DEPARTMENT OF MANAGEMENT

AFDELING FOR VIRKSOMHEDSLEDELSE

Working Paper 2005-6

INVESTMENT IN ADVERTISING CAMPAIGNS AND SEARCH: IDENTIFICATION AND INFERENCE IN MARKETING AND DYNAMIC PROGRAMMING MODELS

Bent J. Christensen and Nicholas M. Kiefer

UNIVERSITY OF AARHUS • DENMARK

ISSN 1398-6228

Investment in Advertising Campaigns and Search: Identification and Inference in Marketing and Dynamic Programming Models*

Bent J. Christensen
School of Economics and Management, University of Aarhus,
Building 322, DK-8000 Aarhus C, Denmark

Nicholas M. Kiefer
Department of Economics, Cornell University

December 6, 2005

Abstract

We treat identification and inference in dynamic programming models of investment in advertising campaigns, search, and marketing. Many insights relevant for general dynamic programming models are motivated by earlier work on the econometrics of the search model. Issues considered include interval identification, the curses of determinacy and degeneracy, the precise role of Bellman's equation in identification, and the dependence of parameter estimates on distributional assumptions in the random utility case.

Keywords: Search, marketing, advertising campaign, likelihood, interval identification, curse of determinacy, curse of degeneracy, random utility, functional equation, optimality principle

1 Introduction

In this paper, we consider econometric analysis of dynamic programming models with applications to investment in advertising campaigns, search, and marketing. Econometrics done as a productive enterprise deals with the interaction between economic theory and statistical analysis. Theory provides an organizing framework for looking at the world and in particular for assembling and interpreting economic data. Statistical methods provide the means of extracting interesting economic information from data. Without economic theory or statistics all that is left is an overwhelming flow of disorganized information. Thus, both theory and statistics provide methods of data reduction. The goal

*We are grateful to the Danish Social Science Research Council for research funding.

is to reduce the mass of disorganized information to a manageable size while retaining as much of the information relevant to the question being considered as possible. In economic theory, much of the reduction is done by reliance on models of optimizing agents. Another level of reduction can be achieved by considering equilibrium. Thus, many "explanations" of behavior can be ruled out and need not be analyzed explicitly if it can be shown that economic agents in the same setting can do better for themselves. In statistical analysis, the reduction is through sufficiency - if the mass of data can be decomposed so that only a portion of it contains all the relevant information, the inference problem can be reduced to analysis only of the relevant data.

Stochastic models are important in settings in which agents make choices under uncertainty, or, from a completely different point of view, in models which do not try to capture all features of agents' behavior. Stochastic models are also important for models involving measurement error or approximations. These models provide a strong link between theoretical modelling and estimation. Essentially, a stochastic model delivers a probability distribution for observables. This distribution can serve as a natural basis for inference. In static models, the assumption of optimization, in particular of expected utility maximization, has essentially become universal. Methods for studying expected utility maximization in dynamic models are more difficult, but conceptually identical, and modelling and inference methods for these models are developing rapidly. The main work horses here are the fundamental dynamic programming model and likelihood analysis.

The classic dynamic programming reference is Bellman (1957) who coined the term. This was followed by Bellman and Dreyfus (1962). Blackwell (1962, 1965) and Maitra (1968) provide the foundations for the modern approach to dynamic programming. Ross (1983) gives an accessible and brief introduction. Computational complexity increases rapidly with the dimension of the state space, leading Bellman (1957) to introduce the term 'curse of dimensionality'. Treatment of identification and inference is a relatively new area. Key contributions are made by Rust (1987). A recent discussion of the curse of dimensionality with approaches to breaking the curse is Rust (1997).

An early application of dynamic programming in economics is the sequential job search model, which is due to Mortensen (1970) and McCall (1970). See also Lippman and McCall (1976a, 1976b) and Mortensen (1986). Mortensen's work has extended the range of applications of the search model from unemployment to labor turnover, research and development, personal relationships, and labor reallocation. His insight, that friction is equivalent to the random arrival of trading partners, generates immediately a stochastic model that lends itself to structural likelihood analysis. This is a key instance of fruitful data reduction through the marriage of statistical analysis and theory of optimizing agents. Kiefer and Neumann (1979) and Christensen and Kiefer (1991) are contributions in this direction. Mortensen (1990) and Burdett and Mortensen (1998) initiated the development of equilibrium dynamic models designed to account for wage dispersion and the time series behaviour of job and worker flows. This leads to the second level of data reduction, obtained by imposing equilibrium conditions on the statistical framework for optimizing agents. Empirical analysis and inference issues are studied by Kiefer and Neumann (1993), van den Berg and Ridder (1993, 1998), Christensen and Kiefer (1997), Bunzel et al. (2001), Christensen et al. (2005) and many others.

In this paper, we treat identification and inference issues in dynamic programming models. To see the issues simply and clearly, we present them in simple discrete state/control settings. A main example running through the sections is a simple marketing model. The study of optimal advertising policy in dynamic models where current investment in advertising affects future demand was pioneered by Nerlove and Arrow (1962). An empirical investigation in the dynamic duopoly case is presented by Chintagunta and Vilcassim (1992). We draw attention to several issues that apply in wide generality in dynamic programming models but have received little or no notice previously, such as identification of parameters only up to intervals, rapid information accumulation, the additional curses of determinacy and degeneracy stemming from the fact that the model predicts a deterministic relation between the state and control, the subtle differences between measurement error, imperfect control and random utility approaches, and the dependence of the parameter estimates on the assumed shock distribution in the random utility case. Many insights derive from earlier work on the econometrics of the search model, and we review some of the relations between the motivating studies of the search model and the implications for general dynamic programming models. That relations exist is natural, of course, since components in one model frequently have equivalents in other models, e.g., unemployment in the search model corresponds to inventory in the marketing model, etc. A fuller account of all these and further general economic modelling and inference issues is provided by Christensen and Kiefer (2005).

The rest of the paper is laid out as follows. In Section 2, we introduce the fundamental dynamic programming model and our marketing example. Section 3 considers discrete states and controls and provides illustrative computations. Section 4 is a preview of important identification issues. Likelihood functions are introduced in Section 5, and identification and inference treated in detail. Measurement error is introduced in Section 6. Section 7 considers the case of imperfect control. Section 8 introduces random utility models, and the case of continuously distributed shocks is treated in Section 9. The connections between the results for general dynamic programming models and the motivating work on the econometrics of the job search model are reviewed in Section 10, and Section 11 concludes.

2 Dynamic Programming: The Marketing Example

The basic components of a dynamic optimization model are the objective function, the state variables, the control variables with any associated constraints, and the transition distribution giving the evolution of states as a stochastic function of the sequence of states and controls. These components are illustrated here in a simple marketing example.

Consider a firm deciding whether to invest in a marketing campaign. Suppose that there are two states of demand, high and low. The marketing decision is to run the campaign or not. Profit in any period is given by demand less costs including marketing costs. Since there are only four configurations of states (demand) and controls (the marketing decision) the profit function is given by four numbers, $u(x, c), (x, c) \in \{0, 1\}^2$. Low demand is indicated by $x = 0$ and marketing by $c = 1$. While simple, our discrete model relates naturally to continuous state/control dynamic models of investment in advertising campaigns, e.g. Chintagunta end Jain (1992). A plausible profit function might be the one in Table 1.

Table 1. Profit Function

	$u(x,c)$	
	$c = 0$	$c = 1$
$x = 0$	7	4
$x = 1$	11	7

Here, the low demand state with no marketing generates the same profit as the high demand state with marketing. The objective function of the firm is to maximize the expected present discounted value of profits $E\Sigma_{t=0}^{T}\beta^t u(x_t, c_t)$ where T is a "horizon" which may be infinite and $\beta \in [0, 1)$ is a discount factor. Dynamics are incorporated in the model by letting the probability distribution of demand next period depend on the marketing decision this period. If the advertising campaign is effective, then the probability that $x_{t+1} = 1$ is greater when $c_t = 1$ than when $c_t = 0$.

A simple case has the distribution of x_{t+1} depending only on c_t and not on the current state. For example, consider the particularly simple case $p(x_{t+1} = 1|c_t = 1) = 1$ and $p(x_{t+1} = 1|c_t = 0) = 0$. Thus the state of demand in period $t + 1$ is determined exactly by the marketing effort in period t. The logic of dynamic programming can be illustrated by considering the 2-period problem. In the last period, clearly there is no benefit from marketing and $c = 0$ is optimal no matter what the level of demand. Now consider the first period. Here there is a tradeoff between current period profits (maximized by $c = 0$) and future profits (maximized by $c = 1$). If the current state of demand is low ($x = 0$) then the current cost of marketing is 3 and the current period value of the gain next period as a result of marketing is 4β. Thus for $\beta > 3/4$ it is optimal to invest in marketing in the first period when the state of demand is low. Suppose the state of demand in the first period is high. Then the marketing effort costs 4 in current profit, and gains only 4β in current value of future profits. Since $\beta < 1$, this effort is not worthwhile. We have found the optimal policy. The logic of beginning in the last period in finite horizon models and working backwards in time is known as "backward recursion" and is generally the way these problems are solved.

A more general and plausible specification would have the distribution of next period's demand depend on both the marketing decision and the current state of demand. As a matter of interpretation, this allows for marketing to have lasting effects (though tapering off over time). This possibility is pursued further below.

3 Discrete states and controls

The discrete state/control dynamic programming model has many applications, can be treated with fairly elementary methods, and allows illustration of most of the important issues of identification and estimation that appear in more general settings. The simple case allows focus on issues of substance rather than details. The latter are important, but can be better handled once the substantive issues are identified. Note that the discrete case is in many ways quite general. For practical purposes machine calculations are discrete, as are data, and indeed applications of continuous models often (but not always) require explicit discretization.

In the finite horizon case it is conventional to index value functions by the number of periods left, rather than distance from the current period (0). Thus, in a T-period optimization, the value function at the outset is

$$V_{T-1}(x) = \max_{\pi} E_{T-1,\pi} \Sigma_{t=0}^{T-1} \beta^t u(x_t, c_t) \tag{1}$$

where $\pi = (\pi_{T-1}, \pi_{T-2}, ..., \pi_0)$ is a sequence of policy functions. A policy function at period t maps the current state x_t and all previous states into the current policy c_t. Thus, we are maximizing over a sequence of functions. The expectation operator is a little more subtle - it is a conditional expectation conditioning on the current value of the state variable (hence the $T-1$ subscript) and on the policy π. Since the transition distribution typically depends on controls, the expectation clearly depends on the policy. The value function in the final period is $V_0(x)$. With $u(x, c)$ the immediate reward from using control value c with state x, and C the set of admissible controls, the final period value function is clearly $V_0(x) = \max_{c \in C} u(x, c)$. The value function with one period left $V_1(x)$ is just the maximized value of the current reward and the discounted expected value of future rewards. But the future reward is $V_0(x)$, so the function $V_1(x)$ is given by

$$V_1(x) = \max_{c \in C} \left\{ u(x, c) + \beta E_1 V_0(x') \right\}. \tag{2}$$

Here, x' is the next period state - a random variable whose distribution is determined by c and x, current controls and state. Iterating, we obtain the whole sequence of value functions $V_0, ..., V_{T-1}$ with

$$V_t(x) = \max_{c \in C} \{u(x, c) + \beta E_t V_{t-1}(x')\}. \tag{3}$$

This recursion is known as *the optimality principle*, or *Bellman's equation.*

The calculations above are simple and can be done straightforwardly using, for example, a spreadsheet program. We illustrate using the marketing model of Section 2. Here, the state variable is a state of demand, $x \in \{0, 1\}$, and the control variable is an advertising decision $c \in \{0, 1\}$. The profit function is given in Table 1.

A heuristic backward recursion argument together with a trivial transition distribution was used to obtain the solution for the 2-period problem in Section 2. Here, we make the problem a little more interesting and realistic by specifying the transition distribution in Table 2.

Table 2
Transition Probabilities

$\mathbf{p}(x_{t+1} = 1 | x_t, c_t)$

	$c = 0$	$c = 1$
$x_t = 0$	0.1	0.85
$x_t = 1$	0.5	0.85

Thus, when demand is low it is unlikely that demand will increase without the marketing effort; when demand is high and marketing efforts are 0 the next period's demand states are equiprobable; in either case marketing improves the probability of next period's demand being high to 0.85. Discounting the future at $\beta = 0.75$ we calculate the value functions and optimal policies by backward recursion for the 1 through 10-period problems and report results in Table 3.

Table 3. Value Functions and Optimal Policies

t	$V_t(0)$	$V_t(1)$	$c_t(0)$	$c_t(1)$
0	7.00000	11.00000	0	0
1	12.55000	17.75000	0	0
2	16.80250	22.36250	0	0
3	20.14638	25.68687	1	0
4	22.64185	28.18746	1	0
5	24.51672	30.06099	1	0
6	25.92201	31.46664	1	0
7	26.97621	32.52074	1	0
8	27.76680	33.31135	1	0
9	28.35976	33.90430	1	0

In the one, two and three period problems it is never optimal to run the marketing campaign. In the 4 period problem, it is optimal to run the marketing program in the initial period if the state of demand is low. This makes sense; it is optimal to try to shock the system into the fairly persistent high-demand state, then abandon the marketing investment. In the longer problems, it is optimal to run the marketing program if demand is low and the remaining horizon is greater than 3 periods. It is never optimal to run the marketing program during the high-demand periods.

Define the operator $\mathcal{T}$ by

$$\mathcal{T}f = \max_{c \in C} \{u(c, x) + \beta E f(x')\} \tag{4}$$

in the infinite horizon stationary case and note that $\mathcal{T}$ is not subscripted with t. Then the optimality equation can be written $V = \mathcal{T}V$ and V is the unique bounded solution to this functional equation. The value function, V solving the optimality equation can be computed as the limit of T-period finite-horizon value functions as $T \to \infty$. Let $V_0(x) = \max_{c \in C} u(x, c)$, the value in the one-period problem. Note that $V_1 = \max\{u(x, c) + \beta E V_0\} = \mathcal{T}V_0$ is the value function for the 2 period problem, where $\mathcal{T}$ is the operator defined in (4). Then

$$\begin{aligned} V_T(x) &= \max_{c \in C} \{u(x, c) + \beta E V_{T-1}(x)\} \\ &= \mathcal{T}V_{T-1} = \mathcal{T}^{T-1} V_0 \end{aligned} \tag{5}$$

has the interpretation of the value of the $T+1$ period problem with final period state-dependent reward $V_0(x)$. We have

$$V_T(x) - V(x) \to 0 \text{ as } T \to \infty, \tag{6}$$

since $\mathcal{T}(\cdot)$ satisfies Blackwell's (1965) conditions of monotonicity and discounting and hence is a contraction (see also Stokey and Lucas (1989)).

To illustrate we consider Table 3 giving the value function in the marketing model for the $T = 1, ..., 10$ period problems and add results in Table 4 for the 45 through 50 period problem.

Table 4. Value Function Iterations

t	$V(0)$	$V(1)$	$c(0)$	$c(1)$
45	30.13856	35.68311	1	0
46	30.13857	35.68313	1	0
47	30.13858	35.68314	1	0
48	30.13859	35.68314	1	0
49	30.13860	35.68315	1	0

Here, the value function iterations are identical to 6 digits - the improvements in the calculations are only relative changes of order 10^{-7}. Note, however, that the policy function appears to have converged much more rapidly. This is frequently the case in discrete state/control models. It is much easier to determine a map from one finite set to another than it is to determine the exact value of a real vector. Notice, on a computer the value can only be determined to a certain level of precision.

4 Identification: A Preview

A central identification issue can be easily illustrated in the context of the marketing model of Sections 2 and 3. The issue will be pursued in detail in Section 5. Parameters are split into two groups: those that relate to the transition distribution at the optimal policy, and those composed of utility, discount factor, and transition probabilities for non-observed transitions. The basic result is that a discrete state/ discrete control model can only determine parameters within certain ranges. More information is required in order to determine some of the economically interesting parameters. The information is sometimes introduced in specification, and we will consider this possibility, too.

Consider the data sequence $\{x_t, c_t\}_{t=0}^{T}$ generated by the optimal policy in the marketing model. The data can be summarized in two tables reflecting the within-period information on the control rule and the intertemporal information on the transitions (see Table 5). Here, $n\,(x=j)$ is the number of time periods t the state variable x_t is observed in state j, and $n_x\,(jk)$ is the number of observed transitions from state j to k.

Table 5. Data Configuration

Policy Information		
	$c=0$	$c=1$
$x=0$	0	$n\,(x=0)$
$x=1$	$n\,(x=1)$	0

Transition Information		
	$x_{t+1}=0$	$x_{t+1}=1$
$x_t=0$	$n_x\,(00)$	$n_x\,(01)$
$x_t=1$	$n_x\,(10)$	$n_x\,(11)$

The transition distribution can clearly be estimated at the optimal policy. Note that the components of the transition distribution at other values of the policy contribute to determination of the optimal policy. Information on these transition probabilities is available only through observation of the optimal policy and any restrictions implied by the functional equation, which depends on alternative transition probabilities and characterizes the optimal policy. This relationship is subtle and we will set it aside for purposes of this example by simply assuming that the transition distribution is known.

Returns $u(x, c)$ may also be observed, in which case the components of the return function corresponding to the optimal policy can be estimated. The other components of the return function, corresponding to state/control configurations never observed, can be identified only through restrictions implied by the optimality equation. Again, the functional equation (4) (or (3), in the finite horizon case) depends on alternative rewards and characterizes the optimal policy. We will set this aside as well by assuming that the reward function is known.

The only remaining unknown parameter is the discount factor β. Suppose we observe the entire infinite sequence $\{x_t, c_t\}_{t=0}^{\infty}$ so that any parameter that can be estimated consistently is known. What can be said about β? Not much. We have seen that the policy $c(0) = 1; c(1) = 0$ is optimal for $\beta = 0.75$. It is obviously optimal for any larger value of β. What about smaller values? It turns out that this policy is optimal for $\beta \geq 0.7143$ (approximately) and the policy of $c(0) = c(1) = 0$ is optimal for smaller values of the discount factor. Thus, the most we can expect the data to tell us is whether $\beta \geq 0.7143$ or not. Of course, β is a continuous parameter with parameter space [0,1), but all we will be able to tell from the data is which of the sets $\{[0,0.7143),[0.7143,1)\}$ the parameter β is in.

This feature of finite state/finite control dynamic programming models, that continuous parameters are identified only up to ranges in the parameter space, is ubiquitous and it is not specific to our example. The situation is not completely hopeless. Of course, continuous parameters corresponding to transition probabilities at the observed policy are typically identified. Similarly, if rewards are observed, rewards corresponding to state/control combinations given by the optimal policy are typically identified. It is the parameters which must be identified through the restriction imposed by the functional equation that are typically underidentified.

Note that the above analysis applies to observations generated by a single decision maker over time, or to panel data on many decision makers following the optimal policy. Panel data do not help.

Observations on the finite horizon control policy are more informative. Here, the policy is not stationary and information on how the horizon changes the policy is informative. However, identification is not possible - instead we increase the number of intervals to which we can assign β. With $\beta = 0.75$, we saw that the policy $c(0) = 1$ and $c(1) = 0$ is optimal if the horizon is 4 periods or longer. At $\beta = 0.72$ that policy is optimal if the horizon is 5 periods or longer. At $\beta = 0.715$, optimality requires 7 periods or longer, at 0.7145 it requires 8 periods or longer, etc. Thus, if we observe panel data on decision makers and we know their horizons we can isolate the horizon at which the policy shifts. In this case it is possible to use the optimality equation to narrow the range of possible β consistent with the policy. But it remains impossible to identify β more closely than an interval.

5 Likelihood Functions

We begin with the likelihood function in the simplest dynamic programming case: A single binary state and a single binary control variable in a stationary infinite horizon problem. This case, while simple, illustrates properties which are general. In the case of accurate observations, the control is a deterministic function of the state, and hence the control policy can be taken as known after a small sample is realized (specifically, after both possible values of the state are realized). Of course, in many applications this setup would be unrealistic, in part because the model is just a model and is not purporting to be an exact description of the world, but it is a useful starting point for considering identification issues. In a sense, this is a situation of maximal information. If a parameter is not identified in this setting, it is hard to argue that additional data information somehow appears when the setting is generalized.

After treating the simple case, we show that the notation and techniques extend to the general discrete model. Then, we consider extensions allowing measurement error, imperfect control and random utility. Identification issues are treated in detail. Parameters are split into two groups: those that relate to the transition distribution at the optimal policy, and those composed of utility, discount factor, and transition probabilities for non-observed transitions. Continuous parameters apart from the transition distribution parameters are typically unidentified. They are only restricted to lie in certain ranges in the parameter space, even asymptotically. The results from our marketing model in Sections 2 to 4 are general.

We focus on the infinite horizon problem

$$V(x) = \max_{\pi} E\Sigma_{t=0}^{\infty}\beta^t u(x_t, c_t) \tag{7}$$

where the expectation is over a Markov transition distribution $p(x_t|x_{t-1}, c_{t-1})$ and hence the optimal policy $(c(x), c(x), ...)$ is stationary, or equivalently

$$V(x) = \max_{c\in C} \{u(x, c) + \beta EV(x')\}. \tag{8}$$

The observables are the state sequence $\{x_r\}_{r=0}^T$ and the control sequence $\{c_r\}_{r=0}^T$, with x_r and c_r in $\{0,1\}$. We assume that the reward sequence is not observed (this will be treated later). The state transition probabilities are state and control dependent. Thus, we have Table 6.

Table 6. Transition Probabilities

	$c_t = 0$		$c_t = 1$	
$x_t = 0$	$p_{00}(0)$	$p_{01}(0)$	$p_{00}(1)$	$p_{01}(1)$
$x_t = 1$	$p_{10}(0)$	$p_{11}(0)$	$p_{10}(1)$	$p_{11}(1)$

In Table 6, $p_{ab}(c)$ is the probability of a transition from a to b when the period t control is c. Given the adding up constraints $p_{a0}(c) + p_{a1}(c) = 1$, there are four transition probabilities in the model. The policy function is a pair $(c(0), c(1)) \in \{0,1\} \times \{0,1\} = \{00, 01, 10, 11\}$. There are four possible policy functions, of which only two are interesting since in the other cases the policy does not depend on the states.

Given the period 0 values of the state and control, the period 1 distribution of the observables conditional on parameter vector θ is $p(x_1, c_1|x_0, c_0, \theta)$. Now, this is a singular distribution, in that c_t is a deterministic function of x_t for any t and in particular for $t = 0$ and 1. This important point has several implications. First, it suffices to condition on one of x or c; we choose to condition on x. Second, the distribution of x and c given parameters is completely described by the distribution of x alone. Nevertheless, the information given by c in addition to x in the likelihood function is enormous. In fact, the deterministic relation between c and x is learned with certainty as soon as the different values of x are seen in the data. Thus, parameters that enter only through the deterministic relation between c and x are learned, if they are identified, rapidly.

We proceed by conditioning. Let x, x' and c, c' refer to current and one-period ahead values of the state and control, respectively. For the present, we suppress dependence on the parameter vector θ. Then

$$\begin{aligned} p(x', c'|x, c) &= p(x'|x, c)p(c'|x', x, c) \\ &= p(x'|x)p(c'|x', x) \\ &= p(x'|x)p(c'|x'), \end{aligned} \tag{9}$$

since there is no point in conditioning on c as well as x, and since c' is a deterministic function of x' and hence x is irrelevant for c' given x'. The second factor is just $I\,(c' = c(x', \theta))$ where I is the indicator function and θ has been reinserted here for emphasis: Given θ, the policy function can be calculated and this factor of the distribution easily evaluated. The likelihood is

$$\begin{aligned} L(\theta, x_r, c_r, r &= 1, ..., T) = \Pi_{r=1}^{T}\, p(x_r|x_{r-1}, \theta)p(c_r|x_r, \theta) \\ &= \Pi_{r=1}^{T} p(x_r|x_{r-1}, \theta)\Pi_{r=1}^{T} p(c_r|x_r, \theta). \end{aligned} \tag{10}$$

Since the second factor is zero for values of θ inconsistent with the data and the first term is positive, only values consistent with the data have positive likelihood. As a practical matter, it is useful to check early on whether there exists any parameter value consistent with the data. In many cases, using this simple specification, there will not be, and hence the model is rejected and must be modified. However, this is a good place to begin the study of identification.

Suppose we have a sample of size T, indexed by subscripts $r, s, t \in \{0, ..., T-1\}$. Consider the state sequence $\{x_t\}_{t=0}^{T-1}$. There are 2^T such sequences. Arrange these sequences in lexicographic order and index these by $i, j, k \in \{0, ..., 2^T - 1\}$. Then the ith sequence is the binary expansion of i. This is a convenient way of thinking about the problem. Our random variable is now i, the position of the realized sequence. Table 7 illustrates for the case $T = 3$.

Table 7. State transitions

i	sequence
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

Elements of the sequences are indexed by r, s, t. Let $\xi(i, s)$ be the sth digit in the binary expansion of i. Let $\kappa(i, r) = \xi(i, r)||\xi(i, r + 1)$ be the rth pair of digits in the binary expansion of i.

The parameters entering $p(x'|x)$ are the transition probabilities corresponding to the optimal policy. We write these $p_{ab}, a, b \in \{0, 1\}$ without an argument to select the appropriate pieces from Table 6. Thus, there are two probabilities to be estimated. Of course, there could be restrictions relating these probabilities. Further, knowledge of the optimal policy could restrict the values of these probabilities (for example if these were the only parameters and the optimal policy was known). However, the restricted estimators can usually be written as functions of the unrestricted, and the unrestricted maximum likelihood estimators (MLEs) are easy to obtain. We will set aside consideration of these issues for the moment. The likelihood function corresponding to the first factor in (10) is

$$p(i|x_0) = \Pi_{r=1}^{T-1} p_{\kappa(i,r)}. \tag{11}$$

Introducing the notation $N\kappa(i, ab) = \Sigma_{r=1}^{T-1} I(\kappa(i, r) = ab)$, the number of ab pairs in the ith sequence, we have

$$p(i|x_0) = \Pi_{a,b} p_{ab}^{N\kappa(i,ab)}, a, b \in \{0, 1\}. \tag{12}$$

Taking logarithms and writing $p_{a1} = 1 - p_{a0}$ yields

$$\begin{aligned} l(p_{00}, p_{10}|i, x_0) &= N\kappa(i, 00) ln p_{00} + N\kappa(i, 01) ln(1 - p_{00}) \\ &\quad + N\kappa(i, 10) ln p_{10} + N\kappa(i, 11) ln(1 - p_{10}) \end{aligned} \tag{13}$$

and the maximum likelihood estimators

$$\widehat{p}_{a0} = N\kappa(i, a0\ )/(N\kappa(i, a0) + N\kappa(i, a1)), a \in \{0, 1\}. \tag{14}$$

The step of taking logarithms is justified only where the probability is positive, that is for values of p_{a0} consistent with the control rule. This important point is illustrated below.

Consider the marketing model of Sections 2 through 4. This is a 2-state, 2-control problem. In Sections 2 and 3 the focus was on solving the dynamic program; here we consider estimation. First, suppose the only unknown parameter is the transition probability p_{00}. Suppose we have a sample of length T, and suppose in this sample each value of x is observed at least once. Then the control rule $c(x) = (c(0), c(1))$ is known. In evaluating the likelihood, only values of the parameter p_{00} consistent with the observed state-control sequence and the hypothesis of optimization are considered (the likelihood is zero for other values of the parameter). How does knowledge of the control rule constrain the value of p_{00}? With the utility function and transition distributions given in Table 8

Table 8. Marketing Model

Profit

	$c = 0$	$c = 1$
$x = 0$	7	4
$x = 1$	11	7

$\mathbf{p}\left(x_{t+1} = 1|x_t, c_t\right)$

	$c_t = 0$	$c_t = 1$
$x_t = 0$	0.1	0.85
$x_t = 1$	0.5	0.85

and with discount factor .75 we found that the optimal policy in the infinite horizon problem is $c(x) = (c(0), c(1)) = (1, 0)$. Here, the value of p_{00}, the $0-0$ transition probability corresponding to $c = 1$, is 1-0.85 = 0.15. Clearly, there are other values of this transition probability for which $c = (1, 0)$ is also optimal. For example, any smaller value would leave the optimal policy unchanged - the purpose of the marketing campaign is to shock the system into the high demand state, so if this becomes easier, it must still be optimal. In other words, the value associated with running the campaign is increased for higher 0-1 transition probability (and hence lower 0-0, our parameter). Hence if it is optimal to run the campaign for $p_{00} = 0.15$, it must also be optimal for any smaller value. In fact, $c = (1, 0)$ is optimal for $p_{00} \leq\ 0.20 = r$.

Thus, once the control rule is known to be $c = (1, 0)$ (i.e., after each value of the state variable has been seen), the information contained in the likelihood factor corresponding to $\Pi_t p(c_t|x_t)$ is exactly that $p_{00} \in [0, r]$. This information is in one sense extremely precise. It is accumulated quickly (this is not a $T \to \infty$ result) and it completely rules out a portion of the natural parameter space [0,1]. It is in another sense extremely imprecise. The exact value of the parameter p_{00} cannot be estimated on the basis of the information in the control rule - only bounded into an interval. There is no more information in the relevant likelihood factor, even as $T \to \infty$. The situation is identical to that discussed in Section 4 relative to estimation of the discount factor β. The difference here is that there is additional information on p_{00} available through the first factor in the likelihood function, corresponding to transition information. We turn now to this factor,

$$L(p_{00}|i, x_0) = N\kappa(i, 00) ln p_{00} + N\kappa(i, 01) ln(1 - p_{00}), \tag{15}$$

which can be maximized subject to the constraint $p_{00} \leq r$. The Lagrangian is

$$L(p_{00}|i, x_0) = N\kappa(i, 00)lnp_{00} + N\kappa(i, 01)ln(1 - p_{00}) + \lambda(r - p_{00}), \qquad (16)$$

where λ is a Lagrange multiplier. The Kuhn-Tucker conditions require

$$\begin{aligned} &N\kappa(i, 00)/p_{00} - N\kappa(i, 01)/(1 - p_{00}) - \lambda = 0 \\ &r - p_{00} \geq 0 \\ &\lambda(r - p_{00}) = 0 \end{aligned} \qquad (17)$$

and the solution possibilities are clearly either $\widehat{\lambda} = 0$ and $\widehat{p}_{cmle} = \widehat{p}_{ml}$ or $\widehat{p}_{cmle} = r$ and $\widehat{\lambda} = N\kappa(i, 00)/r - N\kappa(i, 01)/(1 - r)$. A significant nonzero value for $\widehat{\lambda}$ indicates that the transition information is inconsistent with the control information, and that the model is therefore likely misspecified. This point is pursued later. For the present, assume that the unconstrained MLE satisfies the constraint.

In Section 4 we saw informally that the discount factor was not identified by knowledge of the control rule even when all other parameters including the transition probabilities were known. The parameter was bounded to an interval. In fact, the transition information is informative on the discount factor β, in that the boundary of the interval containing feasible estimates of β depends on the transition probabilities. We consider the case $\theta = (p_{00}, \beta)$; unknown discount factor and transition probability. Here the natural parameter space is $\Theta = [0, 1] \times [0, 1)$. As soon as the control rule is observed, this can be narrowed. Specifically, in our example with the observed control rule $c = (1, 0)$, the parameter space consistent with the observed control rule is $A \subset \Theta$, illustrated in Figure 1.

This is all of the parameter information contained in the control rule.

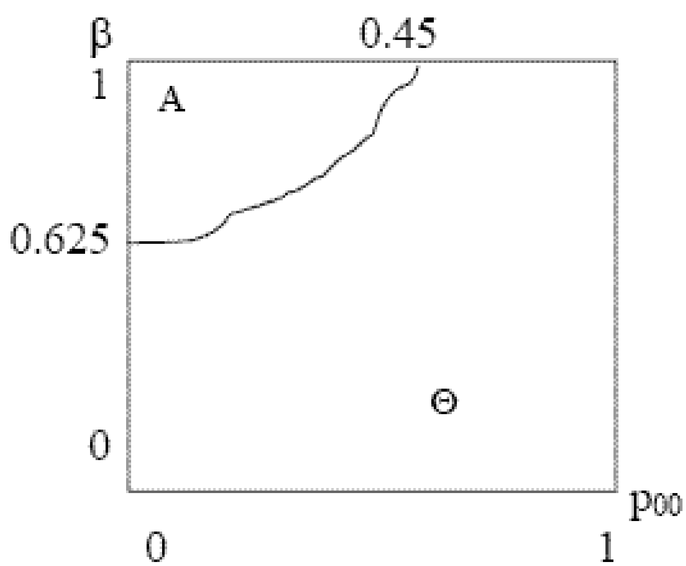
Turning now to the transition information on the two parameters, the log-likelihood is

$$L(p_{00}, \beta|i, x_0) = N\kappa(i, 00)lnp_{00} + N\kappa(i, 01)ln(1 - p_{00}), \qquad (18)$$

to be maximized subject to the constraint that $\theta \in A$. Note that this portion of the loglikelihood does not depend on the discount factor β at all. Consequently the constraint is once again in the form $p_{00} \leq r$ and can be imposed as above. Once again, the estimator will satisfy the constraint asymptotically if the model is well-specified, and if the unconstrained estimator does not satisfy the constraint, the significance of the estimate of the Lagrange multiplier can be used as the basis of a specification test. Turning to the unconstrained estimator $\widehat{p}_{00}$, we note that the transition information is informative on the discount factor β in that the interval in which β can lie and still be consistent with the known control rule depends on the value of $\widehat{p}_{00}$.

Can a general result be obtained in the simple case with the control rule known? Suppose there are K state variables, the ith taking values in the discrete set $\mathbf{H}_i$ with cardinality $|\mathbf{H}_i| = H_i$ and C control variables, the jth taking values

Figure 1. Parameter Space



in the discrete set $\mathbf{J}_j$ with cardinality $|\mathbf{J}_j| = J_j$. The problem can be rewritten as a single discrete state/discrete control model with the single state variable x taking $H = \Pi_i H_i$ values from the set $\mathbf{H}$ and the single control taking $J = \Pi_j J_j$ values from the set $\mathbf{J}$. The control rule is a map $c : \mathbf{H} \rightarrow \mathbf{J}$, a point in $\mathbf{L} = \mathbf{J}^H$, a finite set. Now write the control rule c as a function of the parameter θ, $c(x;\theta)$, where $\theta \in \Theta \subseteq R^k$. Regarded as a function of θ we have $c : \Theta \rightarrow \mathbf{L}$. The identification question is whether we can map backwards from knowledge of the control rule to the parameter θ, i.e., whether the map c is invertible. The general answer is no. Brouwer's Theorem on Invariance of Domain states that there is no homeomorphism between spaces of different dimensions. Thus, if the parameter space is even just an interval in R^1, there is no way to identify the unknown (in this case scalar) parameter from knowledge of the control rule (a point in a finite set). As we have seen, the parameter values can be bounded, but the parameter cannot be estimated without further information.

Further information comes from the transition distribution. There are H origin states and H destination states, hence H^2 transition probabilities less H from the adding-up constraint for a given value of the control rule. Since transitions are only seen under the optimal control, these are the only transition probabilities that can be estimated using transition data. Of course, the other transition probabilities enter the problem in determining the control rule. These can be estimated only from the control rule, and hence can at best be bounded into an interval. The notation developed above generalizes easily. There are H^T possible sequences of states of length T. Arrange these sequences in lexicographic order and index these by $i, j, k \in \{0, ..., H^T - 1\}$. Then the ith sequence is the H-ary expansion of i. Let $\xi(i, s)$ be the sth digit in the H-ary expansion of i. Let $\kappa(i, r) = \xi(i, r)||\xi(i, r+1)$ be the rth pair of digits in the H-ary expansion of i. Finally, let $N\kappa(i, ab) = \Sigma_{r=1}^{T-1} I(\kappa(i, r) = ab)$, the number of ab pairs in the ith sequence (here a, b take on H distinct values). The log likelihood factors according to the origin state, so we have for example for origin state 0

$$l(p_{0a}, a = 0, .., H - 1|i, x_0) = \Sigma_{a=0}^{H-1} N\kappa(i, 0a) ln p_{0a}, \tag{19}$$

defined with the constraint $\Sigma_{a=0}^{H-1} p_{0a} = 1$. The model is clearly in the exponential family. The MLE's are

$$\widehat{p}_{0a} = N\kappa(i, 0a)/\Sigma_{a=0}^{H-1} N\kappa(i, 0a), \tag{20}$$

if these values are consistent with the observed control rule. Thus, $H(H - 1)$ transition probabilities can be estimated from the transition data. Note however that a constraint here (eg, some transitions are impossible, others are necessary, etc.) does not imply that there are degrees of freedom available for estimating utility function parameters. At most, these estimates can be used to refine the bounds on parameters imposed by knowledge of the control rule.

Although the model predicts that a given state should always be associated with the same control, and therefore the control rule is learned rapidly (as soon as each possible state has been observed once), the data will rarely satisfy such a strong requirement. This is the "curse of determinacy." The model, although stochastic, predicts a deterministic relationship between the state and control. This is one of the major difficulties in applying dynamic programming models empirically. There is a number of approaches to modifying the model to be consistent with data not satisfying this deterministic constraint. The approaches are not equally successful.

6 Measurement Error

One natural approach to breaking the curse of determinacy is to allow for measurement error. The idea here is that the model is an accurate description of behavior, but that we are not measuring exactly the quantities entering the optimization problem. We first develop the appropriate notion of measurement for discrete models, working first with the binary situation as above. Then we consider in turn measurement error in the state and in the control. Finally, we consider the case of measurement error in both state and control. We find sensible specifications which do break the curse.

A simple specification for measurement error in a binary model is to allow a constant misclassification probability (crossover probability) ε. Letting x be the true state and x^* the observed, the model for the measurement process is illustrated in Figure 2.

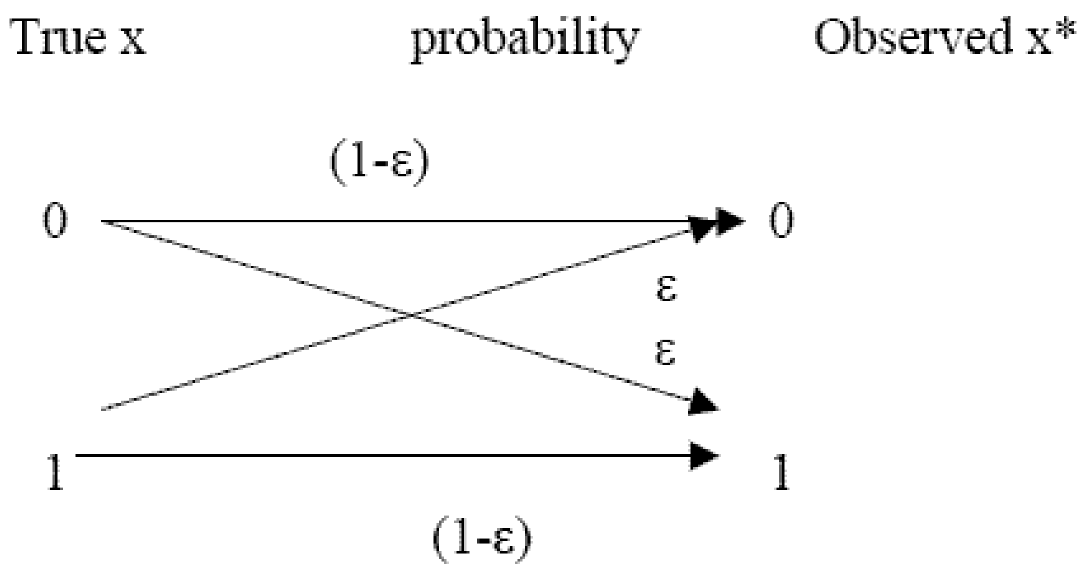
Suppose x is a realization of a sequence from a Markov chain with transition probabilities p_{00}, $p_{01} = 1 - p_{00}$, p_{11}, and $p_{10} = 1 - p_{11}$. Thus, trivially, $P(x_2 = 0|x_1 = 0, x_0 = a) = p_{00} = P(x_2 = 0|x_1 = 0)$. To check, note that $P(x_1 = 0|x_0 = a) = p_{a0}$ and $P(x_2 = 0, x_1 = 0|x_0 = a) = p_{a0}p_{00}$ and dividing gives the result. Suppose x^* is the sequence with measurement error, so that $x_t^* = x_t$ with probability $1 - \varepsilon$ and $x_t^* = 1 - x_t$ with probability ε. Let us calculate $P(x_2^* = 0|x_1^* = 0, x_0^* = a)$. We begin by calculating $P(x_1^* = 0|x_0^* = a)$. Note first that all $2 \times 2 = 4$ sequences of length 2 for the true state variables are consistent with observing the sequence $a0$. Thus ,

$$P(x_1^* = 0|x_0^* = a) = p_{a0}(1-\varepsilon)^2 + p_{a1}(1-\varepsilon)\varepsilon + p_{c0}\varepsilon(1-\varepsilon) + p_{c1}\varepsilon^2 \qquad (21)$$

where the index $c = 1 - a$. Next, calculate $P(x_2^* = 0, x_1^* = 0|x_0^* = a)$. All 2^3 sequences are consistent with the observed pattern $a00$. The probability is thus the sum of 8 terms. The term corresponding for example to the true sequence $a01$ is $p_{a0}p_{01}(1-\varepsilon)^2\varepsilon$, the probability that a process beginning in a is observed to be in a, $(1-\varepsilon)$, times the probability of moving from a to 0, p_{a0} multiplied by the probability of being observed correctly in the second period, $(1-\varepsilon)$, times the probability of moving from 0 to 1 and being incorrectly observed in the final period $(p_{01}\varepsilon)$. Upon adding these terms and dividing we see that

$$\begin{aligned} P(x_2^* &= 0|x_1^* = 0, x_0^* = a) = P(x_2^* = 0, x_1^* = 0|x_0^* = a)/P(x_1^* = 0|x_0^* = a) \\ &\neq P(x_2^* = 0|x_1^* = 0). \end{aligned} \qquad (22)$$

Figure 2: Measurement Error



Thus, the x^* process is not Markovian. This suggests a simple diagnostic before calculating estimates for a complex dynamic programming model. Namely, examine the state sequence to see whether it looks Markovian. If not, either reformulate the state variable specification (sometimes the model can be made Markovian by appropriate choice of the state variables) or consider the possibility of measurement error in the state variable.

To formulate the likelihood for the observed sequence x^*, we first calculate the probability of seeing x^* conditional on the actual underlying sequence x, then essentially marginalize with respect to x. This strategy is attractive since it is easy to calculate the probability of observing the x sequences - realizations from Markov processes. We have

$$p(x^*|i) = \Pi_{t=1}^{T} \varepsilon^{|x_t^* - \xi(i,t)|}(1-\varepsilon)^{1-|x_t^* - \xi(i,t)|}, \tag{23}$$

the probability of observing the sequence x^* when the ith x sequence was actually realized (recall our convention on ordering the sequences). The probability of realizing the ith x sequence of length T is

$$p(i) = \Pi_{r=1}^{T-1} p_{\kappa(i,r)} \tag{24}$$

where $\kappa(i,r)$ is the rth pair of digits in the binary representation of i. Hence, the marginal probability of observing the measured sequence x^* is

$$p(x^*) = \Sigma_{i=0}^{2^T-1} \Pi_{t=1}^{T} \varepsilon^{|x_t^* - \xi(i,t)|}(1-\varepsilon)^{1-|x_t^* - \xi(i,t)|} \Pi_{r=1}^{T-1} p_{\kappa(i,r)}, \tag{25}$$

and in the absence of additional information this would serve as the likelihood function for the unknown parameters $\theta = (p_{00}, p_{11}, \varepsilon)$.

There is however additional information in both the known control rule (given parameters) and the observed control sequence. Knowledge of the control rule as we have seen restricts the range of possible parameter estimators. Observation of the control sequence, however, is equivalent to observation of the true state sequence, given the parameters. The data series consists of the sequence c and the sequence x^*, with joint probability distribution

$$\begin{aligned} p(c, x^*) &= \Sigma_x p(c, x, x^*) \\ &= \Sigma_x p\left(c|x, x^*\right) p\left(x^*|x\right) p\left(x\right), \end{aligned} \tag{26}$$

conditioning on a value of the true sequence x and then marginalizing. This formulation is useful since $p(c|x) = p(c|x, x^*)$ is degenerate at $c = c\left(x\right)$, with $c\left(x\right)$ from the control rule (note that this function does depend on parameters). We treat here for simplicity the 2×2 case with $c(x)$ invertible, noting that the results apply immediately to $K \times K$ models, and treat the noninvertible case below (this is the case with more state than control variables). Thus, the probability $p(c|x)$ is zero except for the ith sequence x, where the ith sequence satisfies $c_t = c(\xi(i,t))$. Hence

$$p(c, x^*) = \Pi_{t=1}^{T} \varepsilon^{|x_t^* - \xi(i,t)|}(1-\varepsilon)^{1-|x_t^* - \xi(i,t)|} \Pi_{r=1}^{T-1} p_{\kappa(i,r)}, \tag{27}$$

yielding the MLEs $\widehat{\theta} = (\widehat{p}_{00}, \widehat{p}_{11}, \widehat{\varepsilon})$ with $\widehat{p}_{ab} = N\kappa(i(\widehat{\theta}), ab)/\Sigma_{ab}N\kappa(i(\widehat{\theta}), ab)$ and $\widehat{\varepsilon} = T^{-1}\Sigma_t|x_t^* - \xi(i(\widehat{\theta}), t)|$ with $i(\widehat{\theta})$ the index of the x sequence satisfying $c = c\left(x; \widehat{\theta}\right)$, where the presence of parameters in this condition is now explicit for emphasis. This is important: When the likelihood is evaluated at a different parameter value θ, it may require different i, as well. The likelihood is thus only piecewise continuous in parameters, so some care must be taken in ensuring that the estimators satisfy the constraint that $x = c^{-1}(c; \widehat{\theta})$ where $\widehat{\theta}$ is the estimator, and that the estimators in fact correspond to a global likelihood maximum. In fact, the control rule and thus $i(\widehat{\theta})$ typically do not depend on the unknown parameter ε, which affects only observation of the data and does not enter the agent's optimization problem. Thus, we have an explicit solution for the maximizing value of ε given the other parameters. This can be substituted back into the loglikelihood

$$\begin{aligned} l(p_{00}, p_{11}, \varepsilon) &= \Sigma_{t=1}^{T} \{|x_t^* - \xi(i,t)| ln(\widehat{\varepsilon}) + (1 - |x_t^* - \xi(i,t)|) ln(1 - \widehat{\varepsilon})\} \\ &\quad + \Sigma_{r=1}^{T-1} \ln(p_{\kappa(i,r)}) \end{aligned} \tag{28}$$

to yield the profile loglikelihood function

$$\begin{aligned} l(p_{00}, p_{11}|c, x^*) &= \Sigma_{t=1}^{T} \Big\{|x_t^* - \xi(i,t)| ln(T^{-1}\Sigma_s|x_s^* - \xi(i,s)|) + \\ &\quad (1 - |x_t^* - \xi(i,t)|) ln(1 - T^{-1}\Sigma_s|x_s^* - \xi(i,s)|)\Big\} \\ &\quad + \Sigma_{r=1}^{T-1} \ln(p_{\kappa(i,r)}) \\ &= T(\widehat{\varepsilon} ln(\widehat{\varepsilon}) + (1-\widehat{\varepsilon}) ln(1-\widehat{\varepsilon})) + \Sigma_{r=1}^{T-1}\ ln(p_{\kappa(i,r)}), \end{aligned} \tag{29}$$

in which it must be emphasized that both $\widehat{\varepsilon}$ and i depend on the parameters $p_{00},\ p_{11}$.

Note that the curse of determinacy has been broken, in that the observed c, x^* pairs need not satisfy $c = c(x^*)$ for every observation. Thus, the data table does not have to be in the form of Table 5 above. The extent to which this form is not satisfied is used to estimate the crossover probability ε. Note also that ε does not enter the optimization problem and is not restricted by knowledge of the optimal control policy. The control rule is not learned quickly and with certainty as in the completely observed case.

Turning now to the case of measurement error in observation of the control, we allow misclassification with probability ε_c. With the state observed without error the argument is completely analogous to the case of observed controls and states with measurement error. That is, observation of the states gives the controls deterministically as a function of parameters. We repeat details briefly. The conditional distribution of the observed controls, given the state sequence is the ith, is

$$p(c^*|i) = \Pi_s \varepsilon_c^{|c(\xi(i,s)) - c_s^*|}(1-\varepsilon_c)^{1-|c(\xi(i,s)) - c_s^*|}, \tag{30}$$

hence the joint distribution is

$$p(c^*, i) = \Pi_s \varepsilon_c^{|c(\xi(i,s) - c_s^*|}(1-\varepsilon_c)^{1-|c(\xi(i,s) - c_s^*|}\Pi_{r=1}^{T-1} p_{\kappa(i,r)}, \tag{31}$$

yielding the MLEs $\widehat{\theta} = (\widehat{p}_{00}, \widehat{p}_{11}, \widehat{\varepsilon}_c)$ with $\widehat{p}_{ab} = N\kappa(i(\widehat{\theta}), ab)/\Sigma_{ab} N\kappa(i(\widehat{\theta}), ab)$ and $\widehat{\varepsilon}_c = T^{-1}\Sigma_s|c(\xi(i(\widehat{\theta}), s) - c^*_s|$. Note that the function $c(x)$ depends on unknown parameters through the optimization problem. The constraints mentioned above in the discussion of measurement error in states must be satisfied at the MLEs. Once again, the profile loglikelihood is easily obtained. It is

$$\begin{aligned} l(p_{00}, p_{11}|c^*, i) &= \Sigma_s \ \big\{|c(\xi(i,s)) - c^*_s|(ln(T^{-1}\Sigma_s|c(\xi(i,s)) - c^*_s|)+ \\ &\quad (1 - |c(\xi(i,s)) - c^*_s|ln(1 - T^{-1}\Sigma_s|c(\xi(i,s)) - c^*_s|)\big\} \\ &\quad +\Sigma_r \ln(p_{\kappa(i,r)}) \\ &= T(\widehat{\varepsilon}_c ln(\widehat{\varepsilon}_c) + (1 - \widehat{\varepsilon}_c)ln(1 - \widehat{\varepsilon}_c)) + \Sigma_r \ ln(p_{\kappa(i,r)}). \quad (32) \end{aligned}$$

Again, the dependence of $\widehat{\varepsilon}_c$ and i on parameters is emphasized.

The curse of determinacy is broken, in that the observed state/control sequence does not have to be degenerate. That is, the same state value can be associated with different observed controls without implying a breakdown of the statistical model. Further, it is not the case that the control rule is learned immediately after each state value has been realized - instead, information is accumulated over time.

Both specifications, measurement error in states and measurement error in controls, break the curse of determinacy. The curse of degeneracy is still present, in that $p\,(c|x) \in \{0, 1\}$, imperfect observations of c or x simply makes it a little more difficult to learn which. Thus, for data in which the summary statistics do not have the structure of Table 5, i.e. for data in which the same state is associated at different time periods with different controls, one of these measurement error models might be appropriate. But they have different implications for observables, in that the observed sequence x is Markov, while x^* is not. Thus, as a practical matter, introducing measurement error in controls might be appropriate if the observed state sequence appears Markovian, measurement error in states if not.

Finally, we turn to a specification with measurement error in both states and controls. Assume that the measurement error in the states and controls are independent, conditionally on the underlying realization of the process. Then

$$\begin{aligned} p(x^*, c^*|i) &= p(x^*|i)p(c^*|i) \\ &= (\Pi_{t=1}^T \varepsilon^{|x^*_t - \xi(i,t)|}(1-\varepsilon)^{1-|x^*_t - \xi(i,t)|}) \times \\ &\quad (\Pi_s \varepsilon_c^{|c(\xi(i,s)) - c^*_s|}(1 - \varepsilon_c)^{1-|c(\xi(i,s)) - c^*_s|}) \quad (33) \\ &= \Pi_{t=1}^{T-1} \varepsilon^{|x^*_t - \xi(i,t)|}(1-\varepsilon)^{1-|x^*_t - \xi(i,t)|} \varepsilon_c^{|c(\xi(i,t)) - c^*_t|}(1-\varepsilon_c)^{1-|c(\xi(i,t)) - c^*_t|}. \end{aligned}$$

The marginal probability $p(i) = \Pi_{r=1}^{T-1} p_{\kappa(i,r)}$; multiplying and marginalizing gives

$$p(x^*, c^*) = \Sigma_{i=0}^{2^T - 1} \Pi_{t=1}^T \varepsilon^{|x^*_t - \xi(i,t)|}(1-\varepsilon)^{1-|x^*_t - \xi(i,t)|} \varepsilon_c^{|c(\xi(i,t) - c^*_t)|}\,(1-\varepsilon_c)^{1-|c(\xi(i,t) - c^*_t|} p_{\kappa(i,t)}, \quad (34)$$

leading to a likelihood function which is substantially more complicated in that it has 2^T terms and no simple closed forms for the estimators. Essentially, when either the state or control is observed without error, then the other is "known" in the sense that it is a deterministic function of parameters. This is not the case when both are measured with error. Nevertheless the likelihood function is not continuous in parameters. This is general, since the controls make up a discrete set.

7 Imperfect Control

Here we study a second approach to breaking the curse of determinacy. That is, we model a decision maker with imperfect control over the action he takes. Thus, the agent may know that $c = 0$ is optimal for $x = 0$, but may only be able to achieve $c = 0$ with high probability, not with certainty. This kind of imperfect control has been used e.g. by Chow (1981).

The extension fits easily into our simple framework. The c variables, which the agent would like to control exactly but cannot, are $c \in \{0, 1\}$. Define the variables $a \in A = \{0, 1\}$ as the variables the agent actually can control; if $a = 0$ is chosen then $c = 0$ with probability p_0; if $a = 1$ then $c = 0$ with probability p_1. Specify without loss of generality that $p_0 > p_1$, so that $a = 0$ is the natural choice if the agent would prefer $c = 0$, etc. By choosing a, the agent chooses a probability distribution over the controls. Let p_a be the probability that $c = 0$ when action a is chosen. We can now apply the dynamic programming framework.

Define the new utility function

$$u^*(x, a) = E_a u(x, c) = p_a u(x, 0) + (1 - p_a) u(x, 1) \tag{35}$$

and the new transition distribution

$$p(x'|x, a) = p(x'|x, c = 0) p_a + p(x'|x, c = 1)(1 - p_a). \tag{36}$$

Then we can simply do dynamic programming using a as the control instead of c. The value function satisfies Bellman's equation

$$V_i(x) = \max_{a \in A} \{u^*(x, a) + \beta E_i V_{i-1}(x')\} \tag{37}$$

in the finite horizon case, and

$$V(x) = \max_{a \in A} \{u^*(x, a) + \beta E V(x')\} \tag{38}$$

in the infinite horizon case. In our simple marketing model from the previous sections, the new utility and transition functions using $p_0 = .8$ and $p_1 = .2$ are given in Table 9 (compare Tables 1 and 2).

Table 9. Imperfect Control

Profit

	$a=0$	$a=1$
$x=0$	6.4	4.6
$x=1$	10.2	7.8

$\mathbf{p}\left(x_{t+1}=1|x_t,a_t\right)$

	$a=0$	$a=1$
$x=0$	0.178	0.700
$x=1$	0.570	0.780

The first 10 value function iterations are given in Table 10 (compare Table 3).

Table 10. Value Functions and Optimal Policies

t	$V_t\left(0\right)$	$V_t\left(1\right)$	$a_t\left(0\right)$	$a_t\left(1\right)$
0	6.40000	10.20000	0	0
1	11.70730	16.62450	0	0
2	15.96201	21.08258	1	0
3	19.25980	24.36055	1	0
4	21.72274	26.82542	1	0
5	23.57096	28.67345	1	0
6	24.95703	30.05954	1	0
7	25.99659	31.09909	1	0
8	26.77626	31.87876	1	0
9	27.36101	32.46351	1	0

Here, we see that it is optimal to try to run the advertising policy in the low demand period if the horizon is 3 or more periods (the exact control case required 4 or more). The policy has converged, though the value functions have not, and this is indeed the optimal policy in the infinite horizon problem. The value function converges to (approximately) $V(0)=29.11526$ and $V(1)=34.21777$. Comparing the analysis of the information on β in learning the optimal policy, we see here that this policy, $(a(0),a(1))=(1,0)$, is optimal for $\beta>0.6694$; in the perfect control case we found that $(c(0),c(1))=(1,0)$ was optimal for $\beta>0.7143$. In contrast to the case with pure measurement error, introducing imperfect control changes the solution to the optimization problem. The lesson here is that observation error and imperfect control are quite different specifications.

The observables remain the $\{x,c\}_t$ sequence, for $t=0,...,T$. Now, the state sequence x_t is observed without error, so we can concentrate on a single sequence (without marginalizing with respect to all possible sequences as in the case of measurement error in both states and controls). The sequence c_t is also observed without error, but it is no longer the control. The controls are the unobserved a_t. However, the observed c_t can be regarded as noisy observations on the actual controls a_t. Thus, an approach very similar to the approach in the case of measurement error can be used in developing this factor of the likelihood. Finally, note that there is more reduced form information on the transition distribution in the case of imperfect control. That is, in the case of perfect control, even with measurement error, the only transitions observed are those corresponding to x, c pairs (the conditioning variables in the transition distribution) which are optimal. The other transition probabilities enter the likelihood only through their effect on the optimal policy. Without perfect control, transitions corresponding to all x, c pairs are observed without error and can be used to estimate the transition probabilities directly. Let us develop the likelihood:

$$\begin{aligned} p(x', c'|x, c) &= p(c'|x, x', c)p(x'|x, c) \\ &= p(c'|x')p(x'|x, c), \end{aligned} \tag{39}$$

since the relation between the state and the control does not depend on lagged values, and neither does the measurement error. The first factor can be simplified to

$$p(c'|x') = \Sigma_{a'} p(c'|a', x')p(a'|x'), \tag{40}$$

where the sum is over all possible $\{a\}_t$ sequences. Note however that the term $p(a'|x')$ is degenerate, in that for given parameters, there is only one a sequence corresponding to the realized states. Thus, all but one of these terms are 0 and hence

$$\begin{aligned} p(c'|x') &= p(c'|a', x') \\ &= p(c'|a') \end{aligned} \tag{41}$$

for that value of a' consistent with x' (and parameters). Let i be the index of the observed x sequence. Then

$$p(c|i) = \Pi_{s \in \{a(\xi(i,s))=0\}} p_0^{1-c_s}(1-p_0)^{c_s} \Pi_{s \in \{a(\xi(i,s))=1\}} p_1^{c_s}(1-p_1)^{1-c_s}, \tag{42}$$

and in the case $p_0 = 1 - p_1$, as in our example, there is further simplification to

$$p(c|i) = \Pi_s p_0^{(1-|a(\xi(i,s))-c_s|)}(1-p_0)^{|a(\xi(i,s))-c_s|}, \tag{43}$$

very like the likelihood in the measurement error case, but with one important difference. That is, here the function $a(\xi(i, s))$ depends on all parameters, including p_0. We can find an illustrative expression for the MLE for p_0, namely $\widehat{p}_0 = 1 - T^{-1}\Sigma_s|a(\xi(i(\widehat{\theta})), s) - c_s|$, but here this is only one of many equations that must be solved simultaneously, since p_0 enters the function a, unlike the parameter ε in the measurement error case, where a similar expression can be used to obtain the profile likelihood.

Summing up so far, we have developed the likelihood function in a compact notation for the discrete state - discrete control setup. The curses of degeneracy (a property of the distribution of the control given the state) and of determinacy (a requirement of the data configuration) are easily illustrated here. The control rule becomes known after a few observations (that is, there is no sampling error - as soon as all of the states have been realized, the control rule is known). In general, knowledge of the control rule is not sufficient to identify underlying real parameters. Two approaches to breaking the curses were examined. The first was measurement error. Here there are two possibilities, measurement error in states or in controls (or, of course, both). Measurement error in states implies that the state-to-state transitions are not Markovian. Thus, this specification might be useful when the transitions are not Markovian. Here, the data do not have to satisfy the unlikely restrictions imposed by the perfect observation case (the curse of determinacy). However, the curse of degeneracy is not broken, essentially because the perfect observation of the controls identifies the states, given the parameters. Measurement error in controls is essentially the same. Observing the states without error identifies the controls, given parameters. Here, however, the state-to-state transitions remain Markovian. Combining both types of measurement error leads to a more complicated likelihood, as it is no longer possible to recover the true states and controls given the parameters. The curse of determinacy is broken in all cases, although the curse of degeneracy remains. Real parameters are typically not identified, although their ranges may be restricted. Imperfect control is an alternative approach. The results are somewhat different from the measurement error case, in that the optimal policies may differ from those in the perfect control setting, although the implications are the same in that the curse of determinacy is broken but that of degeneracy is not.

Neither measurement error nor imperfect control is particularly appealing from an economic modelling point of view, and neither actually solves the problem we wish to solve. Economists have been led almost invariably to a random utility specification, which does solve the curse of degeneracy and allows identification of real parameters, but which does so by introducing "information" in the form of highly specific assumptions. That is, the random utility specification allows estimation of parameters that are not identified if utility is deterministic in a model that is otherwise the same. This is the topic of Section 8.

8 Random Utility Models

A useful and popular approach to breaking the curse of degeneracy is to introduce a random utility specification. Here, the period utility is subject to a control-specific shock. The agent sees this shock before the control choice must be made, but it is not seen in the data. Thus, the choice of the agent may depend on the realization of the shock, and hence the observed optimal control may correspond to different observed states depending on the value of the unobserved shock. Of course, the random utility shock cannot simply be "tacked on" to a dynamic programming problem; it changes the problem, the value and the optimal policy function. The approach of specifying the random utility shock as an unobserved state variable was introduced in empirical dynamic programming by Rust (1987), following McFadden (1973, 1981) on static discrete choice.

Suppose now that the utility is subject to a random shock, so that

$$u^*(x, c, \varepsilon) = \overline{u}(x, c) + \varepsilon(c), \tag{44}$$

where $\varepsilon(c)$ is a random shock. The idea here is that, at time t when the state $x_t = k$ has been realized, a vector of random variables ε is added to the kth row in the utility table, then the choice of control c is made. Thus $c = c(x, \varepsilon)$ is a deterministic function, but given only x, c is a random variable whose distribution depends on x. If ε has a nonzero mean, the mean can simply be absorbed into the utility specification, so it is a harmless normalization to set $E\varepsilon = 0$.

Consider the 2×2 model and suppose the utility shock is $(\varepsilon, 0)$ where ε is a scalar random variable taking the value $-a$ with probability $\frac{1}{2}$ and a with probability $\frac{1}{2}$. Consider for simplicity the case $x = 0$ alone, so we can focus attention on the distribution $p(c|x = 0)$. Let $p_1 = p(c = 1|x = 0)$; of course, p_1 is a function of the parameter θ characterizing preferences, etc. We have

$$p(c|x = 0) = p_1^c(1 - p_1)^{1-c}, \tag{45}$$

and with n independent observations

$$\Pi_t p(c_t|x_t = 0) = p_1^{\Sigma c}(1 - p_1)^{n-\Sigma c}, \tag{46}$$

and $\Sigma c/n$ is a sufficient statistic for p_1. Here, n is the number of observations with $x = 0$. The natural parameter space for the reduced-form parameter p_1 is $[0, 1]$. But what values of p_1 are consistent with the dynamic programming model with random utility shocks?

In the model, the probability that $c = 1$ is chosen is given by

$$Pr(\arg\max_c\{u(0, c) + \beta E_{0c}V(x', \varepsilon') + \varepsilon(c)\} = 1). \tag{47}$$

This probability can be written

$$Pr(\varepsilon < h(\theta)), \tag{48}$$

where $h = u(0, 1) + \beta E_{01}V - u(0, 0) - \beta E_{00}V$ and the generic parameter θ has been introduced as an argument in h for emphasis. Now, we have specified a binary distribution for ε, so this probability can take values in $\{0, 1/2, 1\}$; the precise value depending on $h(\theta)$. Specifically, $c = 0$ could be optimal for both values of $\varepsilon(0)$; and $c = 1$ could be optimal for $\varepsilon(0) = -a$ and $c = 0$ for $\varepsilon(0) = a$, and finally $c = 1$ could be optimal for both values of $\varepsilon(0.)$ Thus, the likelihood function is flat for generic continuous parameters θ, except where the value of θ is such that the probability shifts between 2 of its 3 possible values. The most that can be obtained is that possible values of θ are restricted to intervals, corresponding to values so that the implied choice probability is as close as possible to the sample fraction. Note that this does represent some improvement over the model without shocks - if h is monotonic in the parameter θ, then without shocks, so $p(c|x) \in \{0,1\}$, θ is restricted to one of two intervals (recall estimating β in our marketing model), now with $p(c|x) \in \{0, 1/2, 1\}$, θ is restricted to one of 3 intervals.

The curse of degeneracy has been broken in that $p(c|x)$ is not necessarily in $\{0,1\}$. However, only one new point has been added to that set of possibilities, so it might be better to think of the curse as weakened, not broken. The main point, however, is that merely adding a random utility shock with a completely known distribution appears to add information about a structural parameter, but does not serve to identify fully otherwise unidentified parameters.

Let us generalize the shock distribution slightly and introduce a new parameter. Suppose the distribution of ε is $-a$ with probability $(1-\pi)$ and a with probability π. For π not equal to $\frac{1}{2}$ and a not equal to 0, this distribution has a nonzero mean. That does not really present a problem, as the mean can be absorbed into $u(0,0)$, as we have seen. For a fixed value of π, we have that $Pr(\varepsilon < h(\theta, \pi)) \in \{0, \pi, 1\}$, so again θ can at best be bounded in an interval. By varying π, however, it may be possible to match the sample fraction (it may not be possible, if the probability is zero or one for all θ but this is not an interesting case). Here, the curse of degeneracy is unambiguously broken, in the sense that the sample fraction can be matched by choice of the parameter π. However, this amounts to no more than a reduced-form approach; no identifying information on θ is available and it is at best restricted to an interval.

9 A Continuously Distributed Utility Shock

Since adding a utility shock with a known 2-point distribution adds one point to the model-consistent parameter space for p_1 (and hence possibly restricts the range of the generic parameter θ), it is natural to ask whether a known distribution with support on an interval might tighten things up even more. In fact, the assumption of a continuously distributed utility shock is much more common in applications, for reasons which will become clear. Let us begin the analysis by supposing that, instead of a 2-point distribution on $\{-a, a\}$, ε has a continuous distribution with support $[-a, a]$. Note that we are not making innocuous assumptions by changing the shock distribution - changes here, even with the mean held constant, will affect both the value of the problem and the optimal policy function. To start with, assume $f(\varepsilon) = 1/(2a)$, the uniform distribution. Then

$$p(c = 1|x = 0) = Pr(\varepsilon < h(\theta)) = h(\theta)/(2a) + 1/2. \tag{49}$$

Recall that $h(\theta) = u(0,1) + \beta E_{01}V - u(0,0) - \beta E_{00}V$. This probability now does depend on θ, continuously if h is continuous in θ, so θ can be estimated by setting $h(\theta)/(2a) + 1/2$ equal to the sample fraction and solving for $\widehat{\theta}$, as long as this is a feasible value (again, it could be that $h(\theta)$ is such that $|h| > a$ for all values of θ, and therefore the probability is always zero or one; not an interesting case).

By introducing randomness into the utility specification, we have managed to achieve identification of a parameter not identified without randomness. How can "introducing noise" serve to identify a preference parameter? What is in fact happening is that the assumption that the shock is continuously distributed with a known distribution is crucial. Our specification inserts "information" into the model - information on preferences. This is not necessarily inappropriate, but it is important to realize that θ is being identified completely on the basis of the assumed shock distribution. Indeed, this point can be emphasized with a little further analysis.

Let F be the distribution function for the utility shock. For simplicity, suppose F is a member of a one-parameter family of distributions indexed by γ. Then

$$\begin{aligned} p\left(c=1|x=0\right) &= \Pr\left(\varepsilon < h\left(\theta,\gamma\right)\right) = \int_{-a}^{h(\theta,\gamma)} dF \\ &= H\left(\theta,\gamma\right). \end{aligned} \tag{50}$$

In practice, γ is assumed known (the distribution of the utility shock is fully specified), and $H(\theta,\gamma) = t$ (the sample fraction) is solved for $\widehat{\theta}(\gamma)$. Here, the dependence on the assumed distribution is indicated by the explicit dependence on γ. Assume F is continuously differentiable in γ and h in θ and γ. This assumption rules out sudden shifts of the probability to zero or one (e.g. when h passes the value a); but rather than get involved in details we note that we are really concerned only with properties in the neighborhood of the solution to $H(\theta,\gamma) = t$. Alternatively, we can simply set $a = \infty$ so the distribution function has full support on the real line (indeed this is the most common practice). Using the implicit function theorem,

$$d\widehat{\theta}/d\gamma = -H_\gamma/H_\theta \tag{51}$$

with $H_\gamma = f(h(\theta,\gamma)h_\gamma + F_\gamma(h(\theta,\gamma))$ and $H_\theta = f(h(\theta,\gamma))h_\theta$ where f is the density function $dF(x)/dx$. Writing this out gives

$$d\widehat{\theta}/d\gamma = -h_\gamma/h_\theta - F_\gamma/fh_\theta, \tag{52}$$

the first term giving the tradeoff between γ and θ holding h (the utility difference between using the two controls) constant and the second giving the effect of γ on h through the change in the probability.

This analysis indicates that the solution for the parameter θ is a function of the assumed distribution, here indexed by γ. Further, this derivative is typically nonzero, so the assumption matters. To push this analysis a little further, we concentrate on the parameter β, the discount factor. Then,

$$h(\beta,\gamma) = ((u(0,1) - u(0,0)) + \beta(E_{01}V - E_{00}V), \tag{53}$$

hence $h_\beta = (E_{01}V - E_{00}V) + \beta(E_{01}V_\beta - E_{00}V_\beta))$ and $h_\gamma = \beta d(E_{01}V - E_{00}V)/d\gamma$. Let us evaluate these expressions at $\beta = 0$, in order to get a local tradeoff between the assumed distribution and the estimated discount factor at $\beta = 0$. Here

$$h_{\beta|\beta=0} = E_{01}\left(\max\{u(x,c) + \varepsilon(c)\}\right) - E_{00}(\max\{u(x,c) + \varepsilon(c)\}), \tag{54}$$

a function of γ, and

$$h_{\gamma|\beta=0} = 0, \tag{55}$$

so there is no direct effect of the assumed distribution γ on h, the utility difference, when $\beta = 0$. Hence, all of the effect is through the change in the probability, and this is given by

$$d\widehat{\beta}/d\gamma_{|\beta=0} = -(F_\gamma/f)/[(E_{01}(\max\{u(x,c) + \varepsilon(c)\} - E_{00}(max\{u(x,c) + \varepsilon(c)\})]; \tag{56}$$

all terms in this expression depend on γ, so we have illustrated now in a simple case the correspondence between the specification of the utility shock distribution and the estimated discount factor. Note that this formula gives the effect on the estimate of β of a change in the distributional assumption, holding the data constant. The numerator is the change in the choice probability; the denominator is the density multiplied by the expected utility difference. While this expression clearly depends crucially on F (this is the point, after all), a little more may be said. In particular, if the expected utility difference between the two controls is smaller, then the role of the assumption on F is more important, in that this derivative is larger.

It is sometimes thought that by "freeing up" the random utility distribution F through addition of unknown parameters, one can mitigate the effects of directly and completely specifying an unknown distribution. Let us consider this. First, in the stylized case studied above, with only one choice probability to estimate (i.e., the probability associated with one value of the discrete state variable and a binary control), this proposition is easily rejected. Our analysis shows there is a $1-1$ relationship between values of the parameter γ of the now unknown distribution and the preference parameter θ. Clearly, these are not both identified. Specifically, there is a curve in the (γ, θ) space corresponding to a given value of h, and thus a set of parameter values which serve to match the fitted value h to the sample fraction. Identification requires that the functional form of the utility shock be completely specified.

10 Continuous State and Optimal Stopping: The Search Model

In the previous sections we have drawn attention to several issues in identification and inference in dynamic programming models, such as identification of utility parameters (including the discount factor β) only up to intervals, rapid (finite time) accumulation of information on these, the curses of determinacy and degeneracy, the role of the functional equation in identifying parameters off the optimal path, the subtle differences between measurement error, imperfect control and random utility approaches, and the functional dependence of the parameter estimates on the assumed shock distribution in the random utility case. These issues or close relatives apply generally (and in various disguises) in dynamic programming models and have been illustrated in the discrete state/control case. Many of the insights originate from earlier motivating work on the econometrics of the job search model, although this model differs slightly from the setup of the previous sections. It is an optimal stopping model, so it does have a discrete (binary) control, namely, whether or not to stop, but it has a continuous state variable (the wage). For statistical purposes, information is in many cases only accumulated until the process is stopped, so for asymptotics, a panel is considered. In this section, we briefly review some of the links between the material in the previous sections (general dynamic programming and marketing) and the motivating work on the statistical properties of the search model.

The sequential job search model is due to Mortensen (1970) and McCall (1970). The search model was an early application of stochastic dynamic programming techniques to economic theory. It was among the first models to be estimated with econometric techniques exploiting the dynamic programming structure. In the simplest setup, a worker is assumed to be unemployed and searching for employment. Search consists of sampling, once each period, a wage offer w from a known distribution of offers. Once accepted, a job is held forever. Once declined, an offer is no longer available. The state variable is the outstanding wage offer in the current period. The control is the decision to accept or reject the outstanding offer. Offers are assumed independently and identically distributed (iid), so the distribution of next period's state does not depend on the current state. The state distribution does depend on the control, since offers are no longer received once a job has been accepted. The worker chooses a strategy which maximizes the expected present discounted value of his income stream $E\Sigma_t^T \beta^t w$ where in the simplest models $T = \infty$. This model has been extended and refined and widely applied.

The basic logic of optimal stopping can be illustrated in the infinite horizon model. The "value" for an unemployed worker is heuristically the maximized value of $E\Sigma_t^T \beta^t w$, where the maximum is taken over all possible strategies the worker might follow in his effort to maximize the present discounted value. This value V^u is a constant not depending either on the current (declined, since we are assuming the worker is unemployed) wage offer or on the particular period. The value does not depend on the current offer since we assume offers are iid; it does not depend on the period since we have an infinite horizon problem and the future looks the same from any point. Now suppose the worker gets the next offer w. The value of accepting the offer w is simply $\Sigma_t^T \beta^t w = w/(1-\beta)$. If this value is greater than the value of continued search, namely V^u, the worker should accept the offer; if not he should decline and continue search. Thus we have found the optimal strategy. The worker should decline wage offers until he receives one greater than $r = (1-\beta)V^u$, then he should accept employment and stop searching. In particular, r is the reservation wage.

Note that the logic of backward recursion does not apply here. It would if we considered finite horizon search, and then V^u would depend on the time left before the horizon.

Let b be the current utility of unemployment benefits net of search costs. Then $V^u = b + \beta EV$, and with the state variable $x = w$ we have the value function

$$V(w) = \max\left\{\frac{w}{1-\beta}, b + \beta EV\right\}, \tag{57}$$

the maximum of the value of stopping and accepting the offer w and the value of continuing search. Assume offers arrive at Poisson rate λ_0 during unemployment, so that the probability of receiving an offer any given period is $p = 1 - e^{-\lambda_0}$, and write f for the density of the offer distribution. Then

$$\begin{aligned} EV &= \frac{p}{1-\beta}\int_{(1-\beta)(b+\beta EV)}^{\infty} wfdw \\ &\quad + \left(1 - p\int_{(1-\beta)(b+EV)}^{\infty} fdw\right)(b+\beta EV) \\ &= \mathcal{T}(EV) \end{aligned} \tag{58}$$

with the derivative

$$\begin{aligned} \mathcal{T}'(EV) &= \beta\left(1 - p\int_{r}^{\infty} fdw\right) \\ &\in [0,\beta], \end{aligned} \tag{59}$$

and since $\beta < 1$ the operator $\mathcal{T}$ is a contraction and may be iterated to solve for EV and hence the optimal reservation wage strategy given by $r = (1-\beta)(b+\beta EV)$. Note the simplification: We are iterating on the scalar EV rather than the function V as in Section 3.

Christensen and Kiefer (1991) study the present model from a likelihood perspective. Consider panel data of the form $\{d_i, w_i\}_{i=1}^{N}$ where w_i is the accepted wage of the i'th initially unemployed worker and d_i the unemployment duration. Rapid information accumulation similar to that from the observed control rule in the marketing model in Section 5 occurs in the search model, too, although accumulation does not stop at a finite sample size. Similarly, preference parameters may be restricted to intervals, as in the marketing model in Section 4. For example, if β is the unknown parameter to be estimated in the job search model, the requirement $w_i \geq r$, $i = 1, ..., N$ imposes an interval restriction on β. In the search case, the interval keeps shrinking as $N \to \infty$, and the estimator

$$\widehat{\beta} = r^{-1}(w_m) \tag{60}$$

where $r = r(\beta)$ is inverted with respect to the parameter and $w_m = \min_i \{w_i\}$ is the minimal order statistic converges rapidly (at rate N, as opposed to the usual $N^{\frac{1}{2}}$) to the true value of the discount factor. More generally, for k parameters θ, including also b and parameters of f, Christensen and Kiefer (1991) show that

$$N^{\frac{1}{2}}\left(\widehat{\theta}-\theta_0\right) \rightarrow n_k\left(0, B_0\right), \tag{61}$$

$$rank \;\; B_0 = k-1, \tag{62}$$

a reduced rank limiting normal distribution, with asymptotic variance-covariance matrix B_0 and one parameter being superconsistent. Thus, the rapid information accumulation here takes the form of N-asymptotics in one direction of the parameter space. The control rule is learnt in finite time in Section 5 (discrete state/control) and at rate N in the search model.

Superconsistency here is a result of the fact that with complete observability, data must satisfy $w_m \geq r$, i.e. controls and states must line up such that no wage (state) is accepted (the control decision) unless the parameter dependent inequality is satisfied. This is a case of the curse of determinacy of Section 5. The strict requirement on data may be softened up by introducing measurement error, as in Section 6. This is done by Christensen and Kiefer (1994b) in an application to the 1986 Survey of Income and Program Participation (SIPP) data, and it is shown that regular (rank $B_0 = k$) $N^{\frac{1}{2}}$-asymptotics result. More general theory on information accumulation at different rates on subparameters may be found in Christensen and Kiefer (1994a, 2000), who introduce and study the concept of a local cut (w_m is a local cut in the search model).

Mortensen (1990) and Burdett and Mortensen (1998) introduced the equilibrium version of the search model, showing that in the simplest case it gives rise to an endogenous wage offer distribution with c.d.f.

$$F\left(w\right)=\frac{1+\lambda_1/\delta}{\lambda_1/\delta}\left(1-\left(\frac{q-w}{q-r}\right)^{1/2}\right), w \in [r,h], \tag{63}$$

where λ_1 is the offer arrival rate during employment (the case of on-the-job search), δ is an exogeneous lay-off rate, q is firm productivity and h is an upper bound on the wage distribution. Christensen and Kiefer (1997) determine the minimum panel data structure sufficient for identifying all structural parameters, in particular, $\{d_i, w_i, j_i\}_{i=1}^{N}$, accepted wages along with unemployment and employment durations, and show that both the minimum and maximum wages are local cuts, the reduced rank of the $N^{\frac{1}{2}}$-asymptotic normal distribution now being $k-2$. Again, regularity may be restored by allowing for measurement error, and this is done in an application to Danish data by Bunzel et al. (2001). Christensen et al. (2005) apply a related model with on-the-job search to a Danish panel of matched employer-employee data and use movers and stayers in a firm between two consecutive periods to estimate λ_1, δ and a cost of search parameter. The results imply that on-the-job search explains the employment effect, i.e. the extent of the stochastic dominance of the cross-section wage distribution of employed workers relative to the wage offer distribution.

11 Conclusion

Both statistics and economic theory provide ways to isolate to relevant portions of economic problems and data. Stochastic models are important for inference purposes, and the stochastic dynamic programming model is important when moving to the dynamic case. The sequential job search model of Mortensen (1970) and McCall (1970) is an important early application of dynamic programming in economics. By representing frictions as the random arrival of trading partners, the model is naturally stochastic and leads directly to a likelihood function. This is a key instance of useful data reduction through the productive combination of statistical analysis and theory of optimizing agents. Later theoretical developments starting with Mortensen (1990) allow imposing equilibrium constraints in the statistical analysis of optimizing agents.

In this paper, we have drawn attention to several issues in identification and inference in dynamic programming models which have received little or no prior notice, such as identification only up to intervals, the precise role of the optimality equation in identification, identification off the optimal path, rapid information accumulation, the curses of determinacy and degeneracy, and the dependence of parameter estimates on distributional assumptions in the random utility case. Along with the search model, a simple marketing model of investment in advertising campaigns affecting future demand has been used for illustration. Our discussion shows how earlier work on the econometrics of the search model has led to insights that apply to general dynamic programming models.

12 References

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

Bellman, R. and Dreyfus, S. (1962). *Applied Dynamic Programming.* Princeton University Press.

van den Berg, G.J. & Ridder, G. (1993). Estimating Equilibrium Search Models from Wage Data, in: H. Bunzel, P. Jensen & N.C. Westergård-Nielsen, eds., *Panel Data and Labour Market Dynamics.* New York: North-Holland, 43-55.

van den Berg, G.J. & Ridder, G. (1998). An Empirical Equilibrium Search Model of the Labor Market, *Econometrica,* 66-5, 1183-1221.

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.

Bunzel, H., Christensen, B.J., Jensen, P., Kiefer, N.M., Korsholm, L., Muus, L., Neumann, G.R. and Rosholm, M. (2001). Specification and Estimation of Equilibrium Search Models, *Review of Economic Dynamics,* 4, 90-126.

Burdett, K. and Mortensen, D.T. (1998). Wage Differentials, Employer Size, and Unemployment, *International Economic Review*, 39, 257-274.

Chintagunta, P.K. and Jain, D. (1992). A Dynamic Model of Channel Member Strategies for Marketing Expenditures, *Marketing Science,* 11, 168-188.

Chintagunta, P.K. and Vilcassim, N.J. (1992). An Empirical Investigation of Advertising Strategies in a Dynamic Duopoly, *Management Science,* 38, 1230-1244.

Chow, G.C. (1981). *Econometric Analysis by Control Methods.* New York: Wiley.

Christensen, B J. and Kiefer, N.M. (1991). The Exact Likelihood Function for an Empirical Job Search Model, *Econometric Theory*, 7, 464-486.

Christensen, B.J. and Kiefer, N.M. (1994a). Local Cuts and Separate Inference, *Scandinavian Journal of Statistics*, 21, 389-407.

Christensen, B.J. and Kiefer, N.M. (1994b). Measurement Error in the Prototypal Job-Search Model, *Journal of Labor Economics*, 12, 618-639.

Christensen, B.J. and Kiefer, N.M. (1997). Inference in Non-linear Panels with Partially Missing Observations: The Case of the Equilibrium Search Model. *Journal of Econometrics,* 79: 201-219.

Christensen, B.J. and Kiefer, N.M. (2000). Panel Data, Local Cuts, and Orthogeodesic Models, *Bernoulli*, 6, 667-678.

Christensen, B.J. and Kiefer, N.M. (2005). *Economic Modelling and Inference.* Princeton: Princeton University Press, forthcoming.

Christensen, B.J., Lentz, R., Mortensen, D.T., Neumann, G.R. and Werwatz, A. (2005). On-the-Job Search and the Wage Distribution, *Journal of Labor Economics,* 23, 31-58.

Kiefer, N. M. and Neumann, G. R. (1979). An Empirical Job Search Model with a Test of the Constant Reservation Wage Hypothesis, *Journal of Political Economy*, 87, 89-107.

Kiefer, N.M. and Neumann, G.R. (1993). Wage Dispersion with Homogeneity: The Empirical Equilibrium Search Model, in: H. Bunzel, P. Jensen & N.C. Westergård-Nielsen, eds., *Panel Data and Labour Market Dynamics.* New York: North-Holland, 57-74.

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

Lippman, S.A. and McCall, J.J. (1976b). The Economics of Job Search: A Survey: Part II, *Economic Inquiry*, 14, 347-368.

Maitra, A. (1968). Discounted Dynamic Programming in Compact Metric Spaces, *Sankhya Ser. A*, 40, 211-216.

McCall, J.J. (1970). Economics of Information and Job Search, *Quarterly Journal of Economics,* 84, 113-126.

McFadden, D. (1973). Conditional Logit Analyses of Qualitative Choice Behavior, in Zarembka, P. (ed.), *Frontiers in Econometrics.* New York: Academic Press, 105-142.

McFadden, D. (1981). Econometric Models of Probablistic Choice, in: C.F.Manski & D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications.* Cambridge: MIT Press.

Mortensen, D.T. (1970). Job Search, the Duration of Unemployment, and the Phillips Curve, *The American Economic Review,* 60-5, 847-862.

Mortensen, D.T. (1986). Job Search and Labor Market Analysis: in *Handbook of Labor Economics.* pp 849-919, O. Ashenfelter and R. Layard (eds.) Amsterdam: North Holland.

Mortensen, D.T. (1990). Equilibrium Wage Distributions: A Synthesis, in: J. Hartog, G. Ridder & J. Theeuwes, eds., *Panel Data and Labor Market Studies*, New York: North-Holland, 279-96.

Nerlove, M. and Arrow, K.J. (1962). Optimal Advertising Policy under Dynamic Conditions, *Economica,* 29, 129-142.

Ross, S.M. (1983). *Introduction to Stochastic Dynamic Programming.* New York: Academic Press.

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

Rust, J. (1997). Using randomization to Break the Curse of Dimensionality, *Econometrica,* 65-3, 487-516.

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

Working Papers published by the Department of Management

2005-1: Esben Koling Laustrup and Johannes Raaballe: Udbytteannonceringseffekten i Danmark.

2005-2: Carina Sponholtz: Separating the Stock Market's Reaction to Simultaneous Dividend and Earnings Announcements.

2005-3: Kyle Bagwell and Per Baltzer Overgaard: Look How Little I'm Advertising!

2005-4: Jerker Nilsson and Ole Øhlenschlæger Madsen: Cross-border Mergers Between Agricultural Co-operatives - a Governance Perspective.

2005-5: Bent Jesper Christensen and Morten Ørregaard Nielsen: The Effect of Long Memory in Volatility on Stock Market Fluctuations.

2005-6: Bent J. Christensen and Nicholas M. Kiefer: Investment in Advertising Campaigns and Search: Identification and Inference in Marketing and Dynamic Programming Models.