Optimal Destabilisation, Active Learning, and the Choice of Step Length in Policy Reform

Geoff Frewer*

NUMBER 265

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

UNIVERSITY OF WARWICK
COVENTRY

Optimal Destabilisation, Active Learning, and the Choice of Step Length in Policy Reform

Geoff Frewer*

NUMBER 265

July 1985

* The ideas for this paper originate from discussions with Kevin Roberts and Paul Warren. Thanks also to Morten Hviid, Norman Ireland, and Ken Wallis for their comments on an earlier draft, and to the participants in the Economic Theor Workshop.

ABSTRACT

A theoretical model is presented in which the decision marker's choice of step length for reform is allowed to depend on the consequences which this is expected to have for the accuracy of information about the structure of the economy. Also, destabilisation by means of introducing variance into policies, is desirable under some circumstances since it speeds up the learning process.

1. INTRODUCTION

The acknowledgement of ignorance frequently leads to the prescription of caution. With imperfect information about the parameters of the economic system risk averse economists often recommend piecemeal or marginal reforms instead of global optimisation. Rather than seek to identify the optimum on the basis of uncertain parameter estimates, a less ambitious problem is addressed, namely the directly of welfare-improving changes. A considerable literature has accumulated emphasising that limited movements towards the global optimum may not always be desirable (Lipsey and Lancaster 1956) but that some categories of second best reforms may be beneficial. (See for example Dixit 1975, Green 1962, and Bertrand and Vanek 1971). However, one issue which has received little attention, is that of step length: having identified the right direction how far should you go?

This paper emphasises the importance of information for the decision and allows the nature of the information revealed to depend on the policies chosen. Short steps are desirable in that their outcomes are relatively certain however, longer steps are riskier and hence their outcomes provide more new information.

In the model which follows the government is allowed two instruments. Policies over a period are assumed to be random drawings from a normal distribution whose mean and variance on the government's choice variables. Decisions about changes in the outcome determine the step length, and increases in variance may be interpreted as deliberate 'destabilisation'. The reason why it may be beneficial for the government to introduce uncertainty in the short run lies in learning process.

The passage of time alone would be sufficient to yield more observations, and given some variation in the data, this would produce more accurate parameter estimates; however it may be desirable to speed up the process by introducing policy variation in the short run in order to have a more diverse set of experiences upon which to base policies in the future. This idea is well-known in the control literature, Kendrick (1981) for example distinguishes between three kinds of otimal control strategies associated with different attitudes to learning. The least sophisticated is open-loop control which bases the future trajectory of instruments entirely on the data available at the starting period; no revisions are made in the light of new data and nothing is learnt. The second strategy involves passive learning and is called feedback control: parameter estimates are revised on the basis of new data but the development of instruments is independent of their expected effect on future parameter estimates. Active learning is the most sophisticated procedure and is an example of closed-loop control: policies are deliberately used to perturb the system in order more quickly to learn their influence on the state of the world.

The idea that current actions may influence future information is nothing new in economics: it is essential to concepts such as 'learning by doing', Arrow (1962); 'experimental consumption', Kihlstrom, Mirman, and Postlewaite (1984), and the 'Rothschild effect'. For example, Rothschild (1974) envisages a consumer faced with a two-armed bandit. He is uncertain about the expected payoffs from playing each arm so his actions are influenced by the inclination to play the arm which he believes to pay out most often, and also by the desire to improve his knowledge of the probabilities of winning. Because of sampling errors, the gambler may reach a long-run equilibrium where he plays only the arm with the lower payoff: the general equilibrium

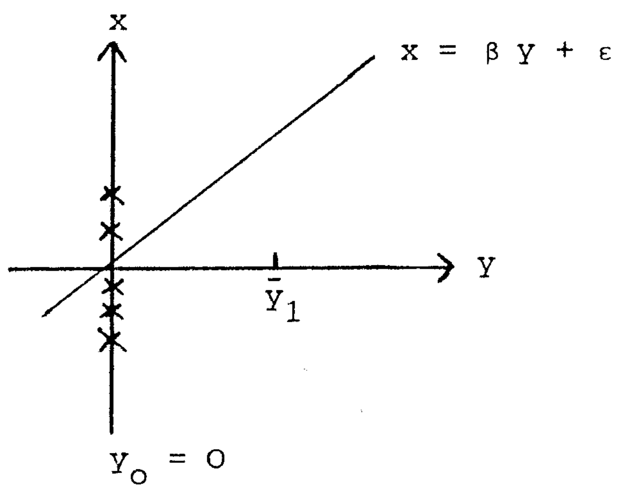
Grossman, Kihlstrom and Mirman (1977) present a closely related model. Faced with a new commodity whose characteristics are uncertain, the consumer experiments in the short run, possibly consuming more of it than he would if he were perfectly informed. My model is simpler because firstly I impose more restrictions on the objective function, and secondly I make sufficient assumptions so that OLS rather than Bayesian learning is appropriate. The main departure is that I allow the decision maker to introduce variance into the policy instrument, as well as choosing its central tendency.

Section 2 gives a preliminary outline of the main argument. Section 3 presents the model and discussion and possible extensions are in Sections 4 and 5. Appendix I gives a brief guide to interpretation of the derivatives of the objective function in relation to choice under uncertainty; and defines some of the concepts taken for granted in the text. Appendix II discusses the relationship between OLS learning and Bayesian updating.

2. PRELIMINARIES

The model in Section 3 envisages a decision maker who is uncertain about the influence of instruments on objectives. For tractability the simplest possible case will be considered, where the objective function may be approximated by a quadratic, and the uncertainty relates to the coefficient in a univariate regression model illustrated in Figure 1.

<u>Figure 1 - Choice of y and variance of β</u>



By examination of the objective function the decision maker knows the most preferred value of the state variable x but is not sure exactly what value for the instrument y will bring this about. This uncertainty results from the stochastic error term whose variance is assumed known, and also from errors in the estimation of β. If since the beginning of time y has ony taken the value zero, the spread of observations along the vertical axis will yield some information about the variance of ε, but this is useless information since σ is already known; however it will be uninformative about β, all lines passing through the origin would fit the data equally well. To learn something about β it would be necessary to either allow some variation around y_o, or fix y at some new level such as $\bar{y}_1$, or as is allowed in Section 3 employ a combination of both strategies.

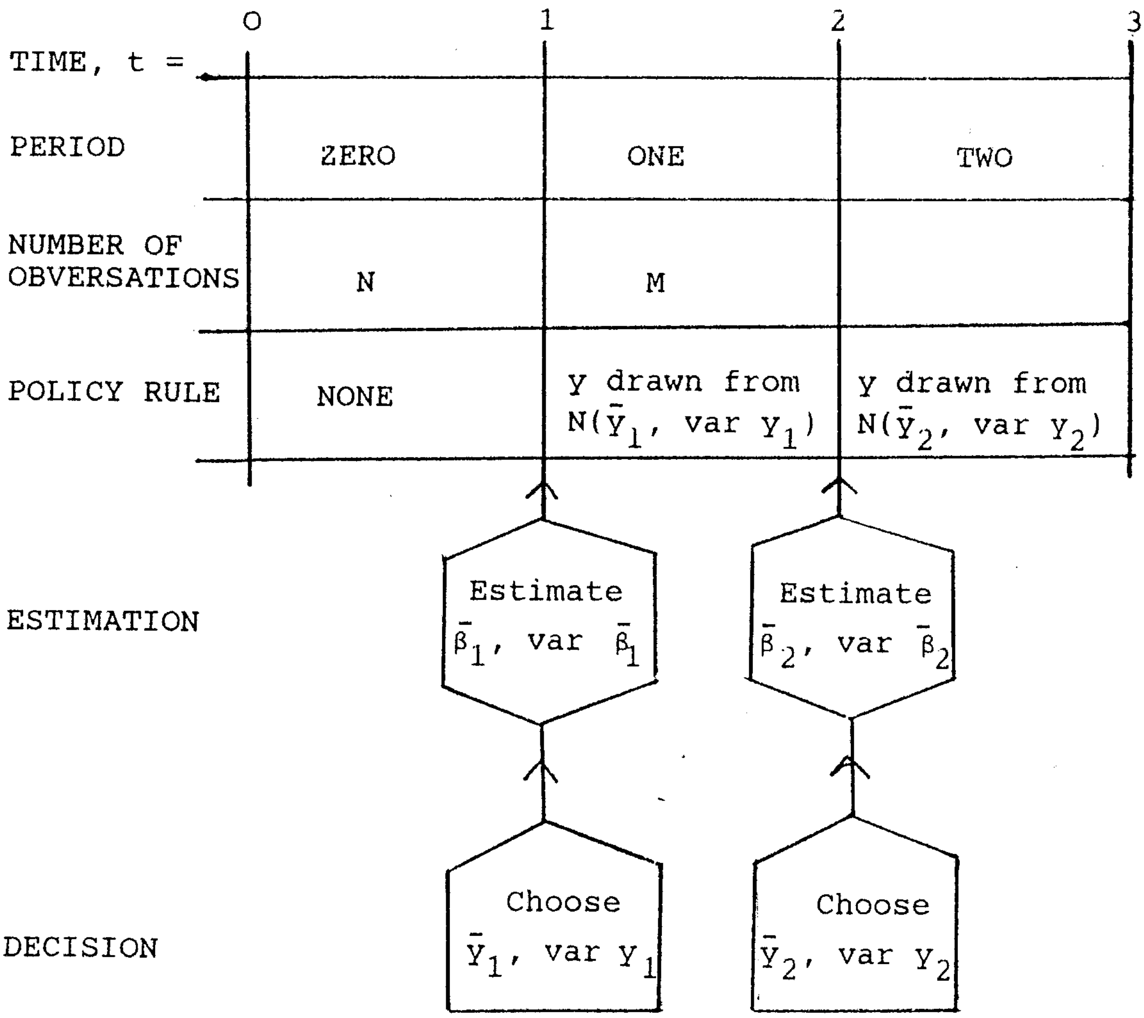
In this setting, there will be trade-offs between the use of policies for the purpose of learning, against aiming as closely as possible at the target value of x. More variation in y would lead to increased accuracy of parameter estimates, however it would also entail more variation in x and hence a higher probability in deviating from the target. Deviations from the status quo y_o are rewarded with greater data variation upon which to estimate β, but higher values of y also result in greater variance of x since the uncertainty is multiplicative.

Dynamic issues are obviously essential to the argument. The only reason for concern about the variance of parameter estimates is that they form the basis for future decisions. Choices based on more accurate information have a higher probability of hitting the target, so there will be conditions, relating to the discount rate and structural parameters, when it is beneficial to deliberately aim short of the target this period in order to have a better change of hitting it next period.

3. THE MODEL

The examination of active learning strategies involving endogenous information requires three time periods. Period zero has the sole purpose of establishing a data set upon which the decision-maker's prior beliefs are formed: no actions are taken until period one. In period two the actions are taken as if there were no learning, since $t = 3$ is the end of time there would be nothing to be gained by obtaining information which could never be used. The essential aspect of learning comes in period one when it is necessary to take account of the influence of current decisions on future information and hence on future policies and outcomes. The order of events is illustrated in Table 1.

<u>Table 1 - Active Learning : the Order of Events</u>



The problem faced in period one is to choose current policies to maximise the expected future stream of discounted utility, taking into account the influence of present policies on future information. The technique used here is dynamic programming: the 'fundamental recurrence relation' for this problem amounts to

$$\max_{\bar{y}_1,\ \mathrm{var}\ y_1} \quad J = EU_1 + \frac{1}{(1+\sigma)} EU_2^* \tag{1}$$

where star denotes maximised value. Thus the solution to the period one problem requires information about the choices that will be made in period two. The solution unfolds as follows: firstly, the period two problem is solved, secondly the optimal instruments are substituted into the objective function to obtain the maximised value EU_2^*, thirdly the period one criterion is established, and finally the first order conditions for this problem are derived and interpreted by means of an example.

The period two problem is to maximise the objective function, defined on the state variable, given that there is uncertainty about the influence of the instrument - that is, multiplicative uncertainty results from the inability to distinguish perfectly between those variations in x which are caused by variations in y, and those that result from the stochastic disturbance ε. The problem may be written

$$\begin{aligned} &\text{Max} \quad U(x_2) \\ &\text{s.t} \quad x_2 = \beta Y_2 + \varepsilon \end{aligned} \tag{2}$$

In the present context, uncertainty about ε and its variance σ are of marginal interest. σ is assumed known, and since the objective function is symmetric* it only plays the role of supporting uncertainty about β.

Taking the Taylor expansion of the objective function around $\hat{x}$

$$U(x_2) \simeq U(\hat{x}) + U_x(x_2-\hat{x}) + \frac{1}{2} U_{xx}(x_2-\hat{x})^2 \tag{3}$$

where $\hat{x}$ is defined by

$$\hat{x} = \bar{\beta}_1 \hat{Y} \tag{4}$$

Bars denote means, and hats points of certainty equivalence†, so $\hat{y}$ is the policy which would be chosen if $\bar{\beta}_1$ were known with certainty, and $\hat{x}$ is the outcome which would be expected if this choice were made. Notice that this implies

$$\frac{\partial U(\hat{x})}{\partial Y_2} \equiv U_2 = 0 \tag{5}$$

since under perfect certainty y would be chosen simply

---

* See Appendix I for a definition of this term, and some support for the assertion.

† See Appendix I for a definition of this term.

by setting U_x equal to zero.

Taking expectations of the last term in parenthesis on the right hand side of (3)

$$E(x_2-\hat{x})^2 = E(\beta y_2 + \varepsilon - \bar{\beta}_1\hat{y})^2$$

$$= E[(\beta y_2 - \beta\bar{y}_2) + (\beta\bar{y}_2 - \bar{\beta}_2\bar{y}_2) + (\bar{\beta}_2\bar{y}_2 - \bar{\beta}_1\bar{y}_1)$$

$$= \bar{\beta}_1^2 \text{ var } y_2 + \bar{y}_2^2 \text{ var } \bar{\beta}_2 + \bar{\beta}_1^2(\bar{y}_2 - \hat{y})^2 + \sigma \qquad (6)$$

To simplify this expression I have assumed firstly that covariances are zero, and secondly that the estimates of the mean of β formulated at the beginning of periods one and two are both unbiased so that $E(\bar{\beta}_1 - \bar{\beta}_2) = 0$.

Taking expectations of (3) and using (5) and (6)

$$EU_2 \simeq U(\hat{x}) + \frac{1}{2} U_{xx} [\bar{\beta}_1^2 \text{ var } y_2 + \bar{y}_2^2 \text{ var } \bar{\beta}_2$$

$$+ \bar{\beta}_1^2(\bar{y}_2 - \hat{y})^2 + \sigma] \qquad (7)$$

which is the maximand in period two. The first order conditions imply that

$$\text{var } y_2 = 0 \qquad (8)$$

$$\bar{y}_2 = \frac{\bar{\beta}_1^2\hat{y}}{(\text{var } \bar{\beta}_2 + \bar{\beta}_1^2)} \qquad (9)$$

It never pays to introduce policy variations in period two because the benefits would be zero in terms of possible future policy improvements and there are non-zero costs implied by risk aversion ($U_{xx} < 0$). (9) is the familiar static result: with symmetry and multiplicative uncertainty the policy maker aims low, and certainty equivalence is the outcome only if σ is zero, and hence β is known with certainty.

Substituting (8) and (9) into (7) and assuming that var $\bar{\beta}_2$ is small, the maximised value may be written

$$EU_2^* \simeq U(\hat{x}) + \frac{1}{2} U_{xx} \left[\frac{\text{var } \bar{\beta}_2 \bar{\beta}_1^2 \hat{Y}^2}{2 \text{ var } \beta_2 + \beta_1^2} + \sigma \right] \tag{10}$$

Notice that this is decreasing in var $\bar{\beta}_2$; given risk aversion there is positive value to new information because it reduces the variance of period two parameter estimates.

The period one problem may be approached in the same way. Taking a Taylor series expansion of $U(x_1)$ around $\hat{x}$

$$U(x_1) \simeq U(\hat{x}) + U_x(x_1 - \hat{x}) + \frac{1}{2} U_{xx}(x_1 - \hat{x})^2 \tag{11}$$

Note that

$$E(x_1 - \hat{x})^2 = \bar{\beta}_1^2 \text{ var } y_1 + \bar{y}_1^2 \text{ var } \bar{\beta}_1 + \bar{\beta}_1^2(\bar{y}_1 - \hat{y})^2 + \sigma \tag{12}$$

having again assumed zero covariances. Taking expectations of (27) using (5) and (12) yields

$$EU_1 \simeq U(\hat{x}) + \frac{1}{2} U_{xx}[\bar{\beta}_1^2 \text{ var } y_1 + \bar{y}_1^2 \text{ var } \bar{\beta}_1$$

$$+ \bar{\beta}_1^2(\bar{y}_1-\hat{y})^2 + \sigma] \qquad (13)$$

Substituting (19) and (13) into (1) the first order conditions for the period one problem may be derived

$$\frac{\partial J}{\partial \bar{y}_1} = \bar{y}_1 \text{ var } \bar{\beta}_1 + \bar{\beta}_1^2 (\bar{y}_1-\hat{y}) + \frac{1}{2}\frac{1}{(1+\delta)} \frac{\hat{y}^2}{(4 \text{ var } \bar{\beta}_2 + \bar{\beta}_1^2)} \frac{\partial \text{ var } \bar{\beta}_2}{\partial \bar{y}_1} = 0 \qquad (14)$$

$$\frac{\partial J}{\partial \text{ var } y_1} = \bar{\beta}_1^2 + \frac{1}{(1+\delta)} \frac{\hat{y}^2}{(4 \text{ var } \bar{\beta}_2 + \bar{\beta}_1^2)} \frac{\partial \text{ var } \bar{\beta}_2}{\partial \text{ var } y_1} = 0 \qquad (15)$$

To interpret these conditions consider the case of OLS learning*. At the beginning of period one OLS estimates of β are based on the N observations over period zero, and at the beginning of period two they are based on regressions pooling the existing N observations with the M new data points

---

* Appendix II discusses the relationship between OLS learning and Bayesian updating. A more fundamental problem arises since actions are chosen on the basis of parameter estimates based on observations which are themselves determined by the actions. Under these conditions there are likely to be no rational criteria for which OLS in the optimal estimation strategy. It has been suggested that the solution to this estimation problem lies in the econometrics of "variable parameters". Rather than aspiring to make a contribution to that debate I assume bounds on the econometric wisdom of the government and hope that OLS turns out to have properties which are not too undesirable.

which become available over period one. Thus

$$\text{var } \bar{\beta}_1 = \frac{\sigma}{N \text{ var } y_o} \tag{16}$$

$$\text{var } \bar{\beta}_2 = \frac{\sigma}{N \text{ var } y_o + M \text{ var } y_1 + \frac{NM}{(N+M)} (\bar{y}_o - y_1)^2} \tag{17}$$

Differentiating (17) with respect to vary$_1$ and $\bar{y}_1$, and substituting into (14) and (15) yields polynomials and it is impossible to write down neat formulae for the optimal values which they imply. However, the nature of the solution may be described by considering the limiting cases set out in table 2. The first four columns show what happens to (16), and (17), and their derivatives, and the last two are obtained by substituting into (14), and (15).

In Table 2, (*) denotes the solution to the quadratic

$$\begin{aligned} &\bar{\beta}_1^2 \, M \text{ var } y_1^2 + 2(2\sigma + \bar{\beta}_1^2 \frac{NM}{(N+M)} \bar{y}_o^2) \text{ var } y_1 \\ &+ \frac{N}{(N+M)} \bar{y}_o^2 (4\sigma + \bar{\beta}_1^2 \frac{NM}{(N+M)} \bar{y}_o^2) - \frac{1}{1+\delta} \frac{\hat{y}^2}{\bar{\beta}_1^2} = 0 \end{aligned} \tag{18}$$

and (**) the soltuion to

$$\bar{\beta}_1^2 \, M \text{ var } y_1{}^2 + 4\,\sigma \text{ var } y_1 - \frac{1}{1+\delta} \frac{\hat{y}}{\bar{\beta}_1^2} \sigma = 0 \tag{19}$$

bearing in mind the possibility of the corner solution since var $y_1 \geq 0$.

TABLE 2 - ACTIVE LEARNING MODEL : RESULTS IN LIMITING CASES

	var $\bar{\beta}_1$	var $\bar{\beta}_2$	$\frac{\partial \text{ var } \bar{\beta}_2}{\partial \text{ var } y_1}$	$\frac{\partial \text{ var } \bar{\beta}_2}{\partial \bar{y}_1}$	$\bar{y}_1$	var y_1
i) $\sigma \to o$, $N \to \infty$, var $y_o \to o$	0	0	0	0	$\hat{y}$	0
ii) $\sigma \to \infty$	∞	∞	∞	∞	0	0
iii) var $y_o \to o$	∞	$\frac{\sigma}{M \text{ var } y_1 + \frac{NM}{N+M}(\bar{y}_o - y_1)^2}$	$-\frac{\sigma M}{[M \text{ var } y_1 + \frac{NM}{N+M}(\bar{y}_o - \bar{y}_1)^2]^2}$	$\frac{2\sigma \frac{NM}{N+M}(\bar{y}_o - \bar{y}_1)}{[M \text{ var } y_1 + \frac{NM}{N+M}(\bar{y}_o - y_1)^2]^2}$	0	(*)
iv) $M \to o$	$\frac{\sigma}{N \text{ var } y_o}$	$\frac{\sigma}{N \text{ var } y_o}$	0	0	$\frac{\bar{\beta}_1^2 \hat{y}}{\frac{\sigma}{N \text{ var } y_o} + \bar{\beta}_1^2}$	0
v) $N \to o$	∞	$\frac{\sigma}{M \text{ var } y_1}$	$-\frac{\sigma M}{[M \text{ var } y_1]^2}$	0	0	(**)
vi) $M \to \infty$	$\frac{\sigma}{N \text{ var } y_o}$	0	$-\frac{\sigma}{\text{var } y_1}$	0	$\frac{\bar{\beta}_1^2 \hat{y}}{\frac{\sigma}{N \text{ var } y_o} + \bar{\beta}_1^2}$	$\frac{1}{1+\delta} \frac{\hat{y}^2 \sigma}{\bar{\beta}_1^4}$

Interpretation of the results in table 2 is fairly straightforward. Taking the rows in order of appearance:

i) The three limits, zero uncertainty about the strcutre of the economy, infinite past data, and infinite variance in past data, are all equivalent and all entail perfect information. Under these circumstances there is no reason to destabilise (var $y_1 = 0$) and the mean of $\bar{y}_1$ steps immediately to the first best optimum. Thus with perfect certainty there is no justification in the model for short reforming steps.

ii) As σ tends to infinity there is complete uncertainty, or total ignorance, nothing is known nor can anything be learnt. Therefore there is no rational basis for any policy. The reform of $\bar{y}$ is either zero or negative depending on whether $\bar{y}_0$ was zero or positive in the past.

iii) With zero variation in y_0 the prior estimate of β is diffuse - has infinite variance - and any choice of $\bar{y}_1$ other than zero would be infinitely risky. However, it still may be beneficial to introduce some policy variations in order to reduce the variance of future parameter estimates. From (18), a sufficient condition for non-zero policy variation is $\bar{y}_0 = 0$ and $\hat{y} > 0$.

iv) As the period of learning decreases, information acquisition becomes increasingly difficult and the model collapses to passive learning. No policy variations are introduced, and although the choice of $\bar{y}_1$ reflects uncertainty about β it takes no account of the expected influence on future parameter estimates. Assuming $\bar{y}_o = 0$, the step length of reform is inversely related to the variance of the estimate of β, and the reform is towards the certainty equivalent optimum.

v) The duration of period zero (N) determines the number of observations upon which prior estimates of β are based. As this duration tends to zero, the estimate $\bar{\beta}_1$ becomes diffuse and the reaction to this infinite multiplicative uncertainty is to choose $\bar{y}_1 = 0$. However, policy variation is still desirable since without it the state of total ignorance would last for ever. Inspection of (19) reveals that in this case destabilisation is generally desirable, provided that σ is not zero, and $\hat{y}$ and $\bar{\beta}_1$ are posivite.

vi) If learning is infinitely easy, the central tendency of y is deployed as if there were no learning; and non-zero policy variations are introduced to ensure perfect knowledge of the structure of the economy in the future.

As expected, these results support the proposition that optimal reforms will take the system closer to certainty equivalence, the smaller is the variance of parameter estimates; and depending on the starting point, more uncertainty will usually lead to smaller step lengths. Also, trial and error policies will be more important when there is little relevant past experience, and when learning is relatively easy. However, the novel feature of the solution is that there are circumstances when it is desirable to use the central tendency of the policy as though no new information were to be revealed, and to introduce variations about this point in order to benefit from active learning. This happens when learning is easy, or there is little evidence from the past about the efficacy of policies.

To get some idea of the magnitudes involves, some numerical examples are presented in figures 2 and 3. The calculations are made on the basis that $\sigma = \bar{\beta}_1 = 1$, there is no discounting ($\sigma = 0$), and the policy maker would, in the absence of uncertainty, like to increase y from the past value of zero to two. The figures show the optimal values for $\bar{y}_1$ and var y_1 (under different assumptions about N and M), for any variance in the past data. The vertical distance between the axis and the line $\bar{y}$ may be interpreted as the step length. The less variation there is in the period zero data, the shorter is the desired step length. Inspection of the line var y_1 shows that destabilisation will only be desirable when the variance of existing data is small, or learning is easy given the extant state of knowledge.

FIGURE 2

ACTIVE LEARNING NUMERICAL EXAMPLES

PARAMETERS $\sigma = \bar{\beta}_1 = 1;\ \bar{y}_o = \delta = 0;\ \hat{y} = 2$

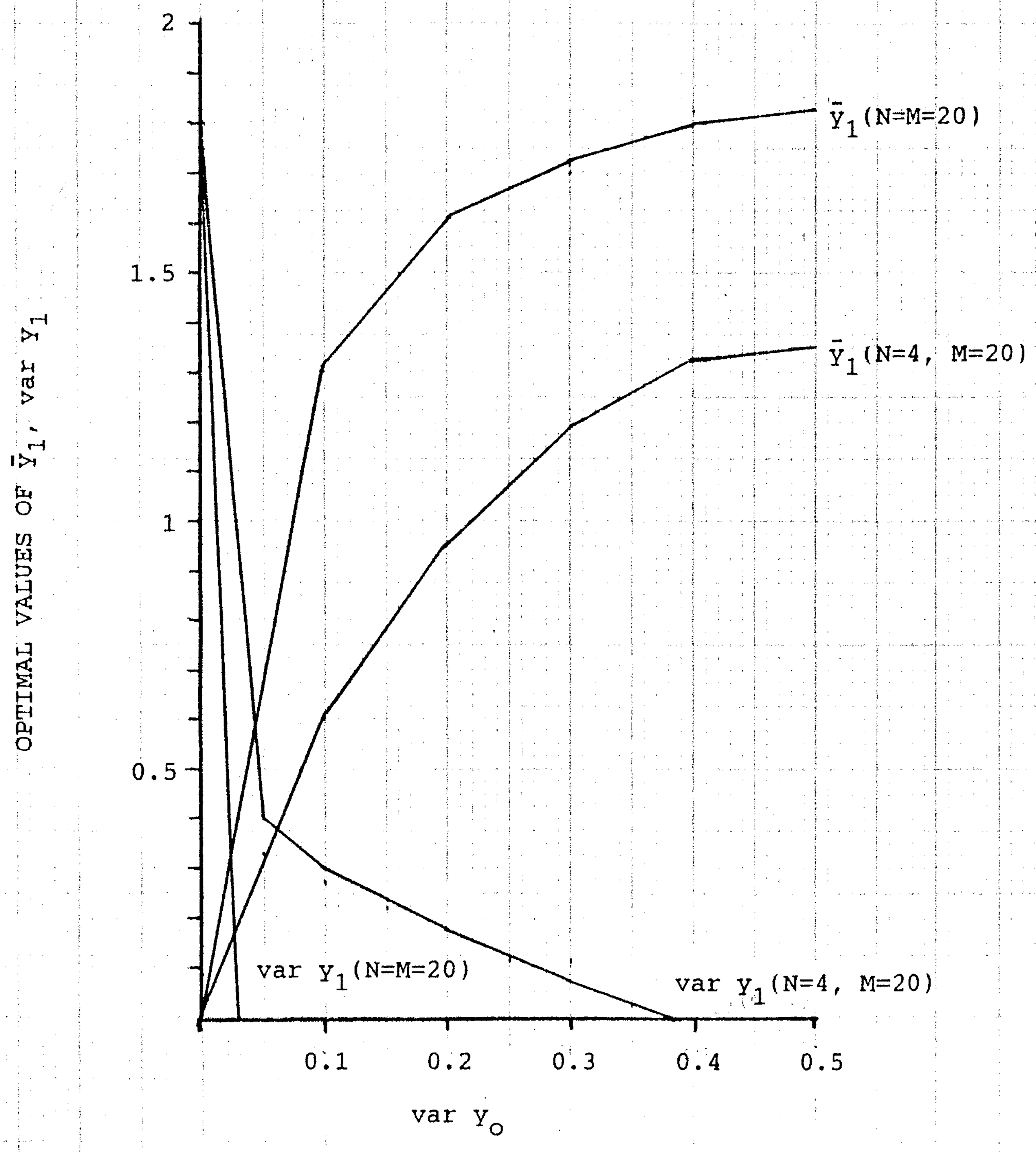
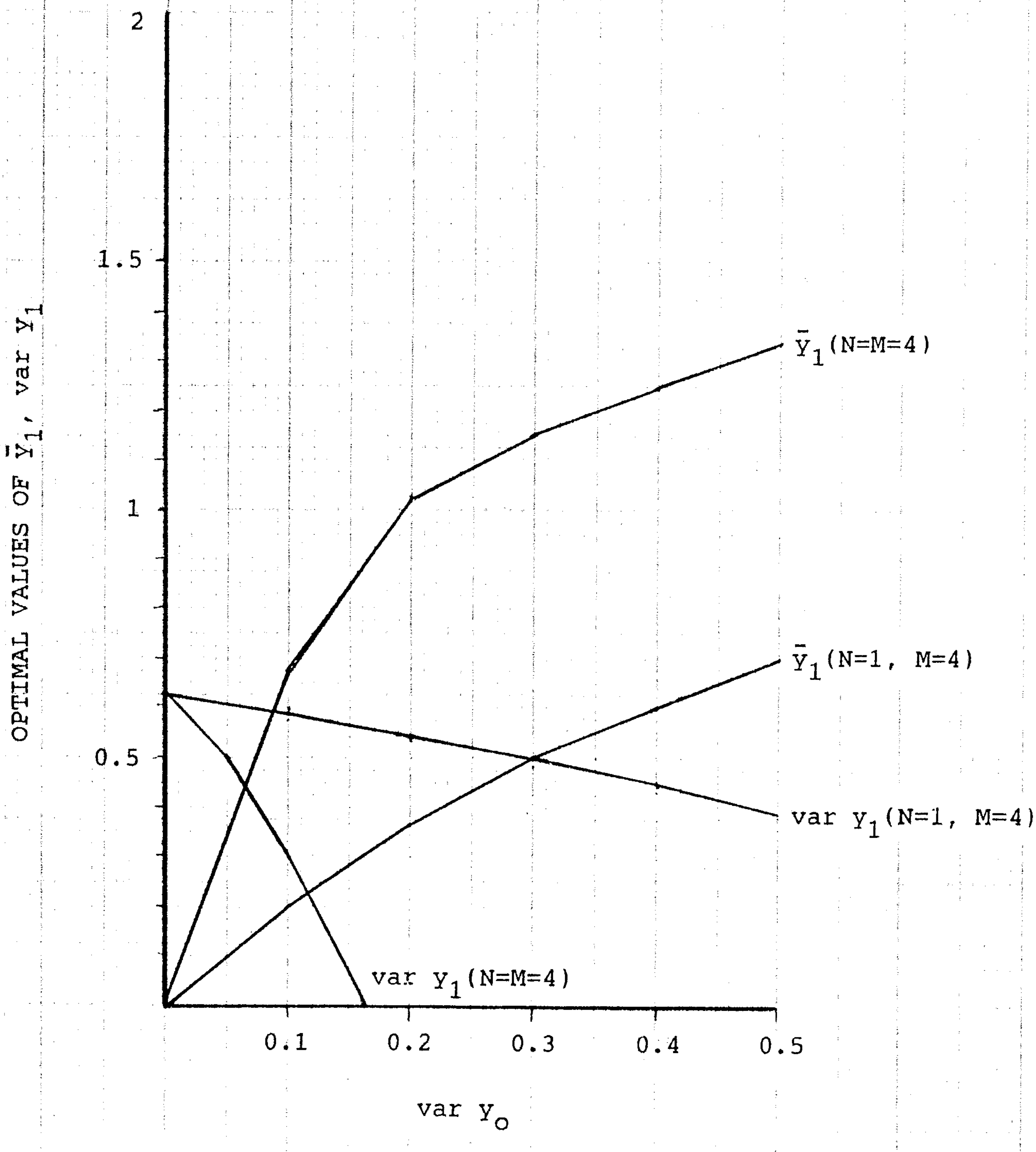


FIGURE 3

ACTIVE LEARNING NUMERICAL EXAMPLES

PARAMETERS $\sigma = \bar{\beta}_1 = 1;\ \bar{y}_o = \delta = 0;\ \hat{y} = 2$



4. DISCUSSION

The active learning strategy described in the preceeding model may be related to the underlying motivation of the tax reform literature. Atkinson and Stiglitz (1980) Chapter 12, Ahmad and Stern (1984), and Deaton (1985) discuss the idea that policy recommendations may be made on the basis of directions for marginal improvement rather than global optimisation. This approach is desirable in that it is informationally much less demanding, and the recommendations are likely to be more robust with respect to parameter uncertainty, and to the value judgements of the underlying social welfare function. However, some regard tax reform as essentially a strategy of passive learning.

Deaton (1985) expresses this view quite explicitly: inadequacies of the data mean that local approximations of parameters are all that can be obtained. Global optimisation under these circumstances is little more than a leap in the dark. Therefore the limitations of the data provide a strong justification for an incrementalist approach to policy changes: marginal reforms are made on the basis of local approximations, and then as information becomes available about the new locality further reforms can be proposed. This is clearly a passive learning strategy, new information is taken into account but the choice of instruments is independent of their expected effect on future knowledge.

Hence, it is possible to envisage an active learning tax reform strategy, where it pays to go further than marginal reforms in order to generate more useful data, and possibly all the way to optimisation treating local approximations as if they were globally true. Since "nature may not have been kind enough to perform the crucial experiments on our behalf", (Deaton, 1985), it may be necessary to take matters into our own hands.

In macroeconomics active learning may be regarded as a potential explanation for apparantly inconsistent behaviour on the part of both govrnment and private agents. A government would appear to be time inconsistent when in fact they were merely trying out an uncertain policy and revising it in the light of new information. Similarly private agents who in fact had unbiased estimates of the parameters of the economy would behave as though they were irrational in order to reduce their variances. Issues of convergence to rational expectations equilibria have been discussed with passive learning, see for example Bray (1982), but it is not clear in advance whether active learning would tend to stabilise or destabilise such models.

5. EXTENSIONS

A more flexible dynamic structure would be desirable for the further analysis of step length. The three period model attempts to capture the rest of eternity as a single final period. This necessarily precludes the examination of subsequent steps: the model asked how actions in the second period affect possibilities in the third but says nothing about the nature and direction of subsequent reforms. An obvious neglected question is convergence to some optimal solution by means of reforming steps . Moreover, if information were allowed to decay, and preferences shift over time, perhaps it would be possible to model a long-run steady state characterised by experimentation and learning.

The basic model in Section 3 may be revised in a more straightforward way to address other related issues.

i) Optimal Procrastination

Hold the variance of Y constant at a small level and allow the decision maker to choose M. There is no discounting, but as M increases the time remaining in period two is reduced. The problem then becomes one of deciding on the optimal duration of delay before a decision is made, given that this decision lasts for the rest of eternity. This is related to the irreversibility arguments of Henry (1974a) and (1974b), and Freixas and Laffont (1984): whereas they envisage an irreversible decision as constraining the value of the instrument to be non-increasing, or non-decreasing over time, I regard it as a decision that can never be altered.

ii) Irreversibility and Learning

Freixas and Laffont (1984) show that if not making as irreversible decision leads to an improvement in information then the result - that uncertainty makes it desirable to delay irreversible decisions - is reinforced. A more interesting case is suggested by the following example. There is uncertainty about nuclear technology, but better information can only be obtained by building nuclear power stations. However, when one is built it can never be demolished and must be maintained. Thus experimental consumption effects may tend to offset irreversibility, and it should be possible to derive conditions under which nuclear investments should be evaluated as if there were no uncertainty or learning.

APPENDIX I

BACKGROUND

Understanding of the active learning model may be enhanced by some discussion of the underlying concepts. These concepts relate firstly to properties of the objective function, and their interpretation in terms of response to risk, and secondly to the distinction between different kinds of uncertainty and their consequences for decision-making As a by-product this discussion also draws together various strands of the literature relating to choice under uncertainty.

Examples of applications to questions of interest in public finance are given.

In general the objective function U may be defined on the state variable x. There may be some overlap between state variables and instruments y, but often it is convenient to define U exclusively in terms of x. This broad class of functions may be approximated with a Taylor series expansion around some arbitrary $\tilde{x}$

$$U(x) \simeq U(\tilde{x}) + U_x(x-\tilde{x}) + \frac{1}{2} U_{xx}(x-\tilde{x})^2 + \frac{1}{6} U_{xxx}(x-\tilde{x})^3 \qquad \text{(A.I.1)}$$

The definition of $\tilde{x}$ allows some flexibility. In different contexts it is convenient to define it as one of the following: $\tilde{x} = 0$ yields an equation containing the moments of the distribution of x; $\tilde{x} = \hat{x}$ (the certainty equivalent value of x) allows the argument to be conducted in terms of moments, and deviations from certainty equivalence.

If the last two derivatives in (1) are insignificant then the function is locally <u>risk neutral</u>. It is well known in the economic texts that with linear utility functions the utility of the expected outcome is the same as the expectation of the sum of the utilities of the outcomes weighted by their probabilities of occurance. If the second derivative is negative the function is said to display <u>risk aversion</u>, and a decision maker with this kind of objective would not accept fair bets, the expected outcome always being prefered to the risky alternative with the same expected payoff.

The third derivative relates to the symmetry of objectives. If it is zero then the objective function is symmetric and a small shift from the optimum in either direction brings about the same decrement to utility. However, if U_{xxx} is negative (positive) then the marginal disutility of a small increase (decrease) in x above the optimum exceeds the marginal disutility of an equal reduction in x. Finally, a policy choice is said to be certainty equivalent if the same choice would be made under complete certainty. This term may also be applied to state variables: a certain equivalent state is the expected outcome of a certainty equivalent policy choice. Having defined these concepts the consequences of the two categories of uncertainty may be examined.

Additive uncertainty may be represented

$$x = y + \varepsilon \qquad \text{(A.I.2)}$$

where ε is normally distributed expectation zero and variance σ. Then necessary conditions for policy optimisation may be derived by substituting into (A.I.1), having defined $\tilde{x}$ as the mean of x, taking expectations and differentiating with respect to y

$$\frac{\partial EU}{\partial y} = U_y + \frac{1}{2} U_{yyy} \sigma = 0 \qquad \text{(A.I.3)}$$

Clearly, with additive uncertainty, symmetric objective functions lead to certainty equivalent policy choices:

$U_{yyy} = 0$ implies that y should be chosen such that $U_y = 0$. Thus Theil (1957), and Simon (1956) find that certainty equivalence follows from additive uncertainty with quadratic objectives, because these functions are symmetric. Clearly the direction in which the optimal policy would diverge from certainty equivalence depends on the sign of U_{yyy}. If it is positive (negative) the decision maker will prefer to aim high (low) since a small overshoot is preferred (inferior to) a small undershoot. Malinvaud (1969) interprets $U_{yyy} > 0$ as describing a good which is a necessity - its marginal utility decreases at a rapidly decreasing rate - and this may lead to increases in the consumption of a good as it becomes riskier. Hahn (1970) for example, discusses this proposition in the context of savings. For a public investment interpretation see Example 1.

Multiplicative uncertainty may be represented

$$x = y(1 + \eta) \tag{A.I.4}$$

where η has expectation zero and variance s. In this case, substitution into (A.I.1), taking expectations and differentiating yields the optimal decision rule

$$\frac{\partial EU}{\partial y} = U_y + U_{yy}ys + \frac{1}{2}\,U_{yyy}y^2 s = 0 \tag{A.I.5}$$

If the objective function is symmetric the last term is zero but even so certainty equivalence does not apply. Inspection if (A.I.4) reveals that the variance of x depends on the choice of y, and when the decision maker is risk averse this dependence will induce him to aim low, i.e. choose a value for the instrument below the certainty equivalent level. Hence multiplicative undertainty with risk aversion means that smaller y's tend to be preferred since they are associated with smaller variance in the outcome. This is the algebraic argument underlying Waud (1976)'s diagramatic discussion of additive and multiplicative uncertainty. For a simple model of intertemporal public investment decisions with uncertain rates of return, see Example 2.

<u>EXAMPLE 1</u>

Choice of expenditure on risky public investments, with additive uncertainty and asymmetric objectives.

The government has to decide how to allocate a fixed budget m between expenditure on conventional and nuclear power stations, y_1 and y_2. The two kinds of electricity x_1 and x_2 which they produce have different characteristics (in terms of the ability to respond to transient changes in demand) so the objective function depends on both, $U(x_1, x_2)$. Conventional technology is certain but nuclear is subject to additive risk: attention is restricted to those uncertainties, such as returns to research and development expenditure, which are independent of the scale of investment

$$x_1 = y_1 \qquad\qquad (A.I.6)$$
$$x_2 = y_2 + \varepsilon$$

where ε is normally distributed with non-zero and variance σ. Using the budget constraint $m = y_1 + y_2$, it is possible to write the governments objective function in terms of the choice variable y_1 and the unknown state ε.

$$U(x_1, x_2) = U(y_1, m - y_1 + \varepsilon) \equiv w(y_1, \varepsilon) \qquad (A.I.7)$$

Taking a Taylor expansion of w around the expectation of ε

$$w \simeq w(y_1, 0) + \frac{1}{2} w_{\varepsilon\varepsilon} \sigma \qquad (A.I.8)$$

Differentiating with respect to y_1 yields the first order condition

$$\frac{\partial w}{\partial y_1} = w_1 + \frac{1}{2} w_{\varepsilon\varepsilon 1} \sigma = 0 \qquad (A.I.9)$$

which implicitly defines the optimal choice y_1^*. Since at a maximum $w_{11} < 0$, inspection of (A.I.9) reveals that

$$\frac{\partial y_1^*}{\partial \sigma} \gtreqless 0 \quad \text{if} \quad w_{\varepsilon\varepsilon 1} \gtreqless 0 \qquad (A.I.10)$$

which is a particular case of Diamond and Stiglitz (1974) theorem 1. From (A.I.7), and using Youngs theorem,

$$w_{\varepsilon\varepsilon 1} = w_{1\varepsilon\varepsilon} = U_{122} - U_{222} \qquad \text{(A.I.11)}$$

Now, if U_{222} is positive and U_{122} is small or negative, $w_{1\varepsilon\varepsilon}$ will be negative and the choice of y_1 will be decreasing in σ. To interpret these third derivatives consider the rates of change of the marginal utilities of x_1 and x_2 with increases in the consumption of x_2 illustrated in Figure A.I.1

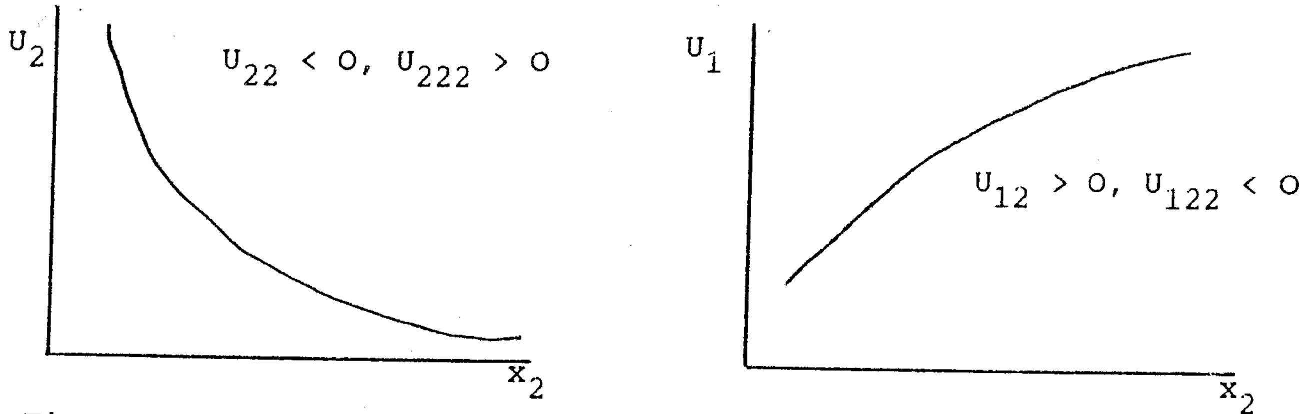


Figure AI.1 Third Derivatives, x_2 a necessity

The signs of these derivatives are consistent with the proposition that x_2 is a necessity: its own marginal utility diminishes at increasing rate; and the marginal utility of x_1 increases with consumption of x_2 but at a diminishing rate. Thus consumption of x_2 in excess of some necessary level adds little to utility, and also the satisfaction resulting from the consumption of other good increases with consumption of the necessity but at a decreasing rate.

Hence, this example shows that if nuclear power is a necessity, reductions in the uncertainties of nuclear technology should lead to a transfer of resources towards conventional generating capacity. The intuition behind this result is that given the definition of necessity, the consequences of having too little nuclear power is disasterous, therefore uncertainty leads the decision maker to aim high. As a result, reductions in uncertainty make it possible to reduce nuclear investment and still be as sure of reaching that target.

EXAMPLE 2

An inter-temporal public investment decision with risky returns, in a model with symmetric objectives, multiplicative uncertainty, and dynamic programming.

The model consists of two periods, in period one the decision maker chooses consumption and investment on the wealth endowment m. The rate of return on investment r is uncertain, and determines the amount available for consumption in the second period. A dynamic programming solution (see Intriligator, 1971, and Kamien and Schwartz, 1981) is presented (although the problem is simple enough for other methods to be applicable) since familiarity with this technique is required for the active learning model in Section 3. Using a Taylor series expansion around zero, and letting $U(0) = 0$, the symmetric objective function at time t may be written

$$U(x_t) = U_x(x_t) + \frac{1}{2} U_{xx}(x_t)^2 \qquad (A.I.12)$$

Obviously, the optimal policy in period two is to consume everything

$$x_2 = (1+r+\eta)\ (m-y_1) \tag{A.I.13}$$

where r is the rate of interest, and η is a stochastic disturbance with mean zero and variance s. By substitution into (A.I.11) the expectation of the maximised value for U_2 is

$$EU_2^* = U_x(1+r)(m-y_1) + \frac{1}{2} U_{xx}[(1+r)^2 + s](m-y_1)^2 \tag{A.I.14}$$

The period one problem is then to maximise present utility, taking account of the influence of current decisions on future opportunities.

$$\max_{y_1}\ J = EU_1 + \frac{1}{1+\delta} EU_2^* \tag{A.I.15}$$

Where δ is the rate of time preference. The first order condition for this problem implies

$$y_1 = \frac{U_x(\delta - r) + U_{xx}[(1+r)^2 + \delta]m}{U_{xx}[(1+\delta) + (1+r)^2 + s]} \tag{A.I.16}$$

The approximation entailed by (A.I.12) is only locally applicable, and certainly breaks down if $U_x + U_{xx}m > 0$ since this would mean a negative marginal utility of money. The expectation of y_2 is given by the budget constraint

$$Ey_2 = (1+r)(m-y_1) \qquad (A.I.17)$$

Thus from (A.I.16) it is possible to show that

$$\delta - s \frac{U_{xx}m}{U_x} > r \Rightarrow y_1 > Ey_2 \qquad (A.I.18)$$

With no uncertainty, if the rate of time preference is greater than the rate of interest then more of the budget will be allocated to current rather than future consumption. The more uncertain is the rate of return on investment and the larger is $-U_{xx}$, the greater is the tendency to plan to consume more at present than in the future.

APPENDIX II

BAYESIAN LEARNING

The following is based on Zellner (1971). The conditions under which OLS is equivalent to maximum likelihood estimation are well known: the errors ε must be normally independently distributed with mean zero and common variance σ , and the independent variable x, if stochastic, must be independent of ε , with distribution not involving ε or the parameter β . Therefore it is sufficient to show that maximum likelihood estimation on the pooled data $d_0 + d_1$ is equivalent to maximum likelihood on d_0 followed by Bayesian updating with d_1. Since parameter estimates come from the derivatives of the likehood functions an appropriate simplification is to consider proportionally rather than equality throughout.

The model may be presented

$$x_t = y_t\beta + \varepsilon_t \qquad \begin{array}{l} t = 1, \ldots, N \text{ in } d_o \\ t = N+1, \ldots, N+M \text{ in } d_1. \end{array} \qquad \text{(A.II.1)}$$

Given the normality of the errors the prior probability density function (given d_o) is

$$P_r(d_o;\ \beta, \sigma) \propto \frac{1}{\sqrt{\sigma^{N+1}}} \exp\left[-\frac{1}{2\sigma}(x_o - y_o\beta)'(x_o - y_o\beta)\right] \qquad \text{(A.II.2)}$$

where subscripts denote the partition corresponding to d_o and d_1. The likelihood of β, σ given the new data in d_1 is given by

$$L(\beta, \sigma;\ d_1) \propto \frac{1}{\sqrt{\sigma^{M}}} \exp\left[-\frac{1}{2\sigma}(x_1 - y_1\beta)'(x_1 - y_1\beta)\right] \qquad \text{(A.II.3)}$$

The posterior is proportional to the product of the prior and the likelihood

$$Pr(d_1, d_o;\ \beta, \sigma) \propto \frac{1}{\sqrt{\sigma^{N+M+1}}} \exp\left[-\frac{1}{2\sigma}\{(x_o - y_o\beta)'(x_o - y_o\beta) + (x_1 - y_1\beta)'(x_1 - y_1\beta)\}\right] \qquad \text{(A.II.4)}$$

This forms the basis of the log likelihood function from which the parameter estimates are derived.

The alternative is to pool the data giving the likelihood function

$$L(\beta, \sigma;\ d_o,\ d_1) \propto \frac{1}{\sqrt{\sigma N+M}} \exp\left[-\frac{1}{2\sigma}(x_o - Y_o\beta)'(x_o - Y_o\beta) + (x_1 - Y_1\beta)'(x_1 - Y_1\beta)\right] \qquad \text{(A.II.5)}$$

and taking a diffuse prior

$$\Pr(\beta, \sigma) \propto \frac{1}{\sqrt{\sigma}} \qquad \begin{array}{c} -\infty < \beta < \infty \\ 0 < \sigma < \infty \end{array} \qquad \text{(A.II.6)}$$

Then the posterior probability density function upon which estimates are formed is identical to (A.II.4).

References

Ahmad, E. and Stern, N., (1984), 'The Theory of Tax Reform and Indian Indirect Taxes', J. Pub E., Vol.25, No.3, December 1984, p259-298.

Arrow, K., (1962), The Economic Implications of Learning by Doing, RES 29, p.155-173.

Atkinson, A.B. and Stightz, J.E., (1980), Lecture on Public Economics, McGraw Hill.

Bertrand, T.J., and Vanek, J., (1971), 'The Theory of Tariffs, Taxes, and Subsidies: Some Aspects of the Second Best', AER, P.925-931.

Boyer, B. and Kihlstrom, R.E., eds., (1984), Bayesian Models in Economic Theory, North-Holland.

Bray, M.M., (1982), 'Learning, Estimation, and the Stability of Rational Expectations Equilibrium', Journal of Economic Theory, Vol.26, p.318-39.

DeGroot, M.H., (1970), Optimal Statistical Decisions, McGraw-Hill, U.S.A.

Deaton, A., (1985), Economic Issues of Tax Design in Developing Countries, in Newbery and Stern, eds., The Theory of Taxation for Developing Countries, World Bank, forthcoming.

Diamond, P.A. and Stiglitz, J.E., (1974), 'Increases in Risk and Risk Aversion', Journal of Economic Theory 8, p.337-360.

Dixit, A., (1975), 'Welfare Effects of Tax and Price Changes', Journal of Public Economics 4, p.103-123.

Freixax, X. and Laffont, J.J., (1984), On the Irreversibility Effect, Ch.7 in Boyer and Kihlstrom, eds., Bayesian Models in Economic Theory.

Freixax, X., Guesnerie, R., and Tirole, J., (1985), Planning Under Incomplete Information and the Ratchet Effect, RES Vol.II(2), No.169, April.

Green, H.A.J., (1962), 'The Social Optimum in the Presence of Monopoly and Taxation', RES, Vol.29, p.66-78.

Grossman, S.J., Kihlstron, R.E., and Mirman, L.J.M., (1977), A Bayesian Approach to the Production of Information and Learning by Doing, Review of Economic Studies, 44, p.633-47.

Hahn, F.H., (1970), Savings and Uncertainty, RES Vol.XXXVII (1), No.109, p.21-24, January.

Henry, C., (1974a), Option Value in the Economics of Irreplacable Assets, RES Symposium, p.89-104.

Henry, C., (1974b), Investment Decisions Under Uncertainty: The Irreversibility Effect, AER, Vol.66, No.6, p.1006-12, December.

Intrilligator, M.D., (1971), Mathematical Optimisation and Economic Theory, Prentice Hall.

Kihlstrom, R.E., Mirman, L.J., and Postlewaite, A., (1984), Experimental Consumption and the "Rothschild Effect", in Boyer and Kihlstrom, eds., Bayesian Models in Economic Theory.

Kendrick, (1981), Stochastic Control for Economic Models, McGraw Hill.

Kamien, M.I., and Schwartz, N.L., (1981), Dynamic Optimisation: The Calculus of Variations, and optimal control in Economics and Management, North Holland.

Levhari, D. and Srinivasan, T.N., (1969), 'Optimal Savings under Uncertainty', RES, Vol.XXXVI (2), No.106, p.153-164, April.

Lipsey, R.G. and Lancaster, K., (1956), 'The General Theory of the Second Best', Review of Economic Studies 24, p.11-32.

Malinvaud, E., (1969), First Order Certainty Equivalence, Econometrica, Vol.37, No.4, p.706-718, October.

Rothschild, M., (1974), A Two-Armed Bandit Theory of Market Pricing, Journal of Economic Theory 9, 185-202.

Simon, H.A., (1956), Dynamic Programming Under Uncertainty with a Quadratic Criterion Function, Econometrica, Vol.24.

Theil, H., (1957), A Note on Certainty Equivalence in Dynamic Planning, Econometrica, Vol.25.

Waud, R.N., (1976), Asymmetric Policy Maker Utility Functions and Optimal Policy Under Uncertainty, Econometrica, Vol.44, No.1., January.

Zellner, A., (1971), An Introduction to Bayesian Inference in Econometrics, John Wiley and Sons, New York.