane found that "lagged income is neither statistically significant nor economically significant in explaining consumption growth. This ... suggests that one possible reason for the difference between Zeldes's and Runkle's results stems from Zeldes's use of a fixed effects estimator in a case that yields inconsistent parameter estimates." (p. 22). Further investigations in this area have been hampered due to the fact that the standard Euler equation approach is inapplicable in the presence of binding borrowing constraints: instead of an Euler equality we have the Euler inequality (4.19) which is not amenable to estimation and testing using existing GMM methods.[30] Development of new estimation methods to handle these problems is high priority since the the existence of liquidity constraints is a generally accepted fact: "Future research should examine the cause, not the existence, of liquidity constraints" (Hayashi 1985 p. 91).

[30] See Hajivassiliou and Ioannides (1991) and Pakes (1991) for possible ways around the problem.

4.4 Empirical Application: Representative Agent Models of Real Business Cycles

If individual data $\{s_t^a, d_t^a\}$ can be treated "as if" it were a realization of a CDP, then under certain conditions we can also treat aggregate time series data for an entire economy $\{s_t, d_t\}$ as if it were a realization of a CDP for a single "representative agent".[31] Kydland and Prescott (1982) used this approach to develope a model of real business cycles. A key fact of business cycles is their persistence: both the levels and the changes in total output and employment are highly serially correlated. To capture this persistence, Kydland and Prescott formulated a CDP model where business cycles are driven by "technology shocks" modelled as an exogenous Markov process. To match the U.S. time-series data, they built two additional sources of persistence into the model: 1) goods are produced in a production pipeline that takes several periods to complete (their so-called "time-to-build" technology), 2) leisure has aspects of a durable good, i.e. leisure in period t yields utility in periods $t + j$, $j \geq 0$. They argued that the latter feature "is crucial in making [the model] consistent with the observation that cyclical employment fluctuates substantially more than productivity does" (p. 1364). Using a calibration approach they chose values of the parameters of (u, p, β) to match certain low order autocorrelations and cross-correlations in U.S. aggregate time series data. They found that the calibrated model "is consistent with the large (percentage) variability in investment and low variability in consumption and the high correlation with output" (p. 1364), and concluded that "the fit of the model is very good, particularly in light of the model's simplicity" (p. 1363).[32]

Kydland and Prescott used the "Hodrick-Prescott" filter to remove indeterministic seasonal and secular components in the data. Singleton (1988) criticized this procedure on the grounds that "the Kydland-Prescott filter leads to large differences between the time series properties of the (prefiltered) seasonally adjusted and unadjusted series. The former, which were studied by Kydland and Prescott, exhibit much weaker dynamic interrelations, less volatility, and different autocorrelations than the unadjusted series. Hence, the 'stylized facts' motivating recent specifications of business cycle models may have been distorted by the prefiltering procedures." (p. 372). A subsequent study by Cogley and Nason (1991) found that "the HP filter can generate spurious business cycle dynamics when applied to integrated time series" (p. 20), and that "Before smoothing, the models often fail to match various sample auto- and cross correlations, but the smoothing operation often improves the fit. After smoothing, artificial dynamics resemble actual dynamics because both display the properties of a smoothing filter." (p. 3).

---

[31] For example, Lucas and Prescott (1971) showed that this will be the case if allocations in a dynamic economy are Pareto efficient since the second Welfare theorem implies that allocations are given by the solution to a certain "social planning problem". The social planning problem maximizes a weighted average of utilities of individuals in the economy, so this weighted average utility function can be interpreted as the preferences of the "representative consumer".

[32] Long and Plosser (1983) also used a calibration approach to study a CDP model of real business cycles and concluded "The time-series properties of [the model] exhibit some major features of observed business cycles. Although this type of model may not be capable of of explaining all the regularities in actual business cycles, we believe that it provides a useful, well-defined benchmark for assessing the relative importance of factors (e.g. monetary disturbances) that we have deliberately ignored." (p. 39).

Kydland and Prescott's "calibration" procedure has also been widely criticized:

> "The use of parameters from microeconomic studies in particular was advocated by Kydland and Prescott (1982) and Prescott (1986) as part of their "calibration" approach to parameter selection. A potentially major limitation of adopting the values of parameters that were estimated in previous studies is that the models considered may have been very different than the model currently under examination. For instance, micro labor market models often allow for nontrivial forms of heterogeneity across consumers and, in some models, consumers are allowed to be at corners with regard to their labor supply decisions. In contrast, aggregate models of labor supply assume interior labor supply decisions and allow only limited forms of heterogeneity, while often accomodating much richer forms of dynamic specifications than in the microeconomic studies. Accordingly the parameters in micro and aggregate analyses are typically identified econometrically using very different sets of population moment conditions that reflect the differences in dynamics or market structures." (Singleton, 1988, p. 380).

Altûg (1989) used maximum likelihood to estimate all the parameters of an extended version of the Kydland-Prescott in the internally consistent manner advocated by Singleton. She found that "an aggregative model in which persistent technology shocks are the only driving force cannot rationalize the *joint* behavior of per capita hours with the remaining series" (p. 912). Eichenbaum (1991) re-evaluated subsequent claims by Kydland and Prescott (1989) that technology shocks account for 70% of the variability in post-World War II output. He concludes that "What the data are actually telling us is that, while technology shocks almost certainly play some role in generating the business cycle, there is simply an enormous amount of uncertainty about just what percent of aggregate fluctuations they actually do account for." (p. 608). Eichenbaum found that the empirical implications of RBC models similar to Kydland and Prescott's model are *fragile:* "(1) Small perturbations to the theory alter the conclusion in a basic way, (2) Small changes in the statistical methods used alter the conclusion in a basic way, (3) Changes in the sample period alter the conclusion in a basic way, (4) And most importantly, our confidence in the conclusion is fundamentally affected once we abandon the convenient fiction that we actually *know* the true values of the structural parameters of standard RBC models." (p. 608).

4.5 Empirical Application: Models of Production, Sales, and Inventory Holdings

The linear quadratic inventory model of Holt et. al. (1960) stimulated substantial theoretical and empirical work on the question of whether firms facing randomly fluctuating sales hold optimal inventories of finished goods to serve as a "buffer stock" and smooth out production. Blanchard (1983) studied inventory holdings of the U.S. automobile industry. He estimated the structural parameters of a linear-quadratic model of firm production/inventory decisions by maximum likelihood, using methods developed by Hansen and Sargent (1980a,b). The structural model implies a set of nonlinear cross-equation restrictions on the coefficients of an autoregression for inventories and sales. Blanchard tested these restrictions and found that "Because of the large number of observations, the χ^2 statistic is large. The model is rejected at the .005 level for five divisions and rejected at the .10 level for eight divisions. The model fits better the divisions of GM than those of Ford or Chrysler. It fits Buick and Cadillac particularly well." (p.

390). Blanchard's overall conclusion is that "inventory behavior is reasonably well explained by the assumption of intertemporal optimization with rational expectations" (p. 365).

West (1986) also estimated a linear-quadratic inventory model using monthly aggregate data on output, sales and inventories of five 2-digit U.S. industries. West used GMM to estimate the structural parameters θ of the following version of the firm's optimization problem:

$$
\begin{aligned}
& max\, E\left\{\sum_{t=0}^{\infty}\beta^t\left[p_tS_t-\theta_1^t[\theta_2(\Delta Q_t)^2+\theta_3Q_t^2+\theta_4(I_t-\theta_5S_{t+1})^2]\right]\right\} \\
& \text{s.t. } Q_t = S_t + I_t - I_{t-1},
\end{aligned}
\tag{4.20}
$$

where p_t denotes price of output, S_t denotes sales, Q_t denotes production, and I_t denotes inventory held at time t.[33] West's specification allows for quadratic costs of both the level and change in production, the latter reflecting costs of adjustment such as hiring and firing costs. The final term represents inventory and backlog costs, a quadratic function of how far inventories are from the target level θ_5S_{t+1}. West found that the GMM estimates of θ were economically plausible, and that for most of the industries, low values of Chi-square statistics did not indicate a rejection of the model's overidentifying restrictions. On the basis of these results we might expect that the estimated optimal decision rule $\delta^*_{\hat\theta}$ ought to provide a good approximation to observed behavior. To test this directly, West computed the expected discounted profits under the estimated decision rule $\delta^*_{\hat\theta}$ and compared this to expected discounted profits under a naive alternative policy δ_a that initially sets inventories to the values observed in the data, increasing this base level from period to period at their historical trend rate of growth. He showed that the difference in expected discounted costs can be rewritten as a simple linear combination of the variances and covariances of production, sales and inventories, and that this difference must be non-negative if the linear-quadratic specification is correct. In fact West found that his "variance bound" statistic was usually significantly negative, indicating the that the naive alternative policy produced significantly *higher* expected discounted profits than the supposedly "optimal" inventory policy $\delta^*_{\hat\theta}$. Indeed in some cases, the naive alternative policy reduced the firm's expected discounted costs of production by as much as 50% relative to the costs under the supposedly optimal policy $\delta^*_{\hat\theta}$. These findings lead West to conclude:

> "It would seem that the linear quadratic inventory model does a poor job of rationalizing these inventory data. In effect, a contradiction results when it is assumed that the actual inventory path chosen is the one that is optimal according to the model. The allegedly optimal path is dominated by a naive alternative path. In the model without a target level for inventories ($\theta_5 = 0$), this follows simply because production is more variable than sales. Inventories therefore cannot be chosen simply to perform their putative function, smoothing production. For the model with a target level, the matter is slightly more complicated. Invetories do usually track their target level (except in the petroleum industry). But this makes production and inventories so variable that inventories cannot be chosen as hypothesized to minimize quadratic inventory, production, and target-level costs. The basic implication of this is that inventories appear to serve some role other than production smoothing." (p. 397-398).

[33] All variables are measured as deviations from trend.

5. Discrete Decision Processes: Implications and Evidence

In DDP's the decision variable can assume a countable number of values: typically the agent faces a finite number of alternatives. In that case the optimal decision rule δ^* is defined by a finite number of inequality conditions:

$$d \in \delta^*(s) \iff \left\{\forall d' \in D(s) \Big| \, u(s,d) + \beta \int V^*(s')p(ds'|s,d) \geq u(s,d') + \beta \int V^*(s')p(ds'|s,d')\right\}, \tag{5.1}$$

so it is clear that we cannot differentiate with respect to d to derive the Euler equations that have proved so useful in the case of CDP's. Instead most estimation and testing of DDP's can be viewed as a type of nonlinear regression problem, in which (u, p, β) is specified up to an unknown parameter vector θ, and a "nested numerical solution algorithm" repeatedly re-solves the DDP problem until it finds an estimate $\hat{\theta}$ such that the implied decision rule $d_t = \delta^*(s_t, \hat{\theta})$ "best fits" data $\{s_t^a, d_t^a\}$ on observed states and decisions of agents $a = 1, \ldots, A$.

If we were able to fully observe agents' states and decisions, and if the primitives (u, p, β) are correctly specified, Blackwell's Theorem has a very strong implication: the estimated decision rule should be able to perfectly predict agents' behavior. However in practice no data set is rich enough to fully measure all aspects of an agent's state. Suppose that s consists of two components $s = (x, \epsilon)$ where x is a state variable observed by the agent and the econometrician and ϵ represents *private information* observed only the the agent. Then even though the decision rule $d = \delta^*(x, \epsilon)$ is non-stochastic from the agent's point of view, the econometrician will not be able to perfectly predict the agent's decision without observing ϵ. Since ϵ generally enters δ^* in a non-additive fashion, it is infeasible to estimate θ by nonlinear least squares. The "method of choice" for estimating θ is maximum likelihood using the *conditional choice probability*

$$P(d|x,\theta) = \int I\{d = \delta^*(x,\epsilon,\theta)\} q(d\epsilon|x), \tag{5.2}$$

where $q(d\epsilon|x)$ is the conditional distribution of ϵ given x (to be defined). We can then evaluate the DDP model using standard statistical procedures. An example is the Chi-square diagnostic test, which compares how well the estimated parametric choice probability $P(d|x,\theta)$ matches up with a non-parametric estimate of P based on sample frequency distribution of different choices d for each cell x. We now define a class of MDP's for which conditional choice probabilities of the form (5.2) are well-defined:

Definition 5.1: A *Discrete Decision Process* (DDP) is a Markov Decision Process 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 $X \times E$, 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:

$$u(s,d) = u(x,d) + \epsilon(d) \tag{AS}$$

where $\epsilon(d)$ is the d^{th} component of the vector $\epsilon \in R^{|D|}$.

- The transition probability $p(ds_{t+1}|s_t,d_t)$ satisfies assumption CI:

$$p(dx_{t+1},d\epsilon_{t+1}|x_t,\epsilon_t,d_t) = q(d\epsilon_{t+1}|x_{t+1})\pi(dx_{t+1}|x_t,d_t). \tag{5.3}$$

where $q(d\epsilon|x)$ is a conditional probability distribution on $R^{|D|}$.[34]

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

$$G(\{u(x,d), d \in D(x)\}|x) = \int_{R^{|D|}} \max_{d \in D(x)} [u(x,d) + \epsilon(d)]q(d\epsilon|x). \tag{5.4}$$

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: $P(d|x)$ equals the partial derivative of G with respect to $u(x,d)$. Using this property we can establish the following theorems characterizing optimal decision rules and their implied choice probabilities in a DDP:[35]

Theorem 5.1: *If $\{s_t,d_t\}$ is a DDP satisfying certain regularity conditions given in Rust (1992), then the optimal decision rule δ^* is given by:*

$$\delta^*(x,\epsilon) = \underset{d \in D(x)}{argmax}[v(x,d) + \epsilon(d)], \tag{5.5}$$

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

$$\Psi(v)(x,d) = u(x,d) + \beta \int G(\{v(y,d'), d' \in D(y)\}|y)\pi(dy|x,d). \tag{5.6}$$

[34] For technical reasons detailed in Rust (1992) I define q as 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 $|D| - |D(x)|$.

[35] Although Theorems 5.1 and 5.2 appear to apply only to infinite-horizon stationary DDP problems, they actually include finite-horizon, nonstationary 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 5.1 and 5.2 continue to hold, with the exception that δ^*, P, G, π, and v all depend on t.

Theorem 5.2: *If $\{s_t, d_t\}$ is a DDP, then the observed components $\{x_t, \epsilon_t\}$ follow a Markov process with transition probability:*

$$Pr\{dx_{t+1}, d_{t+1} \big| x_t, d_t\} = P(d_{t+1} \big| x_{t+1})\pi(dx_{t+1} \big| x_t, d_t), \tag{5.7}$$

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

$$P(d \big| x) = \frac{\partial G(\{v(x,d), d \in D(x)\} \big| x)}{\partial v(x,d)} \tag{5.8}$$

and G is the Social Surplus function defined in (5.4) and v is the unique fixed point to the contraction mapping Ψ defined in (5.6).

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 in Theorem's 5.1 and 5.2. For example if $q(d\epsilon|x)$ is a multivariate extreme value distribution we have:

$$P(d \big| x) = \frac{exp\{v(x,d)\}}{\sum_{d' \in D(x)} exp\{v(x,d')\}}, \tag{5.10}$$

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

$$\Psi(v)(x,d) = u(x,d) + \beta \int log[\sum_{d' \in D(y)} exp\{v(y,d')\}]\pi(dy \big| x,d). \tag{5.11}$$

The extreme-value specification is especially attractive for empirical applications since the closed-form expressions for P and G avoid the need for multidimensional numerical integrations required for other distributions. At the same time, the "dynamic logit model" (5.10) avoids the IIA restriction of static logit models that the odds ratio of choosing any two alternatives is independent of the attributes of other alternatives.[36]

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:

$$\hat{\theta}^f = \underset{\theta \in R^N}{argmax}\, L^f(\theta) \equiv \prod_{a=1}^{A} \prod_{t=1}^{T_i} P(d_t^a \big| x_t^a, \theta)\pi(x_t^a \big| x_{t-1}^a, d_{t-1}^a, \theta). \tag{5.12}$$

or a partial likelihood estimator $\hat{\theta}^p$ defined by:

$$\hat{\theta}^p = \underset{\theta \in R^N}{argmax}\, L^p(\theta) \equiv \prod_{a=1}^{A} \prod_{t=1}^{T_i} P(d_t^a \big| x_t^a, \theta). \tag{5.13}$$

[36] For details see Rust, (1992).

The maximum likelihood estimator has standard statistical properties, allowing us to use a variety of specification test statistics to assess the DDP's model's ability to fit the data. These tests include the information matrix test (White, 1982, Lancaster, 1986), and the Chi-square goodness of fit test (Andrews, 1988).

While the definition of DDP models given in definition 5.1 covers a fairly general class of problems, it does not cover all specifications of dynamic discrete choice models that have appeared in the literature. However all of these models can be mapped into specifications of the form of definition of 5.1, so I refer to all of them as DDP models. The following examples will illustrate some of the alternative specifications as well as some alternative estimation methods.

5.1 Empirical Application: Maintenance, Replacement, and Scrapping of Physical Capital

One of the simplest applications of the specific class of DDP models given in definition 5.1 is Rust's (1987) model of bus engine replacement. In contrast to the macro-economic studies that investigate aggregate replacement investment (e.g. Jorgensen, 1973), Rust's model goes to the other extreme and examines 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:

$$u(x_t, d_t, \theta_1, \theta_2) = \begin{cases} -\theta_1 - c(0, \theta_2) - \epsilon_t(1) & \text{if } d_t = 1 \\ -c(x_t, \theta_2) - \epsilon_t(0) & \text{if } d_t = 0, \end{cases} \tag{5.14}$$

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 (5.14) 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:

$$\pi(x_{t+1}|x_t, d_t) = \begin{cases} g(x_{t+1} - 0) & \text{if } d_t = 1 \\ g(x_{t+1} - x_t) & \text{if } d_t = 0 \end{cases} \tag{5.15}$$

where g is a probability density function. The renewal property given in equations (5.14) and (5.15) defines a *regenerative optimal stopping problem,* and under an optimal decision rule $d_t = \delta^*(x_t, \epsilon_t, \theta)$ 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 (5.12). Figures 5.1 and 5.2 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 ($\beta = .9999$) 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 $\beta = .9999$ 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 $\beta \geq .98$, although with a very slight upward slope as $\beta \to 1$. The latter finding, together with the *Final Value Theorem* (Bhattacharya and Majumdar, 1990) may indicate that Zurcher is minimizing long-run average costs rather than discounted costs:

$$\lim_{\beta \to 1} \max_{\delta} E_{\delta} \left\{ (1-\beta) \sum_{t=0}^{\infty} \beta^t u(x_t, d_t) \right\} = \lim_{T \to \infty} \max_{\delta} \frac{1}{T} E_{\delta} \left\{ \sum_{t=1}^{T} u(x_t, d_t) \right\}. \tag{5.16}$$

The estimated model implies that expected monthly maintenance costs increase by \$1.87 for every additional 5,000 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 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.[37] In this case conversations with Zurcher revealed that implied cost estimates from the structural model corresponded closely to Zurcher's 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.

Kennet (1992) estimated a similar model of an airline's decision of when to remove and overhaul an aircraft engine. In this case x_t represents the accumulated number of operating hours since last replacement. Using data on maintenance records for 42 Pratt and Whitney jet engines Kennet found that the structural model "predicts the observed outcome reasonably well for the first 10 states or so (8,000 engine hours), but not in higher values" (p. 14), a result that may partly be due to the fact that there are relatively few observations above 8,000 hours, which is more than twice as large as the mean time between overhauls. Kennet found that when he split the sample into pre and post deregulation periods, there was a significant change in the structural parameters in the two periods: airlines seemed to perceive that the cost of overhauling an aircraft engine was significantly higher after deregulation, but the estimated monthly cost of maintaining an engine fell. The total effect implied a significant *lengthening* in the amount of time between engine replacements after deregulation, an effect that is consistent with a variety of reduced-form estimates

[37] 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 a discussion of this problem in the context of models for discrete panel data.

of the duration between overhauls. Kennet was also unable to reject the hypothesis that the discount factor $\beta = 0$ in the pre-deregulation era, but was able to reject $\beta = 0$ in favor of $\beta = .9999$ in the post-deregulation era. Kennet's interpretation of these findings is that "maintenance prior to deregulation was super-optimal in that firms did not seem to minimize discounted costs. Deregulation, in this scenario, forced firms to become more cost efficient, driving maintenance effort to a more efficient level. The data also seem to indicate that the perceived cost to firms of engine shutdown is lower since deregulation. For at least one of the engine types, the results are consistent with claims by workers of corporate laxity toward minor airworthiness violations." (p. 18).

Sturm (1991) estimated a DDP model of the decision of when to refuel a nuclear reactor. Refuelings involve a substantial opportunity cost in terms of lost power generation. On the other hand, they provide an opportunity to perform preventive maintenance which helps reduce the frequency of unplanned equipment failures and reactor "scrams" (emergency shutdowns) that can have very costly indirect consequences in terms of lost public and regulatory goodwill. Similar to the Rust and Kennet model, Sturm's model assumes that maintenance done during a refueling regenerates the state of the reactor, so that the rate of unplanned failures is an increasing function of the time since last refueling. The reactor operator's decision problem involves a trade off between the opportunity cost of more frequent planned refueling shutdowns versus the goodwill costs of the increased likelihood of emergency shutdowns as the length of the refueling cycle increases. Sturm used publicly available data on nuclear power plant operations in 5 Eurpoean countries, estimating country-sepcific cost functions and hazard rates for unplanned shutdowns. He found that the estimated model predicts the cross-country variations in the operating cycle length and the rate of unplanned shutdowns that we see in the data: Germany and Switzerland have the shortest mean time between refuelings and the lowest rate of unplanned outages, whereas Belgium, France and Sweden have the longest duration between refuelings and the highest rate of unplanned outages. He concluded that "the model is capable of distinguishing the effects of economic incentives and technical ('X-efficiency') factors on various aspects of nuclear power plant performance. In particular, the results indicate that France and Belgium's relatively good showing in an international comparison with respect to one performance measure, namely availability, is to some extent due to the incentive structure that the operators in these two countries face." (p. 28).

Das (1991) estimated a DDP model of a decision of a cement plant owner whether or not to operate, idle, or retire a cement kiln. Using a dataset on the histories of 987 cement kilns over the period 1972 to 1980, Das estimated a DDP model where the state variable x_t includes the age of the kiln together with three price variables: the hourly wage rate, the price of fuel, and the price of electricity. These additional state variables are necessary to enable the model to track the large increase in retirements and idling of kilns during the energy crisis of the late 1970's. Das found that the estimated model was able to predict firms' responses to the energy crisis, and though it "tends to underpredict utilization and overpredict idleness" she found that "except during 1975-1978, the predicted trends are close to those observed in the data" and "the model tracks retirement quite well until 1977 but fails to explain the 'blips' in retirement in 1978 and 1980." (p. 24). A Chi-squared goodness of fit test of the overall fit of the model was unable to reject the structural model at the 0.5% level of significance.

Rosenzweig and Wolpin (1990) estimated a DDP model of an Indian farmer's decision of how many bullocks to buy or sell, whether or not to breed a bullock, whether or not to buy an irrigation pump, and how much planting inputs to purchase. Their model reflects the existence of credit market imperfections which cause bullocks to serve a dual role as both a productive asset and a financial asset to help smooth consumption in periods of low agricultural output. Their reduced-form estimates show that farmers are significantly more likely to purchase bullocks when income is high than when income is low, consistent with a consumption-smoothing motive. Rosenzweig and Wolpin estimated their model using a CRR utility function with a minimum or "subsistence" level of consumption, and a one-dimensional specification of the unobserved state variable ϵ_t rather than the multi-dimensional specification given in definition 5.1. The unknown parameters were estimated by maximum likelihood using a partial likelihood function similar to (5.13), using a nested numerical solution procedure similar to the one originally developed by Wolpin (1984). The parameter estimates of their structural model indicate that there are very high returns to investing in an irrigation pump regardless of the number of bullocks owned, and that the profit-maximizing number of bullocks is 2. Their estimates allow them to quantify the effect of a spell of bad weather on profits (reducing them by an average of 753 rupees), the cost of a pump (6340 rupees), and the cost of breeding (857 rupees). Their estimate of the parameter of the CRR utility function is $\gamma = .04$, reflecting significant risk aversion, and the estimated minimum consumption level is 1470 rupees, wich is 56% of mean household food consumption. The estimated model predicts ownership of bullocks by small farmers quite well: their Chi-square goodness of fit tests lead them to conclude that "We cannot reject, for any year, the hypothesis that the distribution of bullock stocks predicted by our model is identical with the actual distribution." (p. 19). However out-of-sample predictive tests of bullock holdings by large farmers reveal that "the model performs quite poorly in predicting bullock stocks for large farmers" suggesting that "The poor performance of our estimated model in predicting the behavior of the large-size farmers based on medium-size farmer data is consistent with such farmers having superior abilities to obtain consumption loans or to accumulate alternate assets to be used for consumption-smoothing although it is notable that even the large farmers are still not fully "efficient" in the average holdings of bullocks." (p. 19).

5.2 Empirical Application: Job Search and Career Decisions

In the simplest job search model (Lippman and McCall, 1976) in each period t an unemployed worker receives a wage offer w_t, assumed to be an *IID* draw from a distribution $F(w)$.[38] If a worker accepts an offer he remains in the

[38] If the worker receives no wage offer in period t we set $w_t = 0$. If the worker recieves more than one offer, we let w_t be the highest wage offer.

job forever, earning discounted wages of $w/(1-\beta)$. If the worker is unemployed he receives unemployment benefits u and incurs search costs c. The value function for this problem is given by:

$$V(w) = max[w/(1-\beta), u - c + \beta \int V(y)F(dy)]. \tag{5.17}$$

It is straightforward to show that the optimal search strategy is given by the following threshold rule:

$$\delta^*(w) = \begin{cases} 1 & \text{if } w \geq w_r \\ 0 & \text{if } w < w_r, \end{cases} \tag{5.18}$$

where the *reservation price* w_r is the unique solution to:

$$w_r = \frac{(1-\beta)(u-c) + \beta \int_{w_r}^{\infty} yF(dy)}{1 - \beta F(w_r)}. \tag{5.19}$$

Note that without some sort of error term, the decision rule (5.18) implies that w_r must be less than any accepted wage in the sample: direct maximum likelihood estimation of such a model subject to the constraint that $w_t \geq w_r$ leads to non-standard asymptotic properties since the support of the likelihood generally depends on θ.[39] In addition, this constraint typically yields implausibly low estimated reservation values which imply that unemployed workers would accept virtually any job offer. This is inconsistent with empirical findings such as Blau and Robins (1990) who show that in their EOPP data set[40] unemployed workers reject about one third of all job offers.

One way to deal with this problem is to introduce unobserved state variables $\{\epsilon_t\}$ satisfying the AS-CI assumption of definition 5.1. However empirical studies by Wolpin (1987) and Christensen and Kiefer (1991) treat ϵ_t as measurement error rather than as an unobserved state variable. In Wolpin's model measured wages w_t equal true wages w_t^* plus an *IID* measurement error ϵ_t:

$$w_t = w_t^* + \epsilon_t. \tag{5.20}$$

Wolpin estimated a finite horizon search model where the distribution F_t of true wage offers is $N(\mu_t, \omega_t^2)$ and measurement errors $\{\epsilon_t\}$ are serially independent and independent of $\{w_t^*\}$ with distribution $N(0, \sigma^2)$.[41] Note that the standard search model without measurement errors is a special case when $\sigma^2 = 0$. Christensen and Kiefer's (1991) infinite-horizon search model assumes that measurement errors enter multiplicatively rather than additively:

$$w_t = w_t^* \epsilon_t, \tag{5.21}$$

[39] See Christensen and Kiefer (1991).

[40] Employment Opportunities Pilot Project.

[41] In the alternative log-normal model, we let w_t and w_t^* be the log of measured and true wages, respectively.

where the distribution of true wage offers is shifted exponential:

$$F(dw_t^*) = \gamma_t \, exp\{-\gamma_t(w_t^* - c_t)\}, \quad w_t^* \geq c_t, \tag{5.22}$$

and $\{\epsilon_t\}$ is assumed to be an *IID* process which is independent of $\{w_t^*\}$ with density $g(\epsilon)$ given by:

$$g(\epsilon) = a\epsilon^{-(b+2)} \, exp\{-b/\epsilon\}, \quad \epsilon > 0. \tag{5.23}$$

Note that as $b \to \infty$, the distribution of measurement errors converges weakly to a unit mass at 1, so their model also nests the job search model without measurement error as a limiting case. Both studies estimated the unknown parameters of the search model and the distribution of measurement errors by maximum likelihood. Their empirical findings show that the standard search model does a good job of fitting certain aspects of the data, such as mean accepted wages and the "hazard" or departure rate from unemployment.[42] Wolpin's study used NLS data[43] which followed 144 male high school students from graduation to employment. He found that "The general decline in the observed hazard rate is picked up by the model, although the peculiarly large hazard rates around one year and two years of leaving high school is not picked up by the model. It is possible that some of this clustering in reported job taking at these durations is not real, but either a result of the small sample or of recall error. (p. 812). Wolpin also found that "for the entire sample, the predicted accepted mean wage is 193 dollars, which is quite close to the observed figure of 205 dollars" (p. 813). A χ^2 specification test of Wolpin's model indicates that "the restrictions of the search model are not rejected at conventional significance levels" (p. 813). Christensen and Kiefer used SIPP data,[44] selecting a subsample of 1,031 unemployed males and females. They found that "the model can be fit and interpreted straightforwardly and that the results are sensible. Reservation wages are estimated and turn out to be *a priori* plausible even in samples in which the minimum observed wages are clearly implausible. In other samples, the prototypal model without the addition of measurement error explains the data satisfactorily." (p. 27).

There are two aspects of these models that don't seem to fit the data well: both models imply job offer rates and reservation values which are too low. For example Wolpin estimates weekly job offer probabilities from 1.28% in the first week after graduation to .91% after 54 weeks. In Christensen and Kiefer's model the estimated offer probability depends on the value assumed for β, ranging 1% to 3.4% per week as β ranges from .5 to .7 (on an annual basis). These offer probabilities appear low in comparison to rates of 18% reported by Blau and Robins.[45] Estimated

---

[42] Specifically, this is the conditional probability that an unemployed worker accepts a job offer as a function of duration of unemployment.

[43] National Longitudinal Survey.

[44] Survey of Income and Program Participation.

[45] It should be noted that offer rates in their EOPP are considerably higher than other data. For example Holzer (1987) reports a *monthly* offer rate in the NLS data that is about equal to the *weekly* offer rate in the EOPP data. However the structural estimates of offer probabilities appear low even when we take this into consideration.

reservation wages in Wolpin's model "are quite low and decline continuously. Acceptance probabilities given an offer are conversely very high and continuously increasing. At the time of initial search the acceptance probability ranges from a high of .97 for group one to a low of .88 for group four." (p. 814). Estimated reservation wages are also low in Christensen and Kiefer's model: for example for the group of young white males the estimated reservation wage of \$3.13/hour is below the legal minimum wage of \$3.35/hour. Thus both models imply acceptance rates that are much higher than what we see in the data, which is probably related to the fact that neither Wolpin nor Christensen and Kiefer found significant measurement error in wages: as a result both of their models predict very low reservation wages that are characteristic of models that don't include measurement error.[46]

The latter discrepancies may indicate that the standard job search model is oversimplified: there are other aspects to jobs than just wages. One important aspect is the worker's suitability for the job in question. Another strand of the job search literature, beginning with the work of Jovanovic (1979), has attempted to model the process of "matching" workers with jobs. The result is a theory that can explain why a worker would choose to quit an existing job and look for a new one. A key element of these models is *learning:* a worker is initially unsure of his suitability (productivity) on a given job, but learns over time based on observation of his output. Miller (1984) was the first to estimate and test such a model, formulating the worker's decision problem as a *multi-armed bandit.* These are DDP problems with a special structure that can be best envisioned as the decision as to which of several "one-armed bandit" machines one should play in a gambling casino. Although you do not know the payoff probabilities of the different machines, you can learn about their payoffs via Bayesian updating of your prior distribution over payoff probabilities. The key restriction is that the environment is stationary and the bandits are independent, so you can't learn any new information about any of the other machines while you are playing the current machine. Gittens and Jones (1974) have proved that the optimal decision rule for the bandit problem takes the form of a reservation price strategy: always play the bandit which has the highest reservation price. The reservation prices are computed separately for each bandit. The stationarity assumption implies that the reservation price of the machine you are currently playing is updated as you observe its payoffs, but the reservation prices of the other machines remain constant. An implication of the optimal decision rule is that certain "high risk" bandits should be sampled first since these machines have a high information content in the sense that a one can learn relatively rapidly whether they yield high payoffs or low payoffs.

The translation to the job matching problem is clear: the bandits correspond to different jobs and the payoff is the worker's wage at time t. The worker is uncertain about his true productivity in any job, and adopts a search strategy to maximize his expected discounted earnings. Miller devised a maximum likelihood estimator to estimate

---

[46] Wolpin found that "Little of the total wage variance (1.2 per cent in specification one and 9.2 per cent in specification two) is attributable to measurement error." (p. 812). Christensen and Kiefer found that in three of eight subgroups studied, their estimation algorithm drove $b \to \infty$ indicating no measurement error. In the remaining subgroups the proportion of log wages that is due to measurement error ranges from .12 to .28, except for very high estimated measurement error for older white males and young white females (.75 and .95, respectively).

the unknown structural parameters of the job matching model. Although Miller did not conduct a formal specification test of his model, he concluded that "The empirical findings bolster the theoretical portion of this paper. It argues that since the value of specific experience varies across job types, optimizing behavior induces a stochastic career profile: jobs yielding returns that are subject to a particular form of uncertainty are experimented with first." (p. 1107). Flinn (1986) and McCall (1990) develope more formal tests of the job-matching model. Flinn estimated the structural parameters of a discrete time version of Jovanovic's (1979) model. Flinn showed that the structural model lead to certain restriction on the intertemporal covariance structure of worker's wages. He tested these restrictions and found that the "restrictions on the covariance structure of wages implied by the matching model, as specified here, are not rejected." (p. S104). McCall's model generalizes the multi-armed bandit model by allowing a job to yield both job-specific and occupation-specific information. If occupational matching is significant, then the theory predicts that if a worker chooses a new job in the same occupation, increased tenure in the previous job lowers the likelihood of separation from the new job. Using NLS data, McCall concluded that in general the data are consistent with this implication of the job matching model.

5.3 Empirical Application: Patent Renewal

Pakes (1986) developed a DDP model of a European patent holder's decision of whether or not to pay a renewal fee to keep their patent in force. Pakes's study is innovative in its incorporation of serially correlated unobserved state variables $\{\epsilon_t\}$, and its use of Monte Carlo simulation of the $\{\epsilon_t\}$ 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. To keep a patent in force in Europe, one must pay an an annual renewal fee c_t, which is 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 that the only observable state variable is $x_t = t$, the number of years the patent is kept in force. Assuming that $\{\epsilon_t\}$ is a first order Markov process with transition probability $q_t(d\epsilon_t|\epsilon_{t-1})$, Bellman's equation is given by:

$$\begin{aligned} V_T(\epsilon) &= max\left\{0, \epsilon - c_T\right\}, \\ V_t(\epsilon) &= max\left\{0, \epsilon - c_t + \beta \int V_{t+1}(\epsilon')q_{t+1}(d\epsilon'|\epsilon)\right\}. \end{aligned} \tag{5.24}$$

where T is the statutory limit on patent life. Under fairly general conditions on the transition probabilities $\{q_t\}$ one can establish that the optimal decision rule $d = \delta_t^*(\epsilon)$ is given by a threshold rule similar to (5.18),

$$\delta^*(\epsilon_t) = \begin{cases} 1 & \text{if } \epsilon_t > \eta_t(\theta) \\ 0 & \text{if } \epsilon_t \leq \eta_t(\theta) \end{cases} \tag{5.25}$$

Pakes used a particular family $\{q_t\}$ to simplify the recursive calculation of the cutoffs $\eta_t(\theta)$:

$$q_{t+1}(d\epsilon_{t+1}|\epsilon_t) = \begin{cases} exp\{-\theta_1\epsilon_t\} & \text{if } \epsilon_{t+1} = 0 \\ 0 & \text{if } 0 \leq \epsilon_{t+1} \leq \theta_2\epsilon_t \\ \frac{1}{(\theta_4^{t-1}\theta_5)} exp\left\{\frac{-(\theta_3+\epsilon_{t+1})}{(\theta_4^{t-1}\theta_5)}\right\} & \text{if } \epsilon_{t+1} > \theta_2\epsilon_t. \end{cases} \tag{5.26}$$

The interpretation of (5.26) is that with probability $exp\{\theta_1\epsilon_t\}$ the patent is discovered to be worthless (in which case $\epsilon_{t+k} = 0$ for all $k \geq 1$), otherwise next period returns given by $\epsilon_{t+1} = max\{\theta_2\epsilon_t, z_{t+1}\}$ where $\{z_t\}$ is an *IID* exponential random variable whose density is given by the third term in (5.26).

The likelihood function for this problem requires computation of the probability $\lambda_t(\theta)$ that the patent lapses in year t:

$$\begin{aligned} \lambda_\theta(t) &= \Pr\{\delta_t^*(\epsilon_t) = 0, \delta_{t-1}^*(\epsilon_{t-1}) = 1, \ldots, \delta_1^*(\epsilon_1) = 1\} \\ &= \int_{\epsilon_t} \cdots \int_{\epsilon_0} \left[I\{\epsilon_t < \eta_t(\theta)\} \prod_{s=1}^{t-1} I\{\epsilon_s \geq \eta_s(\theta)\}\right] \prod_{s=1}^{t} q_t(d\epsilon_s|\epsilon_{s-1},\theta) q_0(d\epsilon_0|\theta), \end{aligned} \tag{5.27}$$

where q_0 is the initial distribution of returns at the time the patent was applied for, assumed to be lognormal with parameters (θ_6, θ_7). Computation of the likelihood requires one to "integrate out" the unobserved realizations $\{\epsilon_t\}$. However in practice, the numerical integrations required in (5.27) 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 $\{\tilde{\epsilon}_t\}$ and tabulating the fraction of realizations that lead to drop-outs at year t, i.e. $\tilde{\epsilon}_t < \eta_t(\theta), \tilde{\epsilon}_s \geq \eta_s(\theta), s = 1, \ldots, t-1$.[47]

Pakes found that the estimated structural model fit the data remarkably well. Figure 5.3 plots the actual versus predicted patent drop out rates for Germany and France. The figure compares two different structural models, a deterministic model where patent returns decay according to the law of motion $\epsilon_{t+1} = \theta_2\epsilon_t$ with probability one, and the stochastic model described above. From the figure one can see that the stochastic model does a better job of fitting the data. Pakes' structural estimates allowed him to tabulate the (unobserved) distribution of patent valuations. In the case of France, he found that "Twenty-five per cent of the patents in the French data had realized values of seventy-five dollars or less" (p. 777), but the distribution is highly skewed: "The median of the distribution of realized values ($534) was less than one tenth its mean ($5,631); and the five per cent of the distribution with the highest realized values contribute about half of the total value of a cohort." (p. 777). These findings lead Pakes to conclude that patents are renewed primarily for their "option value":

[47] McFadden (1989) and Pakes and Pollard (1989) subsequently developed a new class simulation estimators that dramatically lowers the number of suimulations (NSIM) needed to obtain consistent and asymptotically normal parameters estimates. In fact, only a finite value of NSIM is required to compute $\hat{\lambda}(\theta)$, where NSIM can be as small as 1.

"Patents are applied for at an early stage in the inventive process, a stage in which there is still substantial uncertainty concerning both the returns that will be earned from holding the patents and the returns that will accrue to the patented ideas. Gradually the patentors uncover the true value of their patents. Most turn out to be of little value, but the rare "winner" justifies investments that were made in developing them." (p. 780).

5.4 Empirical Application: Fertility Decisions

Wolpin (1984) developed a DDP model of 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. There is no uncertainty about fertility in this model, since conception and contraception are assumed to be 100% effective during the fertile period, which ends at age 45 in Wolpin's model. There is uncertainty about child mortality, however. A newborn child has a probability π of dying (which depends on the age of the mother and other family characteristics), however if the child survives its first year of life, Wolpin assumed that the risk of death is zero in subsequent years. Thus, even though families may have perfect control over fertility, they still have imperfect control over completed family size due to child mortality. Wolpin assumed that there is a direct monetary cost to childbirth, which depends on the age of the family and whether or not the family had previous children. Children yield direct utility to the family and are also a productive asset in agricultural households. Wolpin estimated the following quadratic family utility function:

$$u_\theta(x_t, \epsilon_t, d_t) = (\theta_1 + \epsilon_t)n_t - \theta_2 n_t^2 + \theta_3 c_t - \theta_4 c_t^2 \tag{5.28}$$

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 $\{\epsilon_t\}$ is an *IID* Gaussian process with mean 0 and variance σ^2, and that $\{x_t\}$ is Markov process independent of $\{\epsilon_t\}$, the family's optimal decision rule $d_t = \delta_t^*(x_t, \epsilon_t)$ can be computed by backward induction starting in the final period $T = 45$ at which the family is fertile. It is not difficult to see that (5.28) implies δ_t^* is given by a threshold rule similar to (5.25). 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 from the last period of life. The assumption that $\{\epsilon_t\}$ is *IID* $N(0, \sigma^2)$ then implies that the conditional choice probability is given by:

$$P_t(d_t|x_t, \theta) = \begin{cases} 1 - \Phi(\eta_t(x_t, \theta)/\sigma) & \text{if } d = 1 \\ \Phi(\eta_t(x_t, \theta)/\sigma) & \text{if } d = 0, \end{cases} \tag{5.29}$$

where Φ is the standard Normal CDF. Wolpin found that the estimated parameters "conform to strong priors; incremental or marginal utilities are positive and diminishing in surviving children and in the composite good, the cost of bearing a child is positive in all life-cycle periods as is the maintenance cost, and the discount factor translates into a rate of time preference of approximately 0.093." (p. 866). The estimated birth costs are estimated to be very high for the first birth, (22,280 Malaysian $), declining to M$8,121 and M$3,314 the second and third births, and then increasing to to M$25,407 at the end of the fertile period: "The costs are high given that the average level of husband's income was about M$12,000 and that the Malaysian dollar was worth about 33 percent of the U.S. dollar in 1960." (p. 866). This particular pattern of childbirth costs appears to a key factor in the model's ability to fit the data, "If the five cost parameters (the c's) are collapsed into a single parameter, the predicted birth probability profit is essentially flat rather than conforming to the actual." (p. 867). Wolpin's more general 5 parameter specification of the birth cost function allowed the DDP model to track the birth process quite closely: "In only one period is there a discernable difference between the actual and predicted probabilities, and there it seems to be the level that is misrepresented and not the pattern. Thus, if there exists a relationship between the propensity to bear a child in any period and prior net fertility, it is complex and the estimated model seems to capture its essential features." (p. 869).

Montgomery (1992) developed a model of contraceptive choice by U.S. women, which accounts for imperfect control over fertility and contraception. In Montgomery's model women choose one of 5 alternative means of contraception including rhythm, condom/diaphragm, the pill. Similar to Wolpin, Montgomery modelled women as having a utility function over completed family size, but with a desired or "target" number of children and asymmetric costs of exceeding or falling short of the target. Montogmery found that the desired number of children varied depending on religion, equaling 2 for non-catholics and slightly over 3 for catholics. His estimated model fit the data quite closely, and a Chi-square specifcation test was unable to reject his specification at the 1% level.

Hotz and Miller (1991) also estimated a DDP model of a family's contraceptive choice which includes a more extreme alternative, sterilization, which is modelled as an absorbing state, i.e. sterilization is assumed to be irreversible. Surprisingly, voluntary sterilization has become the most common method of family planning in the U.S. in recent years, with nearly 40% of women between the ages of 25 and 34 relying on sterilization, more than twice as many as the next most common contraceptive method, the pill. Hotz and Miller assume the family faces three contraceptive alternatives: sterilization ($d = 1$), temporary contraception ($d = 2$), or no contraception ($d = 3$). State variables in this model include the family's previous birth history, the wife's education level, and family income. The use of complete birth histories imply that the dimensionality of the state space is quite large: for example, assuming that births occur on an annual basis, the state variable must assume 1 million values just to account for the 2^{20} birth histories over a 20 year fertile period. Hotz and Miller developed a new estimation method that estimates the primitives (u, p, β) but avoids the need to numerically solve the DP problem using standard DP methods. They estimated three different structural specifications, the most restrictive (model A) assumes the utility flow and childcare expenses are independent of the child's age while the least restrictive (model C) allows utility flows and childcare costs to be step-functions assuming 5 and 6 different values, respectively, over the 20 year span of the child's upbringing. Their estimation results revealed that many of the structural coefficients were imprecisely estimated, and that "the estimated annual expenditure per child are implausibly low." (p.31). Chi-square specification tests of their specifications revealed that "All three are strongly rejected: the significance level of the test statistic for model A is 0.182×10^{-16}, for model B

0.703×10^{-16} and .0034 for Model C. Confronted with these test statistics, we did explore several other configurations of structural coefficients that are not nested by Model C but failed to identify a specification that was not rejected." (p. 31).

5.5 Empirical Application: Models Female Labor Supply

Eckstein and Wolpin (1989) estimated a DDP model of a married woman's labor force participation decision. Previous work in the area of life-cycle labor supply by Heckman (1976) and Heckman and MaCurdy (1980) treated labor supply a continuous decision, and estimated a subset of the structural parameters from the Euler equation. Eckstein and Wolpin treat labor supply as a binary decision (work at a job/work at home), and estimated all the underlying parameters of (u, p, β) by numerically solving the DP problem. Although their model can also accomodate joint fertility/labor supply decisions, Eckstein and Wolpin estimated a model of post-fertility labor supply decisions for a subsample of women aged 39 to 44. The number of children is therefore treated as a time-invariant state variable, and the husband's labor earnings is treated as an exogenous stochastic process. Thus the woman's decision problem reduces to choosing a decision rule for labor supply that maximizes the family's lifetime utility, accounting for the fact that current work experience leads to higher future wages:

$$log\, y_t = \theta_1 + \theta_2 k_{t-1} + \theta_3 k_{t-1} + \theta_4 S + \epsilon_t, \tag{5.30}$$

where y_t is the woman's wage earnings, k_{t-1} is the total number of periods the woman worked up to period $t-1$, S is years of schooling, and ϵ_t is a random component of wages, assumed to be an *IID* Gaussian process. They assumed that the family utility function has the form:

$$u(x_t, d_t) = \theta_5 d_t + c_t + \theta_6 d_t c_t + \theta_7 d_t k_{t-1} + \theta_8 d_t S + \sum_{j=1}^{J} \theta_{8+j} n_{tj} d_t, \tag{5.31}$$

where n_{tj} is the number of children in age category j at time t, and c_t is family consumption, equal to the sum of husband and spouse labor earnings less child care costs and fixed costs of work. If $\theta_7 \neq 0$ the utility function is not intertemporally separable, reflecting "habit persistence" if $\theta_7 > 0$. The optimal decision rule δ_t^* for this problem takes the form of a threshold rule similar to (5.25): the conditional choice probability is given by a formula similar to (5.29) and the cutoffs $\{\eta_t(x, \theta)\}$ are computed by backward induction from the terminal period T. Eckstein and Wolpin found that the estimated parameters conformed to their priors: "Participation reduces utility $(\theta_5 < 0)$ and the marginal utility of children is lower when the woman participates $(\theta_{8+j} < 0)$ particularly for young children. Also, the value of goods consumption is reduced when the woman participates $(\theta_6 < 0)$ and the disutility of work increases with schooling $(\theta_8 < 0)$, i.e. schooling enhances home production. In addition the disutility of work increases with the amount of prior work $(\theta_7 < 0)$, rather than there being habit persistence." (p. 384). A comparison of actual versus predicted participation rates by years of work experience and age shows that the model fits the data fairly closely: "A chi-square test does not reject the null hypothesis that predicted and actual proportions are the same at the five percent level." (p. 385). Eckstein and Wolpin compared the predictive abilities of their structural model with a reduced-form probit model. They found that

> "A χ^2 test using the probit model ... yields a value of 27.1 which is much greater than ... the structural model (7.53). In a loose sense the structural model fits the observed experience-participation profile better than the probit with one less pararameter. Indeed, in order to obtain the structural non-linearities present in (the cutoffs $\{\eta_t\}$) of the model, an incredibly large number of parameters would need to be estimated in any approximation to the $\{\eta_t\}$'s. Therefore, the dynamic model provides a much more parsimonius representation of the data." (p. 386).

Van der Klauuw (1991) estimated a joint model of female marital status and labor supply decisions. Women are assumed to have a choice set consisting of four alternatives: single and not working, single and working, married and not working, and married and working. Similar to the Eckstein-Wolpin model, Van der Klaauw treats the husband's wage earnings as an exogenous stochastic process and allows the woman's wage earnings to depend on previous labor market experience as in equation (5.30), but abstracts from fertility decisions and the existence of children. Van der Klaauw estimated his model using a two-step procedure in which he first estimated reduced-form coefficients of a flexible linear approximation to the value function, and then used a minimum distance estimator to find structural estimates that minimize a quadratic distance between the reduced-form and structural estimates. The estimation results indicate that 1) higher wages have a positive and significant effect on the probability of working, 2) increases in the husband's wages decrease the gain in utility from working when married, 3) non-whites receive lower utility from being married than whites, 4) more years of schooling increase the utility of working for single women (and also for married women but to a lesser extent), 5) there are large and significant utility costs to divorce, and, 6) unlike the findings of Eckstein and Wolpin, Van der Klaauw finds that the disutility of work decreases with the amount of prior work reflecting possible habit persistence. In terms of the ability of his model to fit the data, Van der Klaauw concluded that "Except of the first and the last few years, the predicted trend in the proportion of women that are married is very close to the actually observed trend", but "The predicted trend in the labor force participation rate ... does not seem to capture the actual trend that well. In particular, it does not predict the drop in participation rates that occurs in the fourth year (after leaving school)." (p. 32).

5.6 Empirical Application: Retirement Behavior

Gotz and McCall (1984) estimated a *dynamic retention model* of an Air Force pilot's decision whether or not to remain in the service or leave for a potentially higher paying job as a civilian pilot. In the air force salaries and job privileges are determined by years of service and rank. There is a well-defined promotion ladder in which pilots progress from the lowest rank (Captain) to the highest (Colonel) in stages. Gotz and McCall model the promotion process as an exogenous Markov chain in which pilots face a probability $\pi_t(r_{t+1}|r_t)$ of being promoted from rank r_t

to r_{t+1} in the t^{th} year of service.[48] If a pilot chooses to leave the service at any point they receive pension equal to a fraction f_t of their basic pay, where f_t is given by:

$$f_t = \begin{cases} 0 & \text{if } t < 20 \\ .025t & \text{if } 20 \leq t < 30 \\ .75 & \text{if } t \geq 30 \end{cases} \tag{5.32}$$

If the pilot leaves the Air Force, he can become employed in a civilian job without jepordizing his military pension. If we let $W_t(r)$ denote the discounted wage earnings from civilian employment expected by a pilot of rank r who leaves the Air Force after t years of service, then the combined pecuniary value to leaving $(d_t = 0)$ is given by:

$$v_t(r,0) = f_t[w(t,r) - a(t,r)] \sum_{j=t+1}^{\infty} s_{tj}\beta^{j-t} + W_t(r). \tag{5.33}$$

where $w(t,r)$ is basic military pay for rank r after t years of service, $a(t,r)$ are allowances not counted in retirement pay calculations, and s_{tj} is the probability of surviving to year j conditional on survival to year t. If the pilot chooses to remain in the Air Force $(d_t = 1)$, the value function is given by:

$$\begin{aligned} v_t(r,\epsilon,1) &= \epsilon + \mu + \beta \sum_{r'} \left[w(t,r') + \int V_{t+1}(r',\epsilon')\Phi(d\epsilon') \right] s_{t,t+1}\pi_t(r'|r) \\ V_{t+1}(r,\epsilon) &= max\left[v_{t+1}(r,0), v_{t+1}(r,\epsilon,1) \right]. \end{aligned} \tag{5.34}$$

where μ is a time-invariant person-specific constant representing nonpecuniary returns to staying in the Air Force (such as camraderie, patriotism, etc.) and ϵ is a time-varying unobserved state variables representing transitory frustations and rewards experienced by a pilot, assumed to follow an *IID* Gaussian process with distribution Φ. It follows that the pilot's stay/leave decision rule is given by:

$$\delta_t^*(r,\epsilon) = \begin{cases} 1 & \text{if } v_t(r,\epsilon,1) > v_t(r,0) \\ 0 & \text{otherwise} \end{cases} \tag{5.35}$$

Gotz and McCall estimated the structural parameters of their model by maximum likelihood, using the partial likelihood estimator with conditional choice probability $P(d|r,\mu)$ computed from (5.35) by integrating out ϵ similar to (5.29), but using a "mixture likelihood" which also integrates out the time-invariant person-specific components μ. Their estimation results revealed substantial person-to-person differences in non-pecuniary valuations of service, reflected in a large estimated variance for the population distribution of μ, although the dispersion in μ was substantially less for graduates of the Air Force academy. Regarding the goodness of fit of their model, Gotz an McCall found that "As is commonly true of the χ^2 statistics with very large sample sizes, the test statistics are almost invariably larger than their degrees of freedom and most are significant at conventional levels." (p. 25). However they also observe that

[48] Demotions are extremely rare the Air Force, so Gotz and McCall asssume that $\pi_t(r'|r) = 0$ for any rank r' which is lower than r.

"Statistically significant errors may not be important errors" and that a cell by cell investigation of prediction errors revealed that "the match between predicted and actual (retention rates) is close. The tendency for actual retention rates to rise as years of service increase is predicted by the dynamic retention model. The dynamic retention model also captures the difference between regular and reserve retention rates. In general, the difference in these retention rates is greatest at EOB[49] and declines in later years of service; this is what the retention model predicts." (p. 27). Gotz and McCall compared the predictive ability of the dynamic retention model to two competing "semi-structural" models, PVCOL (present value cost of leaving) and ACOL (actual cost of leaving). They found the dynamic retention model yielded much more accurate predictions of pilots' behavior than the competing models. Table 5.1 presents χ^2 values for the three models for three aeronautical ratings: pilots, navigators, and nonrated. In out-of-sample predictive tests over the period 1977 to 1981, they found that although "Pilot retention in the later years of service in 1979 to 1981 is persistently overpredicted" (p. 53), in view of "the very large fluctuations in civilian airline hiring prospects for pilots; the promotion opportunities to major, lieutenant colonel, and colonel each increased during the period; and the military received a substantial real pay raise in the last year of the period. The model accounts for these changes fairly well." (p. 55).

Daula and Moffitt (1991) estimated a dynamic programming model of reenlistment decisions by Army personnel. Their model is similar to the Gotz and McCall model, although they make a number of simplifications in the specification and solution of the DP problem. Unlike Gotz and McCall, they estimated rather than fixed the discount factor ($\hat{\beta} = .948$), and found that "reenlistment probabilities are significantly higher for those with more dependents, black enlistees, those with higher educational levels and with higher relative pay grades, and for those coming up for reenlistment prior to 1983." (p. 10). They compared ability of their DDP model with the ACOL model and a naive model that restricts $\beta = 0$. Although they did not perform goodness of fit tests, a comparison of the log-likelihood values revealed that the "dynamic programming model yields the best fit ($L = -314.55$), the naive model the worst ($L = -319.23$), and the ACOL in between ($L = -316.82$)" (p. 14).

Berkovec and Stern (1991) estimated a model of labor force participation at the end of the life cycle. In their model a person has three possible decisions each period: work at a full-time job ($d = 1$), work at a part-time job ($d = 2$) or don't work ($d = 3$). Unlike previous models of retirement behavior, the Berkovec and Stern model does not assume that retirement is an absorbing state: at any point in time an unemployed worker can choose to "unretire" by returning to work on a full or part-time basis. This feature is essential to capture the the significant fraction of reverse transitions made by at least 30% of the population. Berkovec and Stern incorporated unobserved state variables in a specification similar to Gotz and McCall: a time-invariant person-specific state component μ and a time-varying unobserved state variable ϵ_t, treated at an *IID* extreme value process. Condiitonal on μ the extreme value specification

---

[49] End of obligation, the first year in which a pilot can voluntarily leave the Air Force.

implies a tri-variate logit formula for the conditional choice probabilities, $P(d|x,\mu)$ as in (5.10). However rather than integrate the partial likelihood function (5.13) with respect to the population distribution of μ, Berkovec and Stern used a simulation estimator which can be viewed as using monte carlo integration instead of the more computationally intensive numerical integration method of Gotz and McCall. Their model also allows for observed heterogeneity via the variables education, race and health, under the assumption that the latter variable is time-invariant. They estimated two models, a myopic model with $\beta = 0$ and a dynamic model with $\beta = .95$. Their estimation results reveals that in the case of the dynamic model "one more year of education decreases the value of retirement (relative to a new full-time job) by \$76.20 per year and increases the relative value of a part-time job by \$42.70 per year. For both models, bad health decreases the relative value of a full-time job, education decreases the relative value of retiring, and age increases the relative value of retiring. Furthermore, both models explain job attachment mainly through the high cost of changing jobs (\$12,000 for the dynamic model and \$19,000 for the static model). Finally neither model has an estimated flexibility factor ρ close to one. [50] This indicates that the relationship between unobservables in value functions and in wage equations is probably misspecified." (pp. 202–203). Berkovec and Stern did not conduct explicit specification or goodness of fit tests, however it is evident from their plots of the predicted retirement hazard rate that their model fails to predict the observed peaks in the retirement hazard at ages 62 and 65. They conclude that

> "The model estimated in this paper is flawed in that it does not include the Social Security system and it has unreasonable assumptions about the information available to each individual (e.g. health status does not change over time). The first flaw explains the lack of a high hazard rate into retirement at ages 62 and 65. Both of these flaws can be corrected using the same modelling and estimation procedure suggested in this paper, though at an increased cost. We feel that this methodology has great promise in estimating structural dyanmic models." (p. 209).

Phelan and Rust (1991) present estimation results for a simplified version of a DP model of retirement behavior developed by Rust (1989). Similar to the Berkovec–Stern model there are three labor supply decisions, (full-time/part-time/unemployed). In addition, Social Security benefits are explicitly modelled, adding an additional state variables and control variable (apply/don't apply). [51] This decision is only relevant for eligible workers: i.e. those who are at least 62 and who have accumulated 40 quarters of coverage over their working career. The model accomodates heterogeneity via a set of observed state variables rather than via unobserved heterogeneity. The vector of time-varying state variables in this model includes age, income, health status (good health/bad health/dead), employment status (full-time/part-time/unemployed), marital status (married/single), Social Security status (ineligible/eligible), age at which the worker first applied for benefits, and the Social Security average monthly wage. The latter three variables determine the Social Security benefits a worker expects to receive when he retires. Benefits are determined from the Primary Insurance Amount (PIA), a nonlinear function of the worker's average monthly wage. The actual benefits a

---

[50] The flexibility factor is a coefficient of the integrated log-sum term in equation (5.11) which the theory predicts should be equal to 1.

[51] The decision of whether to apply for disability benefits prior to age 62 was not included in this version of the model.

worker receives depends on the age at which benefits are first collected, increasing from 80% of the PIA at age 62 to 100% of the PIA at age 65, with a 1% per year delayed retirement credit for each year of delay after age 65. Benefits are also contingent on the worker's future labor supply decisions: once a worker qualifies for Social Security, $1 of benefits is lost for every $2 of wage earnings above a fairly low threshold known as the "earnings test" level ($2,000 in 1968 dollars). The model was estimated using RHS[52] data which followed 8,131 men aged 58-63 in 1969 through the subsequent decade. Since the RHS has incomplete data on the details of private pension plans, Phelan and Rust focused on a subsample of workers for whom Social Security was the only source of old age benefits.

Given that significant heterogeneity is already accounted for by the observed state variables, Rust's model does not attempt to build-in additional unobserved heterogenity via time-invariant person-specific effects μ. Instead only transient unobserved state variables ϵ_t are allowed, which are assumed to follow an *IID* extreme value process. This implies that the conditional choice probabilities have the multinomial logit form given in equation (5.10). Discretizing income into 21 intervals, the value function $v(x, d)$ and choice probability $P(d|x)$ are arrays with 417,312 elements and the transition probability matrix $\pi(x_{t+1}|x_t, d_t)$ has over 3 million elements, after excluding zero probability transitions.[53] The DDP model was estimated using a utility function of the form:

$$u(x, d, \theta) = \left[\frac{y^{\theta_1} - 1}{\theta_1}\right] exp\left\{\theta_2 + \theta_3 I\{\text{h=2}\} + \theta_4 I\{ms = 2\} + \theta_5 \frac{a}{1+a}\right\} + \sum_{i=1}^{3}\sum_{j=1}^{3} \theta_{ij} I\{e = i, d = j\} \tag{5.36}$$

$$x = (y, h, ms, a, e) \quad e, d \in \{1, 2, 3\} \equiv \{\text{FT,PT,NE}\}$$

This utility function consists of two pieces, a CRR utility of income function is affected by the man's age (a), whether he is in bad health $(h = 2)$, or married $(ms = 2)$. The second component of utility, the double sum term in (5.36), reflects "labor/leisure/search" decisions. The coefficients θ_{ij} represent the utilities associated with different employment state/decision combinations. Thus, θ_{11} represents the utility associated with staying in a full-time job, θ_{33} represents the utility of remaining unemployed, and θ_{31} represents the (dis)utility associated with searching for a new full-time job while unemployed. Similar to the Berkovec-Stern model the model involves an implicit "perfect control" assumption: a worker who decides to return to work is guaranteed to find a job with probability 1. However there is uncertainty about the wage earnings a worker can expect to receive, and these uncertainties are embodied in the estimated transition probability matrix $\hat{\pi}(x_{t+1}|x_t, d_t)$. The Social Security rules are also embodied in $\hat{\pi}$ along with the worker's perceptions of mortality, his chances for separating from his wife, and the chances that his health will turn bad.[54] The DDP model was estimated in two-stage procedure in which π was estimated in the first stage using

---

[52] Retirement History Survey.

[53] For example age satisfies the deterministic relation $a_{t+1} = a_t + 1$: all other values have probability 0.

[54] Due to the high dimensionality of π relative to the available 7,300 person/year observations, π was decomposed as a product of individual transition probabilities for income, health, marital status and death.

semi-parametric techniques, and the parameters θ of the utility function $u(x, d, \theta)$ were estimated in the second stage using the partial likelihood function (5.13). The estimation results revealed significant risk aversion with an estimated value of θ_1 very close to 0. The health status coefficient θ_2 was not significantly different from 0, but the significant negative coefficient on θ_3 was interpreted as a "marriage bonus" in the sense that "the estimated value of $\theta_3 = -.349$ is significantly greater than the value $-.7$ needed to evaluate the utility of income on a per capita basis." (p. 18). In addition, "The θ_{ij} coefficients show that the utility of leisure θ_{33} is higher than the utility of part or full-time work (θ_{11} or θ_{22}), and that older unemployed workers experience considerable disutility in searching for a new full-time job (he experiences a drop of approximately 10 utils in searching for either a full or part-time job compared to not search at all). There is also evidence of a 'retirement party' effect, where workers receive a one-time surge in utility when they quit their jobs." (p. 18). In terms of the ability of the model to fit the data, high values of Chi-squared goodness of fit statistics indicate misspecification, although the actual values of these statistics are overestimated since Phelan and Rust did not adjust the statistic to account for the estimation error in the semi-parametric estimates of $\hat{\pi}$. In a direct analysis of the predictions of the DDP model, Phelan and Rust found that "The DP model does a good job of tracking job search decisions of full and part-time workers prior to applying for Social Security. The DP model also does a good job of predicting the decision of full-time workers aged 60-61 as they apply for early Social Security benefits starting at age 62-63. The DP model also does a tolerable job of predicting the job search decisions of workers who are unemployed." (pp. 23–25). However while the DDP model can capture the peak in early retirements at age 62, it is unable to fully capture the other peak at the normal retirement age of 65. Table 5.2 compares the non-parametric (NP) estimates of $P(d|x)$ with the predictions of the DDP model (DP) for full-time workers between ages 61 and 67. The six columns in the table correspond to the 3 employment decisions (FT,PT,NE) and the 2 Social Security decisions (AB,DB), where AB stands for "apply for benefits" and DB stands for "don't apply for benefits". "Note that up until age 64-65 the DP model predicts actual behavior fairly closely, whereas at age 64-65 it overpredicts the number who continue working without applying for SS and at age 66-67 it underpredicts this same transition (although we are less confident about the 66-67 age group since the number of observations is very small). Overall, transition rates in the data seem to make an abrupt change at age 64-65, when many fewer workers choose (FT,DB) and many more workes choose (FT,AB). By age 66-67 the prediction errors are reversed: now the DP model underpredicts the fraction of workers choosing (FT,DB) and overpredicts the fraction choosing (FT,AB) (and (PT,AB) and (NE,AB) as well)." (p. 27). Phelan and Rust offer two possible explanations for these prediction errors: 1) there may be a sociological phenomenon that 65 is a "customary retirement age" at which workers are expected to retire, or 2) workers applying for Social Security at 65 might simply be exercising a "Medicare insurance option".[55] Phelan and Rust conclude that:

[55] Under Social Security rules, people aren't eligible for Medicare until they apply for Social Security benefit at age 65. Exercising this option is costless in the sense that even if a worker intended to continue working into his 70's, the "automatic recomputation of benefits" provision of the Social Security rules insures that all intervening earning will be credited to his account and his average monthly wage adjusted accordingly.

"we think the [results in table 5.2] cast some doubt on the 'customary retirement age' hypothesis in that if one is 'expected' to retire at age 65, then it seems equally reasonable to suppose one should be 'expected' to retire at later ages. Yet we can see that among age 66-67 workers, retirement rates suddenly fall. We think that the 'Medicare option' hypothesis could provide a much more compelling explanation for this behavior, but we won't know for sure until we have formally incorporated the health insurance motive into the DP model. Despite the fact that there are several areas where the specification of the DP model can be improved, we view our initial results as extremely promising. Indeed, each of the areas where the DP model does relatively poorly can be fixed by making plausible modifications to the structure. This includes incorporating health care costs and modelling the underlying decision to apply for disability insurance benefits. In addition, we think that finer classifications of the health and employment status will enable the DP model to more accurately capture the self-selection and incentive effects induced by Social Security policy." (p. 28)

Lumsdaine Stock and Wise (1991) (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. Needlesss 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 similar in spirit to the PVCOL and ACOL models used by Gotz and McCall and Daula and Moffitt.[56] Two of the models were structural 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:

$$u(x_t, \epsilon_t, d_t, \theta) = \begin{cases} x_t^{\theta_1} + \mu_1 + \epsilon_t(1) & \text{if } d_t = 1 \\ (\mu_2 \theta_2 x_t)^{\theta_1} + \epsilon_t(0) & \text{if } d_t = 0 \end{cases} \tag{5.37}$$

The two components of $\mu = (\mu_1, \mu_2)$ represent time-invariant worker-specific heterogeneity. In the extreme value specification for $\{\epsilon_t\}$, μ_1 is assumed to be identically 0 and μ_2 is assumed to have a log-normal population distribution

[56] Several specifications were tried 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 benfits ("pension accruals"), predicted earnings in the next year of employment, and age.

with mean 1 and scale parameter θ_3. In the Gaussian specification for $\{\epsilon_t\}$, 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 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 final model is the so-called "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^*.[57]

The parameter estimates for the three structural models are presented in Table 5.3. There are significant differences in the parameter estimates in the three models, the Gaussian DP specification predicting 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.[58]

Figures 5.4 to 5.6 summarize the ability of the four models to fit the historical data. Figure 5.4 compares the actual departure rates (solid line) to the departure rates predicted by the option value model (dashed line). Figure 5.5 presents a similar comparison for the two DDP models, and figure 5.6 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 5.4 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 the age trends that are excluded from the other models.

Figures 5.7 to 5.9 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 parmeters predictions were generated from all four models after incorporating the extra-bonus provisions of the

---

[57] 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 (1991) 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.

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

window plan. As is evident from figures 5.7–5.9, 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 were 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 5.4 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.[59] LSW concluded that

> "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 suggest that the option value and dynamic programming models are considerably more successful than the less complex probit model in approximating the rules 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 susing different data sets and with respect to different pension plan provisions." (p. 31).

6. Conclusion

Do people behave according to Bellman's principle? Section 3 showed that without further restrictions on the primitives (u, p, β) the answer is trivially "yes": we can always find a specification that "rationalizes" a given set of observations. However if we willing to impose restrictions on (u, p, β) that reflect our prior notions about specifications for preferences and beliefs that are "reasonable", then the evidence is mixed. While I have not attempted to do an exhaustive review of the empirical literature, my reading of the empirical evidence is that most of the continuous decision problems tested on macro data have been soundly rejected, whereas many of the discrete decision problems tested on micro data have provided fairly good, although by no means perfect, approximations to observed behavior.

It seems likely that the strong rejections of the CDP models are largely a rejection of heroic aggregation and representative consumer assumptions rather than rejections of the basic behavioral model. Although there may be other reasons for estimating representative consumer models using macro data, I would have to agree with Deaton (1991) that "conditions for aggregation to representative agents are implausible, so that a representative agent formulation is even more than usually misdirected" (p. 1241) in cases where we want the CDP model to reflect realistic features such as borrowing constraints, hours of work constraints, or other restrictions and nonconvexities in the choice set. The results in section 4 indicate that CDP models have generally been most successful when tested on micro data, or when heterogeneous agent micro models have been consistently aggregated to test implications at the macro level. Cochrane and Hansen (1992) have also criticized the macro literature on the ground that "Asset market data are often avoided in

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

evaluating macroeconomic models, and aggregate quantity data are often avoided in empirical investigations of asset market returns," arguing that "such practices are mistakes and that security market data are one of the most sensitive and hence useful proving grounds for macroeconomic models." (p.2). However it is striking there has been even less empirical work using *micro* asset and quantity data which would seem to be by far the best way of testing these models. To my knowledge one of the few studies to use such data is MaCurdy and Shoven (1991) who investigated retirement portfolio allocation decisions of individual faculty members. They find that most faculty members allocate the majority of their portfolios to bonds rather than stocks; this behavior appears to be irrational in light of very convincing evidence that stocks systematically out-perform bonds over long horizons.[60]

Abstracting from the problems of using macro data, why have the DDP models generally provided better approximations to behavior than CDP models? A simple answer might be that choosing from a finite set of alternatives may be inherently simpler than choosing from a continuum of alternatives. Indeed, many of the most successful DDP models are restricted to binary choice problems, particularly optimal stopping problems. Green and Osband (1991) conjectured that "It is possible that optimal-stopping problems are somehow psychologically special, and the people's performance with respect to these problems corresponds more closely to expected utility theory than does their general performance." (p. 677). Indeed, as we expand the number of states and available decisions, we quickly run into "the curse of dimensionality" which makes such problems vastly more difficult for us to solve by DP methods. Although we currently have limited experience in estimating very large DDP models, it seems likely that it will be much more difficult to find specifications that fit the data.

Considerations of computational complexity may also force individuals to adopt simple, heuristic decision rules that might be very different than the optimal decision rules calculated by dynamic programming. Indeed, introspection suggests that it is improbable that we literally use backward induction calculations in determining our behavior, at least at a conscious level. It seems much more likely that people use some sort of "forward induction" procedure to prune branches of the decision tree that seem unlikely to yield high payoffs rather than methodically assigning values to each node by backward induction calculations. However a forward induction procedure is not necessarily inconsistent with the results of backward induction: by pruning inessential branches of the decision tree, a person may be able to more clearly focus on the subtree where most of the real action is, and decisions within this subtree may be much more closely approximated by backward induction.[61] Also, decision rules may be fairly robust to variations in valuations attached to nodes far in the future: people may make decisions similar to the way a good AI program plays

---

[60] It may be that the faculty beliefs are irrational and that their behavior may not necessarily be irrational (i.e. non-optimizing) given their beliefs. Although stocks are known to have a higher rate of return, they involve substantial short term risk. The finding that despite this risk, stock returns systematically dominate bond returns in the long run came as a surprise to MaCurdy and Shoven, who subsequently changed retirement portfolios and invested heavily in stocks.

[61] Winston (1992) describes a number of heuristic pruning procedures for decision trees, including "branch and bound" methods, that have been developed in the literature on artificial intelligence.

chess: it assigns more or less heuristic weights to provide summary evaluations of board positions many moves in the future, while using "brute force" calculations to systematically evaluate the consequences of intermediate moves. These sorts of hueristics may result in decision rules that closely approximate an optimal decision rule. This might be one way to interpret empirical work such as Cochrane (1989) who compared the loss in utility involved in using the simple decision rule of setting consumption equal to income instead of the optimal permanent income decision rule. Cochrane found that the utility loss is insignificant: "less than 10¢–$1 per quarter in environments specified by popular tests on aggregate data" (p. 319). Cochrane's result can be interpreted as an empirical discovery of Deaton's (1991) result that "In the limit, when labor income is a random walk, it is optimal for impatient liquidity-constained consumers simply to consume their incomes." (p. 1221). It seems plausible that many consumers might be smart enough to empirically discover such approximate rules as well. For this reason I think research on learning procedures that can "discover" simple, heuristic decision rules that approximate optimal decision rules will have very important payoffs. An example of this approach is the "linear reward-inaction" (LRI) learning algorithm analyzed by Narendra and Wheeler (1986). They studied an agent who was unaware of the functional forms of their preferences u and beliefs p, but who could experience the utility associated with any specific state and decision, $u(x, d)$. If the agent uses the LRI learning algorithm, Narendra and Wheeler demonstrate that the agent eventually "discovers" the optimal decision rule in the sense that expected long run average utility under the LRI procedure coincides with the maximum expected long-run average utility under an optimal decision rule.[62]

This suggests that it may not be very fruitful to pose the problem in overly narrow terms: "are people literally performing backward induction?". Instead, if we pose the issue in terms of Friedman's positivist approach: "do decision rules generated from dynamic programming provide good approximations to actual behavior?", or even better, "do DP models out-perform other behavioral theories?", then I think the results of this survey are encouraging. For example the research on retirement behavior reviewed in section 5.6, the evidence strongly suggests that people are forward looking and do become informed of the important details of their Social Security and pension plan provisions that as they approach retirement age, and that DP models which build in the details of these provisions are able to provide much more accurate predictions of people's behavior than competing reduced-form statistical models or semi-structural behavioral models. It is actually somewhat surprising that these fairly restrictive, parsimonious, and abstract parametric models should do as well as they have, given what we know about expected utility theory. All of these models focus on a very narrow subset of the full set of decisions a person must make within a day, or over a lifespan. As Kreps (1988) has noted, even if the person's global decisionmaking conforms to the expected utility hypothesis, there is no guarantee that the person will act as an expected utility maximizer in any sub-problem:

---

[62] Unfortunately, the LRI proceudre converges very slowly in high-dimensional problems reflecting a much slower learning rate than we typically observe in human subjects.

> "When there is correlation between the gambles over which Totrep is optimizing and other uncertainty is left out of the application, when other decisions are being made simultaneously to those in the application, or when the uncertainty of the gambles being optimized over resolves at dates in the future (after important decisions left out must be taken), then the use of the standard models is *very* suspect and often quite wrong. We have seen those conclusions for the von Neumann-Morgenstern model. Similar conclusions hold for the Savage model. Since this includes just about every important economic/social decision making context that you are likely to encounter, there is obviously a need for care in application." (p. 175).

A cynical view is that the empirical "successes" are just a reflection of the identification problem discussed in section 3: the flexibility of these models gives us considerable latitude to devise simple parametric models that look "reasonable" and do a good job of fitting the data. The ideal place to test Bellman's principle is in the laboratory where we can exercise control over people's preferences and beliefs, and therefore devise much more discriminating tests. I have not reviewed this literature in this paper, partly because it is already covered in Camerer's (1992) comprehensive survey on "Individual Decision Making", and partly because as Camerer notes, "There have been relatively few experimental tests of choice over time." (p. 75). Similar to the econometric results reviewed here, the available experimental results are mixed. Many experiments have tested the implications of the standard search model. Although people's behavior does not conform exactly to an optimal search strategy, Camerer concludes that, "Roughly speaking, people do this task remarkably well."[63] On the other hand, there are examples where human subjects perform abysmally. For example, Johnson, Kotlikoff and Samuelson's (1988) experimental study of the life-cycle model of consumption attempted to induce intertemporal utility functions on student subjects to see if their choices corresponded to the optimal life-cycle consumption rule. They concluded that "The subject's responses suggest a widespread inability to make coherent and consistent consumption decisions. Errors in consumption decision-making appear to be very substantial and, in many cases, systematic. In addition, the experiment's data strongly reject the standard homothetic, time-separable life cycle model." (p. 488). However experiments are subject to their own set of of criticisms, namely, that the decision making environment is artificial and the subject payments may not provide sufficiently high stakes to motivate people to take the right actions. In addition, as Green and Osband have noted, experiments may not provide subjects with sufficient opportunities to learn:

> "It is also possible, though, that the experimental evidence systematically under-represents people's conformity to expected utility theory. This under-representation may be especially severe with respect to decision making by experts. Expected-utility maximization with respect to an area of substantive expertise is a learned, non-verbal, domain-specific skill. Such skill may have to be acquired slowly and through intensive study (for instance, over a period of several years of apprenticeship or postgraduate professional education), and it may not transfer instantaneously to an artificial task in an experiment, even after it has been acquired. These considerations show the great importance of further examining the expected utility hypothesis outside the laboratory, and particularly with respect to a variety of decision situations faced by experts." (p. 678).

---

[63] Camerer, private communication, March, 1992. For a more detailed review, see section 11 of Camerer, (1992).

Thus, it seems that our best hope at success lies in parallel investigations using both structural econometric and experimental methods. In evaluating econometric models, it is particularly important to validate the predictions of estimated DP models using "natural experiments" such as the window plan analyzed by Lumsdaine, Stock and Wise. Too much of the econometric literature has focused on using ever more sophisticated techniques in parameter estimation "exercises", going no further than checking whether the estimated parameters have the "right" signs and magnitdues rather than systematically evaluating how well estimated models predict actual behavior. In addition, whenever possible it is preferable to examine all the implications of the Bellman's principle by completely solving the DP problem rather than relying exclusively on tests of the validity of necessary conditions such as the Euler equation. West's inventory model provides a striking example that an estimated model can appear to have reasonable parameters and pass standard orthogonality tests, yet the optimal decision rule implied by those parameters can be absurd. Only by passing a rigorous battery of in-sample and out-of-sample predictive tests we will start to have enough confidence to use these models in the ambitious policy forecasting exercises suggested by Marschak (1951) and Lucas (1976).

Postscript: What would Bellman have thought about all this? Many mathematicians and statisticians are skeptical that anyone less intelligent than themselves could possibly behave according to their theories. My guess is that Bellman would be quite supportive of the hypothesis that ordinary people behave "as if" they were solving DP problems. In fact, at the time of his death Bellman was collaborating on an empirical project with Michael Hartley, estimating a DP model of the "tree-crop problem" of Indonesian farmers.

7. References

Aedo, M. (1989) "The Schooling Decision: A Dynamic Model" manuscript, University of Minnesota.

Altûg, S. (1989) "Time-to-Build and Aggregate Fluctuations: Some New Evidence" *International Economic Review* **30-4** 889–920.

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

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

Attanasio, O. Browning, M. (1991) "Consumption over the Life Cycle and Over the Business Cycle" manuscript, Stanford University.

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

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

Bertsekas, D. (1987) *Dynamic Programming Deterministic and Stochastic Models* Prentice Hall, New York.

Bhattacharya, R.N. Majumdar, M. (1989) "Controlled Semi-Markov Models – The Discounted Case" *Journal of Statistical Planning and Inference* **21** 365–381

Bierens, H. (1990) "A Consistent Conditional Moment Test of Functional Form" *Econometrica* **58-6** 1443–1458.

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* 5^{th} *Berkeley Symposium on Mathematical Statistics and Probability* **1** 415–418.

Blanchard, O.J. (1983) "The Production and Inventory Behavior of the American Automobile Industry" *Journal of Political Economy* **91** 365–400.

Blau, D. Robins, P.K. (1990) "Job Search Outcomes for the Employed and Unemployed" *Journal of Political Economy* **98-3** 637–655.

Blau, D. (1991) "A Structual Model of Retirement-Related Labor Market Dynamics" manuscript, University of North Carolina–Chapel Hill.

Blundell, R. (1990) "Evaluating Structural Microeconometric Models of Labour Supply" forthcoming in J. Laffont and C. Sims (eds.) *Advances in Economietrics, Sixth World Congress* Econometric Society Monographs, Cambridge University Press.

Brock, W.A. (1982) "Asset Prices in a Production Economy" in J.J. McCall (ed.) *The Economics of Information and Uncertainty* University of Chicago Press.

Brock, W.A. Mirman, L.J. (1972) "Optimal Economic Growth Under Uncertainty: The Discounted Case" *Journal of Economic Theory* **4** 479–513.

Camerer, C. (1992) "Individual Decision Making" forthcoming in J. Kagel and A. Roth (eds.) *Handbook of Experimental Economics* Princeton University Press.

Christensen, B. (1990) "Estimation of Dynamic Programming Models with Applications to Intertemporal Equilibrium Asset Pricing Models" manuscript, New York University.

Christensen, B.J. Kiefer, N.M. (1991) "The Exact Likelihood Function for an Empirical Job Search Model" *Econometric Theory.*

Christensen, B.J. Kiefer, N.M. (1991) "Measurement Error in the Prototypical Job Search Model" manuscript, Cornell University.

Cochrane, J. (1989) "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives" *American Economic Review* **79-3** 319-337.

Cochrane, J. Hansen, L.P. (1992) "Asset Pricing Lessons for Macroeconomics" manuscript, University of Chicago.

Cogley, T. Nason, J.M. (1991) "Effects of the Hodrick-Prescott Filter on Integrated Time Series" manuscript, University of British Columbia.

Constantinides, G. (1990) "Habit Formation: A Resolution of the Equity Premium Puzzle" *Journal of Political Economy* **98-3** 519–543.

Das, M. (1991) "A Micro Econometrica Model of Capital Utilization and Retirement: The Case of the Cement Industry" *Review of Economic Studies.*

Daula, T. Moffitt, R. (1991) "Estimating a Dynamic Programming Model of Army Reenlistment Behavior' *Military Compensation and Personnel Retention.*

Davis, P.J. Rabinowitz, P. (1975) *Methods of Numerical Integration* Academic Press, New York.

de Gorter, H. Fisher, E. (1989) "The Dynamic Effects of Agricultural Subsidy Programs in the United States" manuscript Cornell University.

Deaton, A. (1991) "Saving and Liquidity Constraints" *Econometrica* **59-5** 1221–1248.

Deaton, A. (1987) "Life-cycle Models of Consumption: Is the Evidence Consistent with the Theory?" in Truman Bewley (ed.) *Advances in Econometrics – Fight World Congress* Econometric Society Monographs, Cambridge University Press.

Denardo, E.V. (1967) "Contraction Mappings Underlying the Theory of Dynamic Programming" *SIAM Review* **9** 165–177.

Dynkin, E.B. Juskevic, A.A. (1979) *Controlled Markov Processes* Springer-Verlag, New York.

Eckstein, Z, Wolpin, K. (1989) "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models" *Journal of Human Resources* **24-4** 562–598.

Eckstein, Z. Wolpin (1989) "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience" *Review of Economic Studies* **56** 375–390.

Eckstein, Z, Wolpin, K. (1990) "Estimating a Market Equilibrium Search Model from Panel Data on Individuals" *Econometrica* **58-4** 783–808.

Eckstein, Z. Wolpin K.I. (1992) "Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model" manuscript, University of Minnesota.

Eichenbaum, M. (1991) "Real Business-Cycle Theory: Wisdom or Whimsy?" *Journal of Economic Dynamics and Control* **15** 607–626.

Epstein, L.G. (1988) "Risk Aversion and Asset Prices" *Journal of Monetary Economics* **22** 179–192.

Epstein, L.G. Zin, S.E. (1989) "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework" *Econometrica* **57-4** 937–970.

Epstein, L.G. Zin, S.E. (1990) "'First-Order Risk Aversion and the Equity Premium Puzzle" *Journal of Monetary Economics* **26** 387–407.

Fafchamps, M. (1991) "Sequential Labor Decisions Under Uncertainty: An Estimable Household Model of West-African Farmers" manuscript, Stanford University.

Flavin, M. (1981) "The Adjustment of Consumption to Changing Expectations of Consumption" *Journal of Political Economy* **89** 974–1009.

Flinn, C. (1986) "Wages and Job Mobility of Young Workers" *Journal of Political Economy* **94-3** S88–S110.

Gihman, I.I. Skorohod, A.V. (1974, 1975, 1979) *The Theory of Stochastic Processes* Volumes I, II, III. Springer-Verlag, New York.

Gihman, I.I. Skorohod, A.V. (1979) *Controlled Stochastic Processes* Springer-Verlag, New York.

Gittens, J.C. Jones, D.M. (1974) "A Dynamic Allocation Index for the Sequential Design of Experiments" in J.M. Gani (ed.) *Progress in Statistics* North-Holland.

Gotz, G.A. McCall, J.J. (1980) "Estimation in Sequential Decisionmaking Models: A Methodological Note" *Economics Letters* **6** 131–136.

Gotz, G.A. McCall, J.J. (1984) "A Dynamic Retention Model for Air Force Officers" Report R-3028-AF, the RAND Corporation, Santa Monica, California.

Green, E.J. Osband, K. (1991) "A Revealed Preference Theory for Expected Uility" *Review of Economic Studies* **58** 677–696.

Gustman, A. L. Steinmeier, T.L. (1986) "A Structual Retirement Model" *Econometrica* **54** 555–584.

Häerdle, W. (1990) *Applied Nonparametric Regression* Econometric Society Monographs, Cambridge University Press.

Haavelmo, T. (1944) "The Probability Approach in Econometrics" *Econometrica Supplement* **12** 1-115.

Hajivassiliou, V. Ioannides, Y.M. (1991) "Dynamic Switching Regression Models of the Euler Equation: Consumption and Liquidity Constraints" manuscript, Yale University.

Hakansson, N. (1970) "Optimal Investment and Consumption Strategies Under Risk for a Class of Utility Functions" *Econometrica* **38** 587–607.

Hall, R. (1978) "Stochastic Implications of the Life Cycle – Permanent Income Hypothesis: Theory and Evidence" *Journal of Political Economy* **86-6** 971–987.

Hall, R. Mishkin, F.S. (1982) "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households" *Econometrica* **50-2** 461–481.

Hansen, L.P. (1982) "Large Sample Properties of Method of Moments Estimators" *Econometrica* **50** 1029–1054.

Hansen, L.P. Sargent, T.J. (1980a) "Formulating and Estimating Dynamic Linear Rational Expectations Models" *Journal of Economic Dynamics and Control* **2-1** 7–46.

Hansen, L.P. Sargent, T.J. (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. Singleton, K. (1982) "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models" *Econometrica* **50** 1269–1281.

Hansen, L.P. Singleton, K. (1983) "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns" *Journal of Political Economy* **91-2** 249–265.

Hansen, L.P. Jagannathan, R. (1991) "Implications of Security Market Data for Models of Dynamic Economies" *Journal of Political Economy* **99** 225–262.

Hayashi, F. (1987) "Tests for Liquidity Constraints: A Critical Survey and Some New Observations" in Truman Bewley (ed.) *Advances in Econometrics – Fight World Congress* Econometric Society Monographs, Cambridge University Press.

Heaton, J. (1991) "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specification" manuscript, MIT Sloan School.

Heckman, J.J. (1974) "Life-Cycle Consumption and Labor Supply: An Explanation of the Relationship Between Income and Consumption Over the Life-Cycle" *American Economic Review* **64** 188–194.

Heckman, J.J. MaCurdy, T.E. (1980) "A Life-Cycle Model of Female Labor Supply" *Review of Economic Studies* **47** 47–74.

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> (1990) "Alternative Approaches to the Evaluation of Social Programs: Econometric and Experimental Methods" forthcoming in J. Laffont and C. Sims (eds.) *Advances in Econometrics: Sixth World Congress* Econometric Society Monographs, Cambridge University Press.

Heckman, J.J. (1991) "Randomization and Social Policy Evaluation" NBER Technical Working paper 107.

Hinderer, K. (1970) *Foundations of Nonstationary Dynamic Programming with Discrete Time Parameter* Springer-Verlag, Berlin.

Holt, C. Modigliani, F. Muth, J. Simon, H. (1960) *Planning, Production, Inventories, and Work Force* Prentice-Hall, Englewood Cliffs, New Jersey.

Hotz, V. J. Miller, R.A. (1991) "Conditional Choice Probabilities and the Estimation of Dynamic Models" manuscript, Carnegie-Mellon University.

Hotz, V.J. Miller, R.A. Sanders, S. Smith, J. (1992) "A Simulation Estimator for Dynamic Models of Discrete Choice" manuscript, University of Chicago.

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.

Johnson, S. Kotlikoff, L.J. Samuelson, W. (1988) "Can People Compute? An Experimental Test of the Life Cycle Consumption Model" *American Economic Review* **67** 408–498.

Jovanovic, B. (1979) "Job-Matching the the Theory of Turnover" *Journal of Political Economy* **87-5** 972–990.

Keane, M. P. Runkle, D. E. (1991) "On the Estimation of Panel Data Models with Serial Correlation when Instruments are not Strictly Exogenous" forthcoming, *Journal of Business and Economic Statistics.*

Kennet, M. (1992) "A Structural Model of Aircraft Engine Maintenance" manuscript, Tulane University.

Kocherlakota, N. R. (1990a) "On Tests of Representative Consumer Asset Pricing Models" *Journal of Monetary Economics* **26** 285–304.

Kocherlakota, N.R. (1990b) "On the 'Discount' Factor in Growth Economies" *Journal of Monetary Economics* **25** 41–57.

Krebs, J.R. Kacelnik, A. Taylor, P. (1978) "Test of Optimal by Foraging Great Tits" *Nature*, 275-7, 27-31.

Kreps, D.M. (1988) *Notes on the Theory of Choice* Underground Classics in Economics, Westview Press, Boulder. Kreps, D.M. Porteus, E.L. (1978) "Temporal Resolution of Uncertainty and Dynamic Choice Theory" *Econometrica* **46-1** 185–200.

Kreps, D.M. Porteus, E.L. (1979) "Dynamic Choice Theory and Dynamic Programming" *Econometrica* **47-1** 91–100.

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. Prescott, E.C. (1982) "Time to Build and Aggregate Fluctuations" *Econometrica* **50** 1345–1370.

Leland, H. (1974) "Optimal Growth in a Stochastic In A Stochastic Environment" *Review of Economic Studies* **41** 75–86.

Levhari, D. Srinivasan, T. (1969) "Optimal Savings Under Uncertainty" *Review of Economic Studies* **36** 153–163.

Lippman, S.A. McCall, J.J. (1976) "The Economics of Job Search: A Survey: Part I" *Economic Inquiry* **14** 155–189.

Long, J.B. Plosser, C. (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-Rohcester Conference on Public Policy, North Holland.

Lucas, R.E. Jr. (1978) "Asset Prices in an Exchange Economy" *Econometrica* **46** 1426–1446.

Lucas, R.E. Jr. Prescott, E.C. (1971) "Investment Under Uncertainty" *Econometrica* **39–5** 659–681.

Lucas, R.E. Jr. Stokey, N.L. (1984) "Optimal Growth with Many Consumers" *Journal of Economic Theory* **32-1** 139–171.

Lumsdaine, R. Stock, J. Wise, D. (1990) "Three Models of Retirement: Computational Complexity vs. Predictive Validity" NBER Working Paper 3558.

Machina, M. J. (1987) "Choice Under Uncertainty: Problems Solved and Unsolved" *Journal of Economic Perspectives* **1-1** 121–154.

MaCurdy, T.E. (1981) "An Empirical Model of Labor Supply in a Life-Cycle Setting" *Journal of Political Economy* **89** 1059–1085.

MaCurdy, T.E. (1983) "A Simple Scheme for Estimating an Intertemporal Model of Labor Supply and Consumption in the Presence of Taxes and Uncertainty" *International Economic Review* **24-2** 265–291.

MaCurdy, T.E. (1985) "Interpreting Empirical Models of Labor Supply in an Intertemporal Framework with Uncertainty" in J.J. Heckman and B. Singer (eds.) *Longitudinal Analysis of Labor Market Data* Cambridge University Press, London.

MaCurdy, T.E. (1987) "A Framework for Relating Microeconomic Evidence and Macroeconomic Evidence on Intertemporal Substitution" in T.Bewley (eds.) *Advances in Econometrics: Fifth World Congress* Econometric Society Monographs, Cambridge University Press.

Mangel, R. Clark, Colin, (1991) *Dynamic Modelling in Behavioral Ecology* Cambridge University Press.

Manski, C. (1991) "Dynamic Choice in a Social Setting: Theory and Application to Schooling Decisions" manuscript, UW-Madison.

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.

McCall, B.P. (1990) "Occupational Matching: A Test of Sorts" *Journal of Political Economy* **98-1** 45–69.

McCall, B.P. (1991) "A Dynamic Model of Occupational Choice" *Journal of Economic Dynamics and Control* **15** 387–408. McFadden, D. (1973) "Conditional Logit Analysis of Qualitative Choice Behavior" in P. Zarembka (ed.) *Frontiers of Econometrics* Academic Press, New York.

Mehra, R. Prescott, E.C. (1985) "The Equity Premium: A Puzzle" *Journal of Monetary Economics* **15** 145–161.

Mehra, R. Prescott, E.C. (1988) "The Equity Risk Premium: A Solution?' ' *Journal of Monetary Economics* **22** 133–136.

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.

Montgomery, M. (1992) "A Dynamic Model of Contraceptive Choice" forthcoming, *Econometrica.*

Montgomery, M. (1991) "Migration and Job Search in Malaysia: Estimates from Dynamic Models" manuscript, SUNY-Stony Brook.

Montgomery, M. (1990) "The Transmission of Human Capital in the Marriage Market: The Case of Singapore" manuscript, SUNY-Stony Brook.

Narendra, K.S. Wheeler, R.M Jr. (1986) "Decentralized Learning in Finite Markov Chains" *IEEE Transactions on Automatic Control* **AC-31-6**, 519–526.

Narendra,K.S. (1986) *Adaptive and Learning Systems: Theory and Applications* Plenum Press, New York.

Newey, W. West, K. (1987) "A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix" *Econometrica* **55-3** 703–708.

Pakes, A. (1986) "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks" *Econometrica* **54** 755–785.

Pakes, A. Pollard, D. (1989) "Simulation and the Asymptotics of Optimization Estimators" *Econometrica* **57-5** 1027–1057.

Pakes, A. (1991) "Estimation of Dynamic Structural Models: Problems and Prospects Part II: Mixed Continuous-Discrete Models and Market Interactions" forthcoming in C. Sims and J.J. Laffont (eds.) *Proceedings of the 6th World Congress of the Econometric Society* Barcelona, Spain, Cambridge University Press.

Phelps, E. (1962) "Accumulation of Risky Capital" *Econometrica* **30** 729–743.

Phelan, C. Rust, J. (1991) "U.S. Social Security Policy: A Dynamic Analysis of Incentives and Self-Selection" manuscript University of Wisconsin.

Pinkus, A. (1985) *n-Widths in Approximation Theory* Spinger-Verlag, Berlin.

Porteus, E. L. (1980) "Overview of Iterative Methods for Discounted Finite Markov and Semi-Markov Decision Chains" in R. Hartley et. al. (eds.) *Recent Developments in Markov Decision Processes*, Academic Press.

Powell, M.J.D. (1981) *Approximation Theory and Methods* Cambridge University Press.

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.

Resnick, S. Roy, R. (1991) "Random USC Functions, Max-Stable Processes and Continuous Choice" *The Annals of Applied Probability* **1-2** 267–292.

Reitz, T. A. (1988) "The Equity Risk Premium: A Solution" *Journal of Monetary Economics* **22** 117-131.

Rosenzweig, M. Wolpin, K. (1989) "Credit Market Constraints, Consumption Smoothing and the Accumulation of Durable Assets in Low Income Countries: Investment in Bullocks in India" manuscript, University of Minnesota.

Runkle, D. E. (1983) "Liquidity Contraints and the Permanent Income Hypothesis: Evidence from Panel Data" Ph.D. dissertation, M.I.T.

Runkle, D.E. (1991) "Liquidity Constraints and the Permanent Income Hypothesis: Evidence from Panel Data" *Journal of Monetary Economics* **27** 73–98.

Rust, J. (1985) "Stationary Equilibrium in a Market for Durable Assets" *Econometrica* **53-4** 783–806.

Rust, J. (1986) "When Is It Optimal to Kill Off the Market for Used Durable Goods?" *Econometrica* **54-1** 65–86.

Rust, J. (1987) "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher" *Econometrica* **55-5** 999–1033.

Rust, J. (1988) "Statistical Models of Discrete Choice Processes" *Transportation Research* **22B-2** 125–158.

Rust, J. (1988) "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, 359–398.

Rust, J. (1991) "Estimation of Dynamic Structural Models: Problems and Prospects Part I: Discrete Decision Processes" forthcoming 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. (1992) "Structural Estimation of Markov Decision Processes" forthcoming in D. McFadden and R. Engle (eds.) *Handbook of Econometrics* **volume 4** North Holland.

Rust, J. (1992) *Stochastic Decision Processes: Theory, Computation, and Estimation*, manuscript, University of Wisconsin.

Salant, S. (1990) "For Sale by Owner: When to Use a Broker and How to Price the House" *Journal of Real Estate Finance and Economics* **4** 157–173.

Salant, S.W. Danchick, R. (1987) "Air Force Academy Attrition: A New Perspective on the College Dropout Problem" Working Paper 87-36, University of Michigan.

Salant, S. Kalat, K. Wheatcroft, A. (1992) "Surviving Winter: A Fitnss-Based Explanation of Hoarding and Hibernation" manuscript, Department of Economics, University of Michigan.

Samuelson, P.A. (1969) "Lifetime Portfolio Selection by Dynamic Stochastic Programming" *Review of Economics and Statistics* **51** 239–246.

Sanders, S. (1991) "A Dynamic Model of Welfare Participation" manuscript, Carnegie-Mellon University.

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.

Simon, H.A. (1956) "Dynamic Programming Under Uncertainty with a Quadratic Criterion Function" *Econometrica* **24** 74-81.

Singer, I. (1970) *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces* Springer-Verlag, Berlin.

Singleton, K. (1988) "Econometric Issues in the Analysis of Equilibrium Business Cycle Models" *Journal of Monetary Economics* **21** 361–386.

Singleton, K. (1990) "Specification and Estimation of Intertemporal Asset Pricing Models" in B. Friedman and F. Hahn (eds.) *Handbook of Monetary Economics* North-Holland, 583-626.

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.

Stavins, R. Jaffe, A. (1990) "Unintended Impacts of Public Investments on Private Decisions: The Depletion of Forested Wetlands" *American Economic Review*

Stern, S. (1991) "Approximate Solutions to Stochastic Dynamic Programming Problems", manuscript, University of Virginia.

Stock, J. Wise, D. (1990) "Pensions, The Option Value of Work and Retirement" *Econometrica* **58–5**, 1151–1180.

Stokey, N.L. Lucas, R.E. Jr. (with Prescott, E.C.) (1989) *Recursive Methods in Economic Dynamics* Harvard University Press, Cambridge, Massachusetts.

Sturm, R. (1991) "A Structural Economic Model of Operating Cycle Management in European Nuclear Power Plants" manuscript, RAND Corporation.

Tauchen, G. (1990) "Solving the Stochastic Growth Model by Using Quadrature Methods and Value Function Iterations" *Journal of Business & Economic Statistics* **8-1** 49–51.

Tauchen, G. Hussey, R. (1991) "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models" *Econometrica* **59-2** 371–396.

van den Berg, G. (1990) "Structural Dynamic Analysis of Individual Labour Market Behavior" Ph.D. Dissertation, Tilberg University, Tilberg, Holland.

Van der Wall, S. (1990) *Food Hoarding in Animals* University of Chicago Press.

VanderHart, P. (1991) "Optimal Reduction of Housing Equity at the End of the Life-Cycle" manuscript, University of Wisconsin.

van der Klaauw, W. (1990) "Female Labor Supply and Marital Status Decisions: A Life-Cycle Model" manuscript, Brown University.

Wald, A. (1947) *Sequential Analysis* Wiley, New York.

Weil, P. (1989) "The Equity Risk Premium Puzzle and the Risk-Free Rate Puzzle" *Journal of Monetary Economics* **24** 401–421.

West, K.D. (1986) "A Variance Bounds Test of the Linear Quadratic Inventory Model" *Journal of Political Economy* **94-2** 374–401.

Whittle, P. (1982) *Optimization Over Time: Dynamic Programming and Stochastic Control* Volumes I and II. J. Wiley, New York.

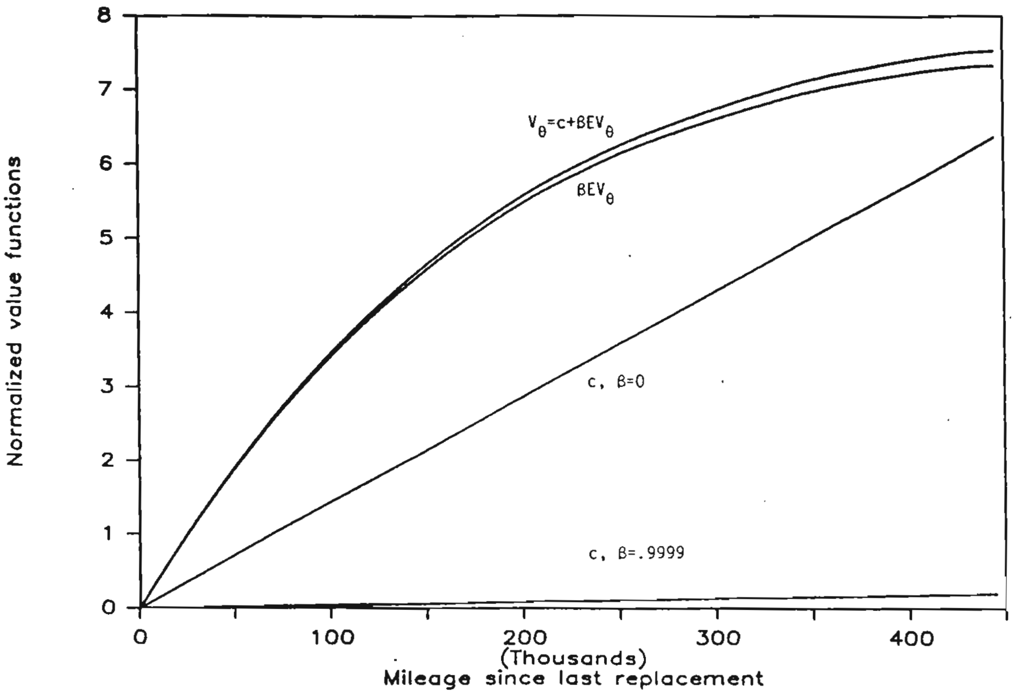
Wolpin, K. (1984) "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality" *Journal of Political Economy* **92-5** 852–874.

Wolpin, K. (1987) "Estimating a Structural Search Model: The Transition from Schooling to Work" *Econometrica* **55** 801–818.

Wolpin, K. (1989) "The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth", manuscript. UNiversity of Minnesota.

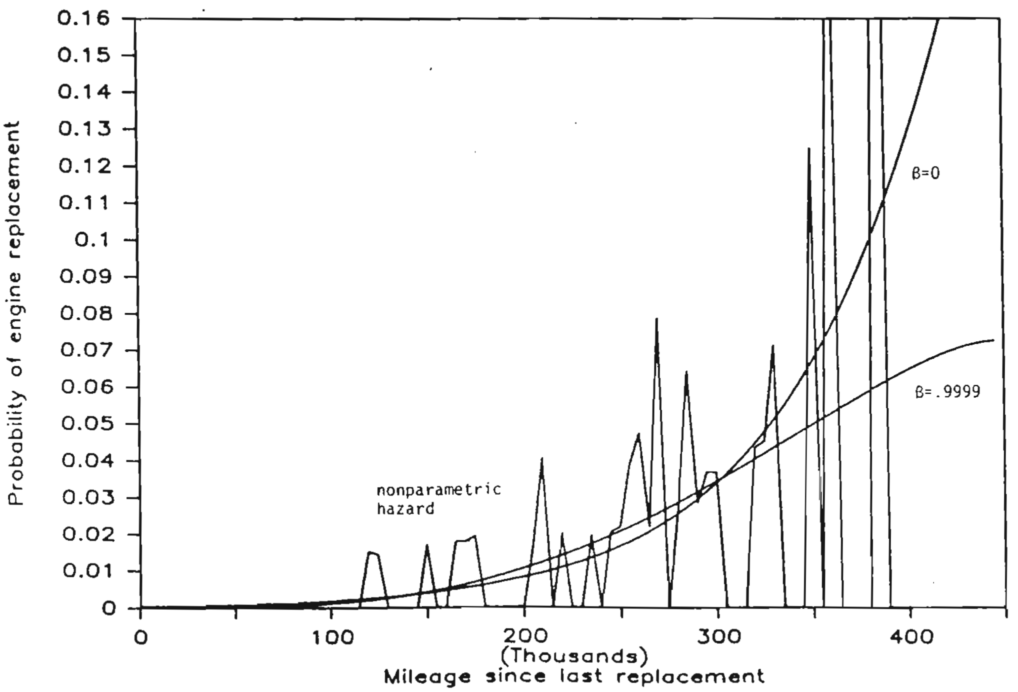
Zeldes, S. (1989) "Consumption and Liquidity Constraints: An Empirical Investigation" *Journal of Political Economy* **97-2** 305–346.

Zeldes, S. (1989) "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence" *Quarterly Journal of Economics* 275–298.



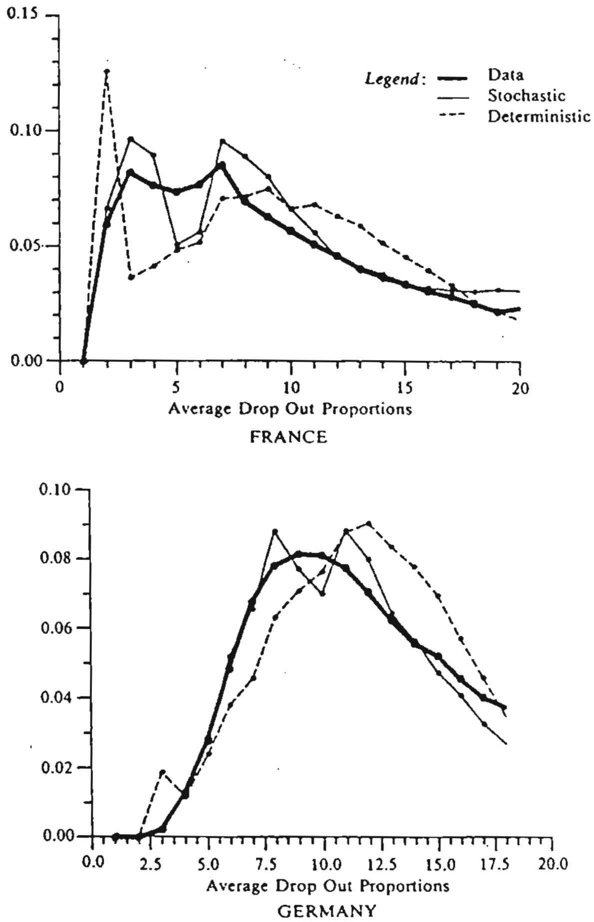
Estimated value functions for Model 11.

Figure 5.1



Estimated hazard functions for Model 11.

Figure 5.2



—Comparisons of average drop out proportions.

Figure 5.3

Predicted versus actual 1980 departure rates and implicit cumulative departures, by age: option value model (2).

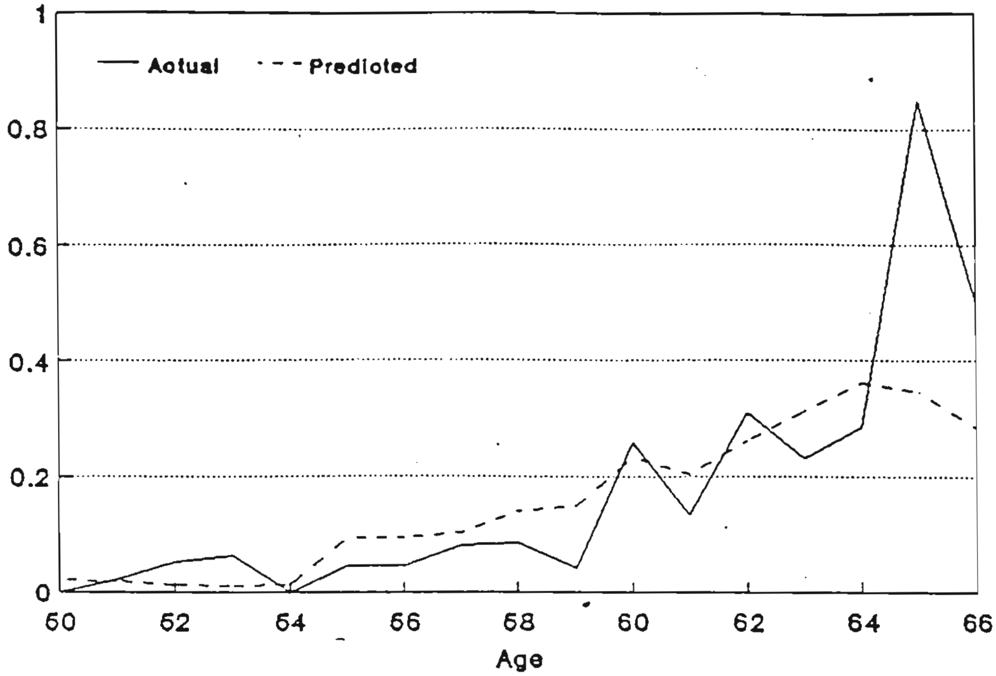


Figure 5.4

Predicted versus actual 1980 departure rates and implicit cumulative departures, dynamic programming model, by age: extreme value distribution (model 2) and normal distribution (model 4).

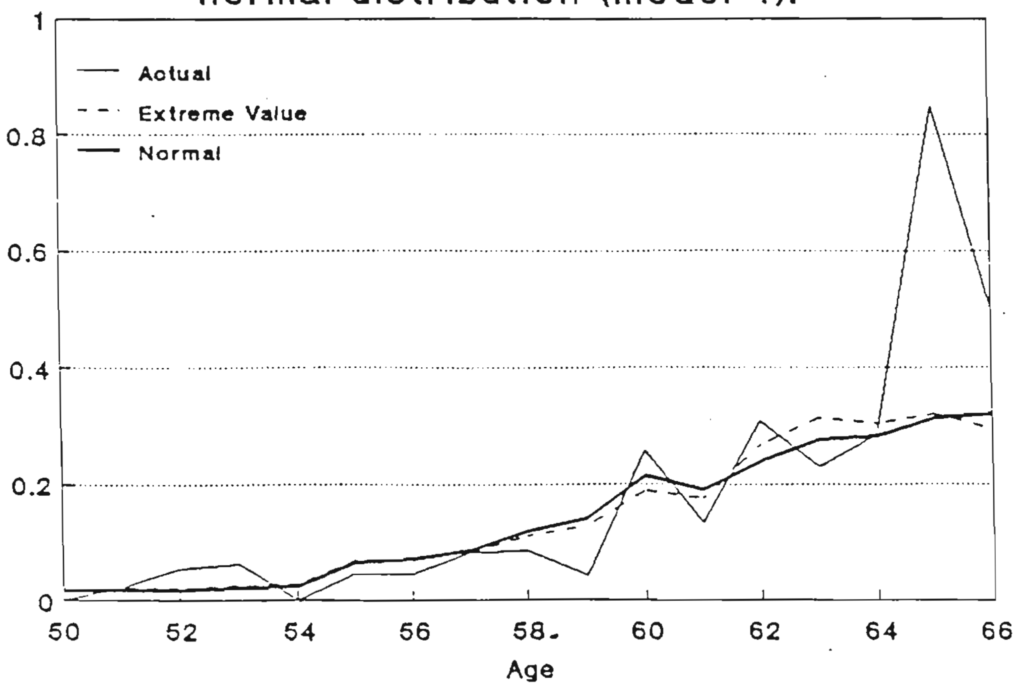


Figure 5.5

Predicted versus actual departure rates and implicit cumulative departures, by age: probit model (2).

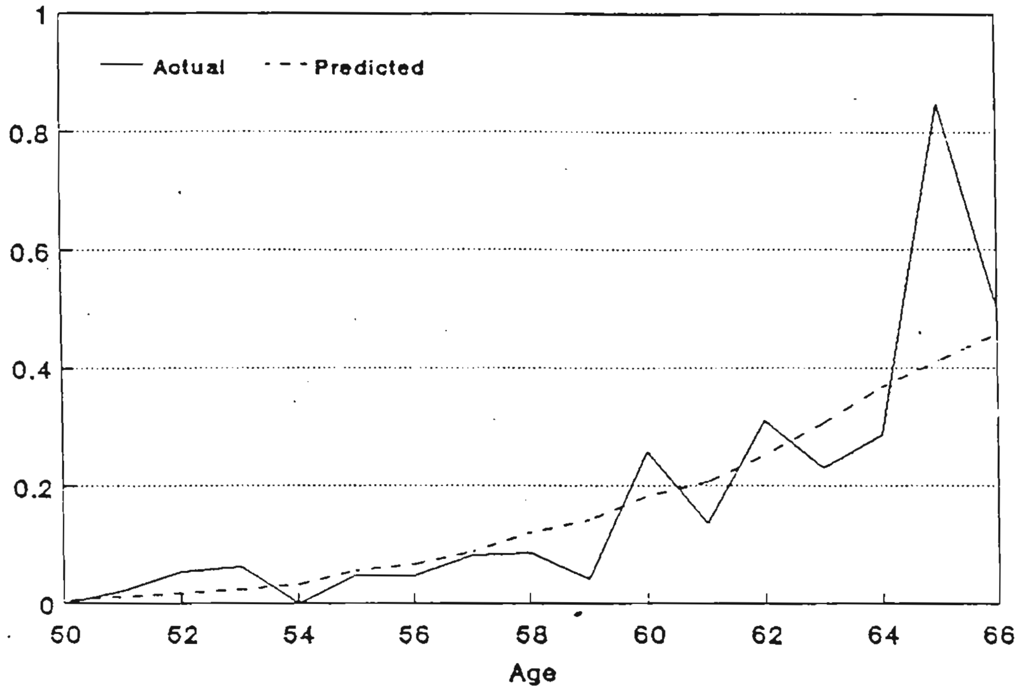


Figure 5.6

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 (2).

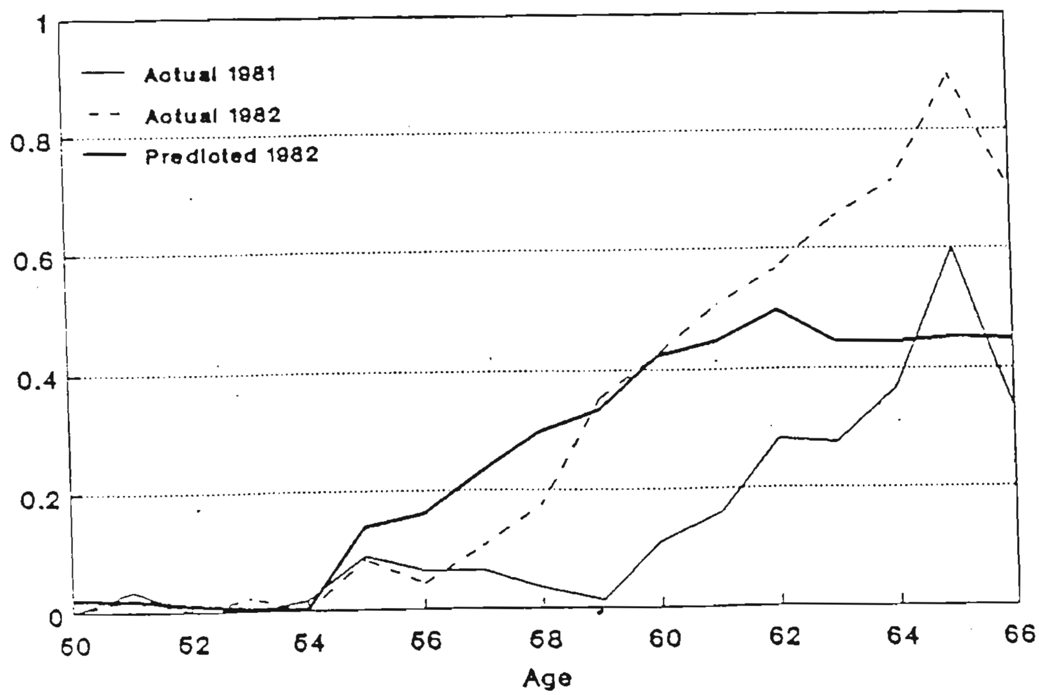


Figure 5.7

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 model 2 (extreme value distribution) and model 4 (normal distribution).

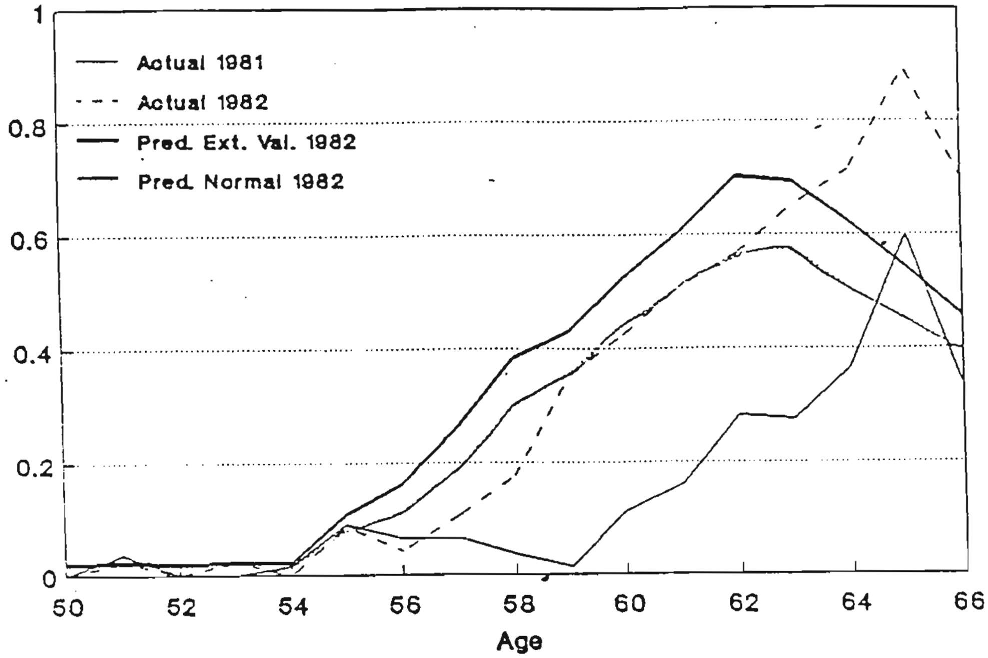


Figure 5.8

Predicted versus actual departure rates and implicit cumulative departures under the 1982 window plan, based on 1980 parameter estimates, and 1981 actual rates: probit model (2).

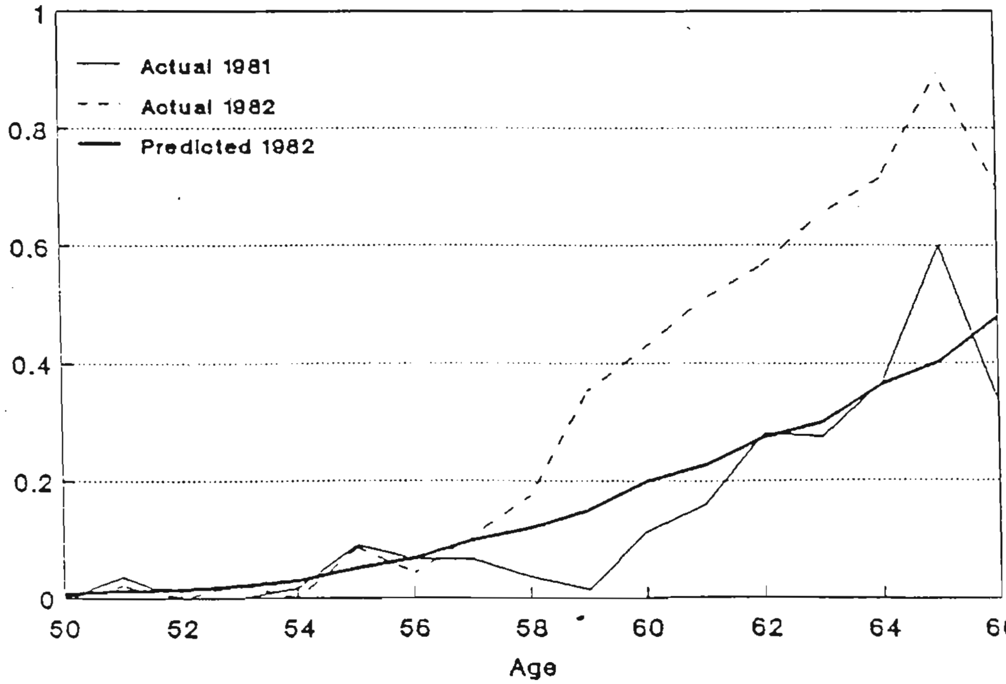


Figure 5.9

χ^2 STATISTICS AND DEGREES OF FREEDOM FOR THE COMPETING MODELS

Rating and Source of Commission	Dynamic Retention Model		ACOL		PVCOL	
	χ^2	d.f.	χ^2	d.f.	χ^2	d.f.
Pilots						
Academy	175	74	445	75	484	75
ROTC	422	132	3,632	134	1870	134
OTS/other	1132	126	16,210	128	2731	128
Navigators						
Academy	21	36	22	37	30	37
ROTC	124	104	268	106	488	106
OTS/other	378	122	318	124	715	124
Nonrated						
Academy	32	24	30	25	38	25
ROTC	1041	86	2760	88	4137	88
OTS/other	682	77	1510	79	2376	79

Table 5.1

E=FT, A=60–61, 1117 observations

	FT,DB	PT,DB	NE,DB	FT,AB	PT,AB	NE,AB
NP	70.81	4.39	0.54	14.77	6.18	3.31
DP	75.02	8.31	0.62	12.92	1.02	2.12

E=FT, A=62–63, 1131 observations

	FT,DB	PT,DB	NE,DB	FT,AB	PT,AB	NE,AB
NP	30.86	1.33	0.35	47.48	15.30	4.69
DP	31.84	3.96	0.73	46.18	12.01	5.28

E=FT, A=64–65, 254 observations

	FT,DB	PT,DB	NE,DB	FT,AB	PT,AB	NE,AB
NP	08.66	1.57	0.39	57.09	16.54	15.75
DP	29.08	3.99	1.31	44.89	13.86	6.88

E=FT, A=66–67, 18 observations

	FT,DB	PT,DB	NE,DB	FT,AB	PT,AB	NE,AB
NP	44.44	0.00	0.00	38.89	11.11	5.56
DP	27.76	4.20	1.88	43.06	14.56	8.53

Table 5.2

Parameter estimates for the option value and the dynamic programming models.

Parameter	Option Value Models (1)	Option Value Models (2)	Dynamic Programming Models: Extreme Value (1)	Extreme Value (2)	Extreme Value (3)	Normal (4)	Normal (5)	Normal (6)
θ_1	1.00*	0.612 (0.072)	1.00*	1.018 (0.045)	1.187 (0.215)	1.00*	1.187 (0.110)	1.109 (0.275)
θ_2	1.902 (0.192)	1.477 (0.445)	1.864 (0.144)	1.881 (0.185)	1.411 (0.307)	2.592 (0.100)	2.975 (0.039)	2.974 (0.374)
β	0.855 (0.046)	0.895 (0.083)	0.618 (0.048)	0.620 (0.063)	0.583 (0.105)	0.899 (0.017)	0.916 (0.013)	0.920 (0.023)
θ_3	0.168 (0.016)	0.109 (0.046)	0.306 (0.037)	0.302 (0.036)	0.392 (0.090)	0.224 (0.021)	0.202 (0.022)	0.168 (0.023)
θ_4	.	.	0.00*	0.00*	0.407 (0.138)	0.00*	0.00*	0.183 (0.243)
<u>Summary Statistics</u>								
$-\ln \mathcal{L}$	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. The option value model is described in Section II.A and the stochastic dynamic programming model is described in Section II.C. All monetary values are in $100,000 (1980 dollars). See the notes to Table 1.

*Parameter value imposed.

Table 5.3