ACTIVE LEARNING, CONTROL THEORY AND AGRICULTURAL POLICY

Gordon C. Rausser
Harvard University

In recent years the potential variability of the world food system has become increasingly obvious. Especially during 1972-73 the magnitude of increases in farm product and food prices surprised most everyone within the public as well as the private sectors. To U.S. government officials who were struggling to contain inflation, especially in the administered price sectors of the economy, this tremendous increase in food prices was indeed a bitter disappointment. Moreover, from the standpoint of existing price forecasting models, it became clear that current perceptions and methods of predicting food price system events were no longer viable.

Correspondingly, our perceptions have undergone change and are in the general state of flux. Due to the increasing complexity within the system, it is indeed difficult to form precise perceptions on the basis of intuition. This lack of intuitive understanding of some outcomes resulting from interactions within the system makes it clear that what in the short run may seem to be an unqualified desirable policy may lead in the long run to undesirable, even deleterious results.

In the above setting, it remains crucial to isolate the basic nature of the agricultural and food system. Even in the face of changes which have occurred simultaneously and in magnitudes never before perceived, conventional wisdom seems intact. Generally, conventional wisdom characterizes the system by (Brandow [1976]) highly inelastic aggregate demand; low income elasticity of demand; rapid technological change; asset fixity; atomistic structure of the production sector; life cycles of plant and animal growth;

Presented at AAEA/WAEA joint meeting, San Diego, July 31- Aug. 3, 1977.

the growing nature of inventories; climatic and weather uncertainties; labor immobility; and the demand for and the propensity of governments to actively intervene in the agricultural and food systems. One of the principal implications of these characteristics is instability. A policy frequently advanced for dealing with instability is inventory or buffer stocks. Such policies can assume many different forms and their effects are highly uncertain (Rausser [1977]). Other instruments of government intervention are trade oriented; they include export subsidies, export controls, foreign aid, import tarriffs and quotas, concessional sales, and efforts to liberalize trade relations. The last set of policies often applied to food and agricultural systems are chiefly oriented toward production and include regulation of output prices, quantity control, input controls, and input taxes and subsidies. Here, due largely to uncertainty, the issue of price versus quantity controls becomes a nontrivial problem (Rausser and Riboud [1977]).

In assessing each of the above policies, a number of critical uncertainties arise for which little in the way of empirical evidence has been accumulated. For example, in the context of trade policies, is the paradigm offered by neoclassical theory sufficiently robust or must governmental behavior and resulting trade distortions be introduced endogenously? In the case of buffer stock policies, the nature of the relevant demand and supply relationships become critical. The distribution of gains and losses from price stabilization can be drastically altered by various specifications of these relationships (Rausser [1977]). Moreover, the risk levels within various commodity systems and the distribution of risk and their effects on behavior have certainly not been precisely quantified. Much remains to be learned about the equity effects of such policies along both qualitative and quantitative dimensions. In a more general sense, to provide useful policy assessments, much remains to be learned about the (i) nature of structural change, (ii) parameter variation, and (iii) expectation formation patterns of various participants in food and agricultural systems.

Over the last few decades, a number of methods have been developed to assist us in the assessment of evaluation of alternative policies. Some of the more interesting methods emanate from the work of both electrical engineers and economists on control theory. As early as 1966, Fox et al provided much direction to the development and application of these methods by economists. Subsequently, the American Agricultural Economics Association held a special session at the Winter Meetings in 1969 on the applicability of control theory to agricultural policy and decision making. Emphasis here was on deterministic control theory which has been widely applied to various agricultural decision problems. These methods have been extended to stochastic as well as adaptive control formulations. Such formulations have been emphasized in the six conferences held by the National Bureau of Economic Research over the period 1972-77. An indication of the apparent value of these methods is reflected a glowing article which appeared in the May 1973, *Business Week*:

> "Control theory has swept into the economic profession so rapidly in the past two or three years that most economists are dimly aware that it is around. But for econometricians and mathematical economists and for the companys and government agencies that use their skills, it promises an improved ability to manage short run economic growth investment portfolios and corporate cash positions."

Although these methods have not achieved the degree of success implied by this report, its potential as a supporting device for governmental and private decision making remains promising. In the case of agriculture, this potential is documented in the recent volume on the survey of quantitative methods in agricultural economics (Judge, et. al. [1977]). The work surveyed in this volume emphasizes the importance of fully integrating the array of various quantitative approaches to problem solving and testing of theoretical constructs. The editors of this volume, Judge et. al. [1977] note that the compartmentalization typical within the agricultural economics profession seems increasingly more artifical as the quantitative methods and approaches to data generation are viewed in their generalized forms. Econometric methods, optimization, systems and simulation or more pragmatic data collection efforts should not be viewed as separable tasks. Instead, problem oriented research requires that these apparently separate tasks and associated methods be more fully and effectively integrated. One framework that provides a consistent approach to this integration is provided by adaptive control theory. The adaptive control formulation includes as special cases the deterministic as well as stochastic versions of control theory.

From a pragmatic standpoint, the adaptive control framework recognizes that as a system progresses through decision periods more data become available with which to update or revise estimates of the influence of alternative control variable settings on various performance measures. In general, revisions of system representations should not be regarded as separate from the derivation of an optimal policy. To be sure it is possible that various decisions may reveal more or less information about the actual system via different sets of the resulting data obtained. The inherent benefits of the additional information depends upon whether or not an "improved representation" of the structure results in superior future control. The incurred cost of such information emanates, in part, from choosing a current decision which is less than optimal from the pure control point of view. In a concrete context, as Dreze [1972, page 15] suggested "a monopolist may wish to depart from the price which maximizes expected profit, simply to learn more about his demand function."

The adaptive control approach to economic policy corresponds to Bellman's [1961, pages 198-209] as well as Feldbaum's [1965, pages 24-31] third class of control systems. This class is characterized by some unknown quantities about which uncertainty changes by learning as the process evolves. This class includes the active accumulation of information, i.e., the accumulation of information does not take place independently of the control process. In fact, optimal adaptive controls require a simultaneous solution to a combined control and sequential design of experiments problem and thus are dual in nature. This dual feature of the adaptive control framework was first recognized by Feldbaum [1965]. Early examinations of the dual control technique advanced by Feldbaum are available in Sworder [1966], Aoki [19~~76~~ 67], and Meier [1966]. From this work, the anatomy of the dual control approach may be characterized by three principle elements, viz., direct control, learning, and design of experiments.

Neglecting the relationship between current information and future measurements, the control element takes into account the direct effect of decisions on the criteria function. These effects may include values of the decision variables themselves, and through their influence on the system, their indirect effects on current and future values of the state or performance measures. The learning element is motivated by the existence of a set of sufficient statistics which are conditional on information related to the current state of the system and on the most recent estimate of the probability distribution function for the unknown of the decision problem. With respect to the latter, as we proceed into the future, additional sample information becomes available. The use of these sample additions allows learning to take place regarding unknown parameters, unknown states, and the like.[1]

Adaptive control strategies may be represented as conditional functions depending, in part, on the moments of the probability distribution function for the unknown components of the decision problem. The measurement of the criteria function resulting from these strategies is in general improved the more concentrated the probability distribution functions are about their expected values. As noted above, via some learning process the properties of these probability distribution functions depend upon available sample information. The sample information in turn can be influenced by alternative settings on current controls. Since the generation of such sample information has direct bearing on future estimates of the probability distribution functions and these in turn influence the performance of alternative decisions, an experimental dimension is involved in current adaptive control actions.

Unfortunately, it has not yet been possible to express the adaptive of dual controls solutions in analytical form. Although numerous authors present the relevant recursive equations (see for example Feldbaum [1965] and Aoki [1967]), these equations do not yield explicit solutions. For this reason, much in the way of research effort has been directed toward the development of analytical approximation and numerical feasible adaptive control rules. The performance of these approximate rules are usually tested by simulations which are problem specific. Since the first computationally feasible algorithms for general adaptive control systems were constructed only in the last few years, the total number of applications in economics and related fields is rather small.

A significant issue is whether the systems encountered in agricultural and natural resource economics can indeed benefit from an adaptive approach which induces learning about its parameters.[2] There seem to be three general cases where the importance of adaptive control is apparent. The first is when a new system is introduced and its parameters are generally unknown. Examples would be the demand equation for a new product; the advertising-sales reaction for a new advertising copy of a new product; employee reactions to a new incentive scheme; the production function of a new technology, say a new seed variety or the production of gas from animal wastes; or the effectiveness of a new capital device in tapping existing natural resources or controlling the adverse (environmental) effects of natural resource utilization. The second case is where the system is established, but the parameters tend to vary over time. In order to ensure an acceptable control performance these parameters should be monitored and actively learned.

The third category, technically viewed as a subset of the first two, is where improvement of the model representing the system is sought. In particular, if two candidate models are competing, and the control is determined on the basis of a composite model formed by a convex combination of the two, the weighted scheme can be modified continuously as a result of learning. If the two models are nested, the discrimination between the two may be improved by adaptive control and hence the performance of the system may be enhanced.

In what follows, the adaptive control framework is examined by first formulating the general problem and defining the associated information elements (section 2). In particular, control policies are classified with regard to the information usage. This classification, although not exhaustive allows various adaptive control solution approximations to be generated. In order to motivate the need for these approximations, section 3 presents a simple two period problem to illustrate the inherent analytical difficulties in obtaining solutions. This problem is also utilized to demonstrate the distinction between *passive learning* and *active learning*. Schemes based on the former approach are briefly surveyed; they admit only accidental learning and thus do not involve an experimental dimension. The various actively adaptive schemes are then presented; they represent a generalization of what Raiffa and Schlaifer [1961] have referred to as "preposterior analysis." For these schemes, the value of information is anticipated by using statistics of future measurements via algorithms based on the notion of *closed-loop* control. Such control schemes involve an experimental dimension which probes the system in anticipation of the value of information to be derived from future observations. This section also records the available results on the comparative performance of both passive and active schemes. Finally, a concluding section presents a brief overview of past and current research efforts as well as those issues of significance to the further development of adaptive control methods.

2. GENERAL MODEL SPECIFICATION AND CLASSES OF STOCHASTIC CONTROL PROBLEMS

For the general adaptive stochastic control problem, the state of the system at time k, $\underline{x}(k)$ is presumed to evolve according to:

$$(1) \quad \underline{x}(k+1) = f\{k, \underline{x}(k), \underline{p}(k), \underline{u}(k), \underline{v}(k)\}, \ k=o,\ldots,N-1$$

where $\underline{p}(k)$ is the vector of parameters, $\underline{u}(k)$ is the vector of controls or decisions applied at time k and $\underline{v}(k)$ is the process noise.

At time k, prior to applying the control, the state of the system $\underline{x}(k)$ may be observed. However, as a result of the measurement noise $\underline{w}(k)$, only partial observation y(k) may be achieved, where[3]

$$(2) \quad \underline{y}(k) = \underline{h}[k, \underline{x}(k), \underline{w}(k)], \ k=1,2\ldots,N.$$

The variables $\underline{x}(k)$, $\underline{p}(k)$, $\underline{u}(k)$, $\underline{v}(k)$, $\underline{y}(k)$, $\underline{w}(k)$ are vectors with the following dimensions: (n x 1), (r x 1), (m x 1), (n x 1), (s x 1), and (s x 1) respectively. The statistics of the random elements are:

$$\underline{x}(0), \ \{\underline{v}(k)\}_{k=0}^{N-1}, \ \{\underline{w}(k)\}_{k=0}^{N-1}$$

and the functional form f(.), h(.) are assumed known.

The objective function is represented by the minimization (maximization) of the expected cost (gain)

$$(3) \quad J = E\,\{C(N, X^N, U^{N-1})\}$$

where the expectation is taken over all the underlying random variables, C is a real-valued function, and U^k and X^k, k=0,...,N are defined as:

$$(4) \quad U^k \equiv \{\underline{u}(i)\}_{i=0}^{k}$$

$$X^k \equiv \{\underline{x}(i)\}_{i=0}^{k}$$

Accordingly, the minimization of the expected cost is performed with respect to the sequence of decisions U^{N-1} applied during the N-stage control process.

The set of observations from time k=1 to time j, after the sequence U^{j-1} has been applied, is denoted by

(5) $Y^j \equiv \{\underline{y}(k)\}_{k=1}^{j}, \; 1 \leq k \leq j \leq N.$

One possible additional specification frequently found in the literature (Rausser and Pekelman [1977], treats the case where the vector of parameters $\underline{p}(k)$ evolves according to

(6) $\underline{p}(k+1) = \underline{g}(k, \underline{p}(k), \underline{\theta}(k))$

where g(.) and the statistics of $\underline{p}(0)$ are known. This formulation admits the special case of constant, but unknown parameters. Clearly, p(k) can be regarded as a state and augmented to $\underline{x}(k)$ so that the new state vector is

(7) $\underline{x}'(k) = [\underline{x}(k), \underline{p}(k)]$

The various informational elements contained in the above system may be captured by a few additional definitions. In particular, the information about the dynamics of the measurement system between time 0 and time j may be denoted as:

(8) $m^j \equiv \{h(k,.,.)\}_{k=0}^{j}$

i.e. m^j represents the structure of the observation program up to time j. Along with this definition, the following joint density distributions prove useful in characterizing alternative adaptive control schemes,

$$(9) \quad L^k \equiv dP[\underline{x}(0), \underline{p}(0), \{v(i)\}_{i=0}^{N-1}, \{\theta(i)\}_{i=0}^{N-1}, \{w(i)\}_{i=0}^{k}], \ k=1,\ldots,N-1.$$

$$L^o \equiv dP[\underline{x}(0), \underline{p}(0), \{v(i)\}_{i=0}^{N-1}, \{\theta(i)\}_{i=0}^{N-1}],$$

Note that this representation implies that L^o does not contain any information on the measurement statistics.

The various classes of stochastic control problems can be defined by the amount of information utilized to determine the control for each period. These classes are not exhaustive and are advanced only to provide some overall structure. It is particularly important to note the difference between the feedback policy and the closed loop policy. For both policies, available information is utilized to determine current decisions, however only in the case of the latter policy does the decision maker anticipate future information and take it into account in his current control calculations.

2.1 The open loop policy

During the horizon [0,...,N-1] no information is available to the controller, i.e. all decisions are made on the basis of the information available at time 0. Therefore the control rule has the form

$$(10) \quad \underline{u}(k) = \underline{u}[k, U^{k-1}, L^o], \ k=0,\ldots,N-1.$$

2.2 The Open-Loop Feedback Policy

At each period k, as the data becomes available, the controller observes this information but in his determination of the control he assumes that no future observations will be available to him. In other words, the controller for each period k is presumed to have no knowledge about future observations. Hence, the control rule is this case has the form:

(11) $\underline{u}(k) = \underline{u}[k, U^{k-1}, Y^k, m^k, L^k]$, $k=0,\ldots,N-1$.

2.3 The M-measurement feedback policy

For this policy, in addition to the current data, the subsequent M-measurements with their statistics are available to the controller, i.e. the control decision is based on the current measurement and M additional anticipated measurements. Here we have:

(12) $\underline{u}(k) = \underline{u}\ [k, U^{k-1}, Y^k, m^{k+M}, L^{k+M}]$, $k=0,\ldots,N-1$.

2.4 The closed loop policy

In this policy the current information as well as all future anticipated information and their corresponding statistics are taken into account, i.e. it is known that the loop will be closed through all periods of the planning horizon. For this policy, the rule will have the form:

(13) $\underline{u}(k) = \underline{u}[k, U^{k-1}, Y^k, M^{N-1}, L^{N-1}]$, $k=0,\ldots,N-1$.

Note that for a deterministic system all four classes are identical.

2.5 Illustration of alternative policies

Further insight into the above policies may be gained by examining a stochastic problem which is not adaptive in its parameters, namely, a production or acreage control scheme where the decision variable is the quota imposed, and observations are made on demand, supply and inventory carryover.realizations. For a N-period problem the open loop policy will consist of N-decisions, one for each of the N periods. These quota decisions will be made at time zero, on the basis of the inventory level at that time and the demand and supply distributions for all future periods. No further observations on future demand and supply realizations will be made. In the feedback policy the quota decision is made in each period k after observation of the realized demand, supply, and inventory of that period. However, it will be assumed that no future observations will be available, i.e. corrective quotas in future periods will not be possible. Note that what constitutes a feedback policy for the horizon [0, N] is actually the set of a first-period solutions of an open loop policy with horizons [0,N], [1,N],...[N,N]. This policy will differ from what is commonly coined the "rolling horizon" problem, i.e. when the length of the horizon remains constant and the set of first-period solutions is generated for the horizons [0,N], [1, N+1],...,[N,2N].

To illustrate the M-measurement policy consider the case where during the N-period problem observations of demand, supply and inventory realizations will be available only for the first M periods. Then the quota decision at time k=0 will be based on the knowledge that corrective quotas can be made during the coming M periods. In period k=1 there will be only M-1 corrective quotas in the future and therefore the policy will be an M-1 measurement policy and in general for period k, $k \leq M$, we have an M-k measurement policy until k=M where the policy will become a feedback or zero-measurement policy. The closed loop policy is obtained when the expected effects of current quotas on future demand, supply, and inventory realizations are taken into account for all subsequent periods. For an N-period planning horizon the closed-loop policy is an N-measurement policy.

In general, the optimal policy for a stochastic control problem will be closed loop. However, due to analytical difficulties in deriving the closed loop control rule the other policy classes are commonly used as an approximation. The order of the analytical difficulty usually decreases as we reduce the amount of information available to the decision maker.

The common formulation of the adaptive control problem has some inherent nonlinearities since the unknown parameters are usually the coefficients of the state and control variables. Therefore, the adaptive control problem typically constitutes a nonlinear stochastic control problem and requires an approximate solution. In the following section we first demonstrate by a particular example the nature of the closed loop solution and its inherent analytical difficulties. We then introduce the notion of actively adaptive and passively adaptive solutions and subsequently relate them to various approximation methods.

3. APPROXIMATE SOLUTIONS

To properly classify various schemes, it is important to first distinguish bewteen approximations performed on the original system and those performed in the process of deriving the optimal rule. Most approximations resulting from the first approach will fall in the category which we define as passively adaptive schemes while the latter type approximations will generally be actively adaptive. In order to motivate the need for approximation we solve a simple two-period problem and show the analytical difficulty in deriving the optimal solution. We utilize this example to illustrate as well the nature of passively adaptive versus actively adaptive solutions.

Consider the two-period problem:

$$(14)\quad J = \min E \sum_{k=1}^{2} (x_k + 1/2 \cdot u_k^2)$$

subject to

$$(15)\quad x_k = \alpha u_k + \varepsilon_k,\ k = 1,\ 2$$

where $\varepsilon_k \sim N(0,q)$ for $k = 1, 2$ cov $(\varepsilon_1,\varepsilon_2) = 0$; $E(\alpha,\varepsilon_k) = 0$; and α is the unknown parameter with the prior distribution at time $k = 0$ given as:

$$(16)\quad \alpha \sim N(\alpha_o,\ \sigma_o).$$

The first decision will be u_1 which when imposed will result in x_1. After the realization of the random variable x_1, the distribution of the parameter α will be reestimated. Using a Bayesian procedure, the a posteriori mean and variance of α, denoted as α_1, and σ_1, are:

$$\alpha_1 = \frac{\alpha_o q + \sigma_o u_1 x_1}{u_1^2 \sigma_o + q} \tag{17}$$

$$\sigma_1 = \frac{\sigma_o q}{\sigma_o u_1^2 + q} \tag{18}$$

Clearly the selection u_1 and the resulting x_1 will effect these estimates. In particular, if $|u_1|$ is large, the updated variance of α will be small. Hence a large control value may assist in learning about α. In order to determine the optimum value of u_1, the value or usefulness of additional information on α must be captured. This value, in the context of a given control problem, allows the trade off between current control and learning for future control to be estimated. If the estimate of α_1 can be improved a better decision u_2 can obviously be made. This argument may be referred to as the active learning part of the decision since deliberate learning in terms of improved estimation is exercised.

The other dominant consideration in determining u_1 is current control, i.e., the impact on the objective function for the current period. Clearly, the first period control which minimizes current expected cost is $u_1 = -\alpha_o$; this setting must be weighted against the desire to set $|u_1|$ at higher levels to achieve additional learning. In this particular example, the only interaction between the two periods is through learning.
The underlying system is inherently static since the state x_{k+1} is independent of x_k. Therefore, we will be able to isolate the static terms corresponding to current control and the dynamic one which corresponds to learning.

Let us turn now to the solution of the problem. Using dynamic programming we have:

(19) $J_3^* \equiv 0$

(20) $J_2^* = \min_{u_2} E\{x_2 + 1/2\ u_2^2 | x_1, u_1) + J_3^*\}$

(21) $J_1^* = \min_{u_1} E\{x_1 + 1/2\ u_1^2 | \alpha_o, \sigma_o) + J_2^*\ (u_1)\}$

Substituting for x_2 and applying the expectation operator we obtain:

(22) $J_2^* = \min_{u_2} \{\alpha_1 u_2 + 1/2\ u_2^2\}$

which yields

(23) $u^*_2 = -\alpha_1$

and

(24) $J^*_2 = -\ 1/2\ \alpha_1^2$

where $\alpha_1 = E\{\alpha | x_1, u_1\}$, i.e., α_1 is the a posteriori mean of the parameter α after the decision u_1 is taken and the realization x_1 is obtained. Substituting (17) into (24) we have:

(25) $J_2^* = -1/2 \left(\dfrac{\alpha_o q + \sigma_o u_1 x_1}{u_1^2 \sigma_o + q}\right)^2$

which is a random variable due to x_1. The cost of learning is negative and is represented by the expected minimum cost

(26) $E(J_2^*) = -\ 1/2\ \left(\alpha_o^2 + \dfrac{\sigma_o^2 u_1^2}{u_1^2\ \sigma_o + q}\right)$

This result indicates that learning is beneficial (negative cost) to future control.

Shifting one stage backward and substituting (25) for $J_2^*(\ )$ into (21) we have

(27) $J_1^* = \min E\ \{(\alpha u_1 + \varepsilon_1 + 1/2\ u_1^2) - 1/2\ \left(\dfrac{\alpha_o q + \sigma_o u_1 x_1}{u_1^2\ \sigma_o + q}\right)^2\ |\alpha_o, \sigma_o\}$

and after taking expectations

$$(28)\quad J_1^* = \min_{u_1} \{(\alpha_o u_1 + 1/2\, u_1^2) - 1/2\, (\alpha_o^2 + \frac{\sigma_o^2 u_1^2}{u_1^2 \sigma_o + q})\}$$

where the first term of (28) refers directly to control while the second term refers directly to learning. The optimum value for the u_1 must achieve some balance between present control (the first term) and learning for future control (the second term).

Unfortunately, the minimization of the two "dual" cost terms in (28) is not a simple analytical task; the optimal decision rule is highly nonlinear in the control variable u_1. As shown in Figure 1, the current control, $C_1(u)$, and the future learning, $E\ J_2^*(u)$, components of the dual cost suggests that (where α_o is assumed to be positive): (i) if q is small, the optimal u_1 is approximately $-\alpha_o$; (ii) if q is moderate, the optimal u_1 lies to the left of $-\alpha_o$; and (iii) if q is very large, the u_1 is approximately $-\alpha_o$. For case (i), the small value of q implies unplanned (or accidental) learning is sufficient to reduce the updated variance (18) while for case (iii), a large value of q suggesting high noise intensity, implies planned learning does not pay off and thus there is no need to look beyond a one-stage optimization problem. Only if q is moderate size does planned learning prove beneficial. Note that similar conclusions can be drawn for different values of σ_o.[4/]

Although the system (15) is static, dynamic interactions are introduced by the reestimation of α after the realization u_1, x_1, i.e., the effect of u_1 and x_1 on α_1 and therefore on the decision in the second stage. The decision u_1 in the first stage and hence the realization x_1 may improve the estimation of α and consequently the second period decision u_2 along with the resulting payoff J_2^*. Hence, when the decision u_1 is made the value J_2^* is taken into account. This is represented by the second term in expression (28).

In this static example, the effects of the control setting do not propagate in time through the system; hence, the effects of learning and control can be nicely distinguished. In a dynamic system, current controls will influence the future state trajectory and thus the dual effects of current control levels would indeed be difficult to separate. However, even for the above static system, to proceed analytically one more period backward would be impossible since a closed form solution for u_1^* cannot be found. Therefore, some approximations are required to obtain a solution.

When the decision is concerned with parameter system learning as well as with the control we will call it active learning. If the active learning of α is ignored, i.e., if we eliminate the second term in the J_1^* expression, the problem will be completely static, and the optimal strategy is simply:

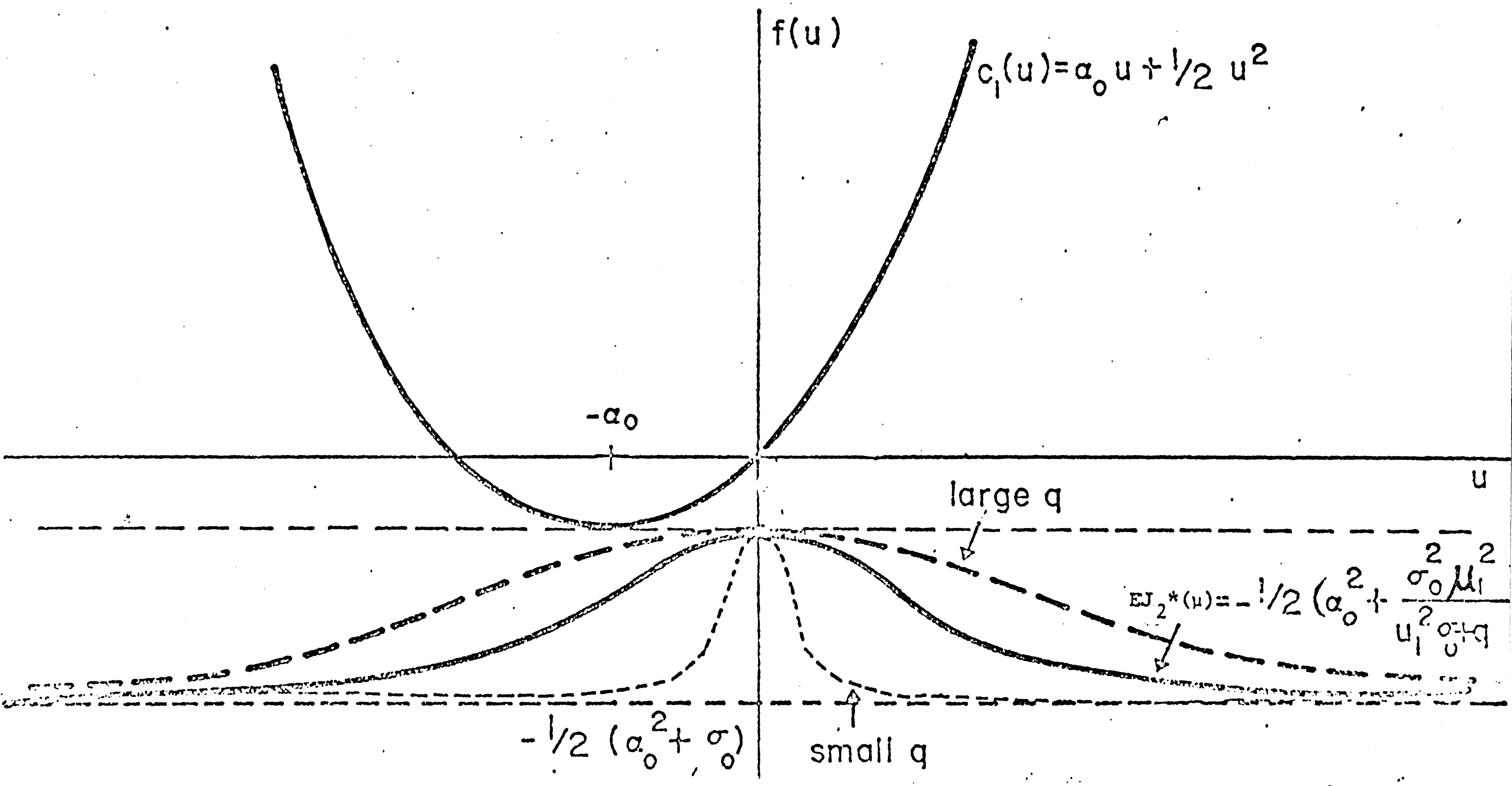
(29) $$u_1^* = -\alpha_o$$

(30) $$u_2^* = -\alpha_1$$

The learning element will still exist, since α will be reestimated, however it will be completely passive, i.e., the amount of learning about α will not be deliberately directed by any previous decision. Various approximation approaches involve the partial or complete neglect of active learning. In this example, an obvious approximation is generated by adding a simplifying assumption to the original system. Specifically, the mean estimate α_1 is treated as though it was the true value of α. This assumption results in the removal of the learning term in (28).

An example of an approximation performed during the derivation of the rule, rather than on the basic system, would be an expansion up to a second order of the active learning term in J_1^* around some nominal values. If we select the nominal value in this case to be $u_1 = -\alpha_o$, we will obtain

Figure 1. Current Control Versus Future Learning



after the expansion to the second order the following:

$$\text{(31)} \qquad J_1^o = \min_{u_1} \{\alpha_o u_1 + 1/2\, u_1^2 + Q_o + \frac{\partial Q_o}{\partial u_1}(u_1+\alpha_o) + \frac{\partial^2 Q_o}{\partial u_1^2}\, 1/2\, (u_1 + \alpha_o)^2\}$$

where Q_o is the active learning term in J_1^* evaluated at the nominal trajectory $u_1^o = -\alpha_o$. Differentiating and substituting we have:

$$\text{(32)} \qquad J_1^o = \min_{u_1} \{\alpha_o u_1 + 1/2\, u_1^2) - 1/2\, (\alpha_o^2 + \frac{\sigma_o^2\, \alpha_o^2}{\sigma_o \alpha_o^2 + q})$$

$$- \frac{\alpha_o\, \sigma_o^2\, q}{(\alpha_o^2 \sigma + q)^2}(u_1 + \alpha_o) + 1/2\, [\frac{3\alpha_o^2\, \sigma_o^3 q - \sigma_o^2 q^2}{(\alpha_o^2 \sigma_o + q)^3}]\, (u_1 + \alpha_o)^2\}$$

and hence a closed form expression for u_1 exists, and multiple-period problems can then be solved recursively. Note that the active learning part is approximated but will still appear in the control rule.

To recapitulate, two approximate forms were introduced, one on the original system by ignoring the uncertainty on the parameter and the other on the optimal rule after the expectations were taken. The first approximation generates a control rule with passive learning while the second form preserves the element of active learning. For many of the common adaptive control approximations this general observation holds. That is, when the approximation is on the original system, the control rule is often passive while if the approximation is performed in the deviation of the rule some of the active learning elements are preserved.

3.1 Passive Adaptive Schemes

The passive schemes which have been advanced to solve the adaptive control problem are simply generalizations of stochastic control methods found in Chow [1975] or Rausser and Hochman [1978]. They include updated certainty equivalents, various versions of open loop feedback, sequential stochastic and linear quadratic Gaussian schemes. All of these schemes involve some modification of original adaptive control system and admit only accidental learning.

3.1.1 Updated certainty equivalence. In this scheme it is assumed that the certainty equivalence property holds i.e., that the mean value of all random variables in the system is their true value. For every period k, the current estimate of the mean value of the random variables is used to replace the random variables and the problem is solved in a deterministic mode. In terms of the information used, the C.E. control rule has the form:

(33) $\underline{u}^{ce} = u^d[k, U^{k-1}, Y^k, m^k]$.

The rule is in feedback form and the parameters and the states are reestimated every period as additional data becomes available. However, no account is taken for future measurements and thus future uncertainty.[5/] Note that the certainty equivalent rule is the same as the deterministic rule $u^d(.)$. However, the time paths of the optimal C.E. controls and associated state variables and criterion function can differ substantially from their deterministic counterparts. This rule and the associated certainly equivalence property does not hold for the general adaptive control problem as a result of the nonlinearity introduced by the product of the parameters and the states.

The C.E. result is closely related to the separation property employed in electrical engineering literature (Witsenhausen [1971]. This property holds if the information about the parameters and the states utilized in the optimal control rule can be captured by their mean value, where the functional form of the feedback equation may differ from the deterministic rule $u^d(.)$. For C.E. formulation, the control rule is equivalent to the deterministic one and the random variables can be replaced by their mean values. The separation property has only the second characteristic and therefore the C.E. property implies the separation property but not vice versa.[6]

3.1.2 Open loop feedback. This approach was originally suggested by Dreyfus [1965] and has been investigated by Aoki [1967], Bar-Shalom and Sivan [1969], Curry [1969], Ku and Athans [1973] and Tse and Athans [1972]. At time k when the control u_k is calculated it is assumed that no measurement will be obtained in the future. The stochastic elements of the parameters as well as those of the states are taken into account. The use of information in this scheme is identical to our previous definition of the feedback class.

In the process of deriving the optimal rule for these schemes, we face the difficult task of determining the expectation of the product of the parameters and the states along with the parameters and the controls. Tse and Athans [1972], for example, avoid the problem by assuming that the parameter multiplying the state is nonrandom, while Ku and Athans [1973] assume that the expected product of the parameter and the state is equal to the product of the expectation.[7] Both of these assumptions obviously represent misspecifications.

3.1.3 Sequential stochastic control. Rausser and Freebairn [1973], Zellner [1971], Chow [1975] and Prescott [1971] have investigated this approach. The derivation of the control rule is based on the assumption that observations will indeed be available in the future. However, they will not be used to adapt the probability distributions on the parameters. Chow [1975] has referred to this approach as control without learning. For the case of a linear system and quadratic objective function the optimal rule here is linear in the mean value and variance of the state. In essence, all of this work treats the unknown parameters as independent, identically distributed random variables for each period of the planning horizon.

3.1.4 Linear quadratic Gaussian approximations. This scheme is thoroughly surveyed in Athans [1971]. If the system is nonlinear and/or the objective function nonquadratic, an expansion around some nominal values may be performed. The nominal values can be generated by imposing on the original problem the certainty equivalence assumption. Given nominal paths for the states, controls, and the parameters, the criterion may be stated as a quadratic function of deviations from these paths. One principal limitation of this approach is that it requires a nominal path for parameter vector and subsequently minimizes deviations from this path. Clearly, the performance of this scheme will deteriorate rapidly if the uncertainty in the parameters or the input noise is "large." Moreover, this scheme is readily operational only for "tracking problems," i.e., where the nominal paths are treated as target trajectories. For meaningful applications, the target trajectories should be inherent in the nature of the decision problem.

3.2 Actively Adaptive Schemes

For situations in which the need for learning is explicitly recognized, knowledge can be accumulated in an active or dual control fashion. A number of schemes have been advanced in the literature for approximating actively adaptive control problems. Most of these schemes provide a consistent approach to the entire planning horizon and thus are approximations to the optimal closed loop solution. When these schemes are applied only to a subset of periods within the planning horizon, they approximate the M-measurement feedback solution.[8/] Each of these schemes views the future by utilizing in various fashions what is currently known about the information to be obtained later.

The key distinguishing feature of the alternative schemes is how they deal with the dependence of future information on present controls. For the general problems, (1) - (4), the information state along with the optimal cost-to-go (value function) associated with each information state must be characterized. The information state, denoted by Ψ_k, is directly influenced by the conditional density, $p(x_k | Y^k, U^{k-1})$. The general expression for how future observations will be made and utilized may be represented by the optimal-cost-to-go, $I^*\{\Psi_{k+1}, k+1\}$. Given these definitions along with a criterion function (4) which is separable across k, a stochastic dynamic programming solution expresses how $I^*\{\Psi_k, k\}$ can be computed (in principle) recursively by

$$(34) \qquad I^*\{\Psi_k, k\} = \min_{u_k} E\{C[x_k, u_k, k] + I^*\{\Psi_{k+1}[\Psi_k, u_k], k+1\} | Y^k, U^{k+1}\}$$

where $\Psi_{k+1}[\Psi_k, u_k]$ monitors the evolution of the information state. Two of the principal difficulties which arise in attempting to deal with (34) are : (i) the information is either infinite dimensional or finite but grows with time; and (ii) the optimal cost-to-go associated with the information state is generally not an explicit function. Approximations are offered by the following schemes to deal with these difficulties and thus all the schemes involve some simplification of the experimental or active learning dimension of dual control.

3.2.1 Tse, Bar-Shalom, and Meier (TBM). The most widely publicized actively adaptive scheme is based on the work of Tse, Bar-Shalom and Meier [1973], Bar-Shalom, Tse and Larson [1974], and Bar-Shalom and Tse [1976]. In essence, the wide-sense dual control procedure advanced by these authors decomposes the complete adaptive control problem into three components (a) current control, (b) future deterministic control, and (c) a future perturbation control. The perturbation or experimental control component is partitioned into a caution and probing term. The caution term reflects the effects of uncertainty at time k and subsequent system noise on the criterion function. The probing term summarizes the effect of uncertainties when subsequent decisions are made.

To handle difficulties (i) and (ii), TBM approximate the information state by maintaining only the first two moments of the state estimate, namely the mean and covariance updated estimates.[9/] These updated estimates can be computed by any one of a number of methods including the extended Kalman filter (Sorenson [1966], Jazwinski [1970]), the adaptive filter (Jazwinski [1970], Mehra [1972]), the second order filters, (Athans et al. [1968]) or the optimal filter (Bucy and Senne [1970], Alspach and Sorenson [1970]. The optimal future benefits or cost-to-go are then associated with this approximate information state at time k+1.

Moreover, the optimal cost-to-go is also approximated. This approximation is motivated by results which are available for the special class of linear-quadratic-Guassian problems. Clearly, if the nonlinear joint control estimation problem can be somehow transferred into a linear quadratic Gaussian problem, an explicit form of the optimal cost-to-go is readily available. These authors make this transference by associating with each control setting u_k a future (fictitious) nominal control sequence. The choice of the nominal is quite flexible and dependent upon the class of problem under consideration. For each nominal control trajectory, there is a corresponding nominal state trajectory and trajectory of variances and covariances.[10/] Perturbation analysis is then carried out around these nominals from which approximate cost-to-go can be obtained that explicitly reflects future learning and control performance.

This approach involves a search over the space u_k which consists of the following basic steps First, select a value of values of u_k, then evaluate explicit functional relationship between u_k and the future covariance matrices, and finally use the resulting functional values to select the next values of u_k to examine. The scheme proceeds until satisfactory covergence is obtained.[11/] Clearly, this active scheme is computationally expensive and for systems with multiple control this expense will often prove to be prohibitive.

3.2.2 Chow [1975]. An actively adaptive scheme has also been developed by Chow [1975] for the case of a quadratic but non-additive criteria function, a linear system (1) and active learning with regard to the unknown parameters. For this specification, a complex nonlinear stochastic control problem is obtained which Chow's method approximates by applying a second order Taylor series expansion in perfectly measured states of the relevant value function. This expansion is taken about some tentative path for the states over the complete planning horizon using numerical differentiation.

Although the Chow second order scheme is a closed-loop approximation, it differs from the TBM approximation in a number of important respects. First, the TBM procedure treats the unknown parameters by augmenting them to the state vector while the Chow approximation does not. Second, the TBM approach makes the computation of the nominal path endogenously, i.e., the nominal trajectory for the entire planning horizon is a function of the current controls. In contrast, the Chow procedure treats the computation of the nominal path exogenously. Third, the Chow approach approximates the entire optimal expected cost-to-go by a second order expansion about a tentative path after taking expectations using numerical derivatives, while the procedure advanced by TBM applies the second order expansion prior to taking the expectations. The latter approach allows the derivation of an explicit representation for the perturbation costs.

A principal limitation of the Chow procedure is the exogenous nature of the nominal path. There is clearly no assurance that the selected control settings will conform to globally optimal controls. However, this scheme can be easily implemented particularly if targets appearing in a quadratic criterion function are prespecified, i.e. the decision problem is in essence a tracking problem.

More recently, Chow [1976] has advanced another approximate solution procedure for the case in which the system representation is nonlinear. That is, the criterion function remains quadratic, the parameters are unknown, but the system representation is nonlinear. In the development of this procedure, Chow [1976] assumes a solution exists and advances a simple first-order Taylor series expansion of the system representation to obtain a linear approximation.

3.2.3 Adaptive Covariance. One actively adaptive scheme which admits a relatively simple computational framework has been developed by MacRae [1975, 1972] who assumes the relevant parameter matrices entering a linear state equation are specified as constant but unknown and the criterion is quadratic. Essentially the adaptive covariance approach alters the original dual control structure of the problem so that equations defining the approximate solution may be derived while maintaining a relationship between current controls and the future information state. Moreover, the equations defining the approximate solution readily admit terms associated with such economic concepts as the price of information and the value of estimation.

Given the above specification, the conditional mean and covariance completely characterize the relevant information state. MacRae introduces a simplification in the updating equations for conditional mean and covariance matrix which eliminates the need for the conditional mean updating relationship. Specifically, by replacing unknown future observations by expected observations in the conditional mean and covariance equations, a temporally invariant result on the conditional mean updates is obtained. Hence, MacRae is left only with a covariance update equation. This covariance update relationship is quadratic in the controls and is introduced into the value function of a deterministic constraint along with an associated matrix of Lagrangian multipliers. Applying conventional dynamic programming procedures the augmented criterion function which remains quadratic in the control leads to analytical results for the control setting which are in linear feedback form. The obvious advantage of this approach is computational simplicity; it is readily tractable for the case in which the criteria function is quadratic in the states.[12/] Of course, to solve for those linear feedback controls the relevant gain matrix must be determined by solving a two-point boundary-value (TPBV) problem.[13/]

3.2.4 Other Adaptive Control Schemes. Prescott [1972] has examined control in the context of a very simple model involving one unknown parameter, a single control and state. His actively adaptive controls were derived numerically with the aid of dynamic programming. In other words, he provides no new method of solution; his results were computed by enumeration of all possibilities. This approach is of course computationally infeasible for the multivariate case.

An approach motivated by the notion of M-measurement feedback control has been advanced by Sarris and Athans [1973] which they characterize as the two-step optimal adaptive control. The optimization for this scheme pertains to only two more periods and imposes the assumption that the unknown parameters are constant and equal to the predicted magnitudes of these parameters. On the basis of this assumption, a control law is derived which is a highly nonlinear function of the state variables. The interesting feature of this scheme is that it takes into account future adaptation of the conditional means but not the variances and covariances associated with estimates of the unknown parameters. This dual control scheme has the advantage of not involving the solution to any iterative system of equations that are encountered with most of the other methods.

Due to the complexity of the above schemes, especially the TBM scheme, some interest has emerged in further simplifications of these schemes for large models. In particular, Norman [1976] has introduced a first-order scheme which involves certain simplifications of the TBM scheme. Like the TBM approach, the first order scheme decomposes the problem into a current control, future deterministic control, and future perturbation control components. However, in contrast to the TBM approach, the first-order scheme neglects all second-order terms involving the interaction between random parameters and states in the derivation of the future perturbation component. In essence, this simply means that the basic difference between the first-order scheme and the TBM procedure is that the former contains no covariances of the states and unknown parameters.

Two versions of the first-order scheme have been investigated by Norman [1976]. The first augments the original state vector with the unknown parameters while the second does not. The second version has some rather obvious advantages, particularly in the computation of the relevant Ricatti matrices which are cubic in the number of states. Hence, the nonaugmented first-order scheme incurs smaller computations which indeed may be significant for larger models.

Another approach which is useful when the underlying probability distributions are non-Gaussian or nonunimodal has been developed by Alspach [1972]. Following some earlier work, he utilized the Gaussian sum approximation to the required a posteriori densities generated by adaptive control approach. Based upon these approximations suboptimal adaptive controls can be computed. Unfortunately, the procedure is not feasible in real time or even off-line if the number of terms in the a posteriori densities are very large or if the number of total stages over which it must be performed are large. The two principal limitations of the approach are the cost of computing the approximate a posteriori densities by the Gaussian sum method and the curse of dimensionality. For these reasons, Alspach suggests employing a version of the M-measurement feedback control approach.

Most of the remaining schemes available in the literature generally apply to specialized structures. For example, Aoki [1967], Saridis and Lobbia [1972], and Wieslander and Wittenmark [1971], have examined problems for which the criterion function is separable across the states and controls. For this specification, a one period step optimization problem is obtained and the availability of future information becomes irrelevant in the computation of u_k. In this sense the method is similar to the open loop feedback method. This approach has been extended by Alster and Belanger [1974] to admit dual control with an ad hoc specification on a guaranteed level of learning. Murphy [1968] and Gorman and Zabordsky [1968] have also investigated a specialized structure of the adaptive control problem, viz. where only the parameters associated with the controls are unknown.

A number of other methods, often appearing under the heading of self-organizing control, are also available. For these methods, the control scheme is treated as a fixed structural form with some adjustable parameters. In general, the parameters are adjusted adaptively by one or two alternative procedures. The first adjustment procedure is based on improved learning of the system parameters and is often referred to as parameter-adaptive schemes (Saridis and Stein [1968], Saridis and Dao [1972], and Tsypkin [1973]). In a similar vein, Deshpande et. al. [1972], for a quadratic criteria function and a linear representation of the system, specify a method of computing the a posteriori probabilities which incorporate some elements of learning. It is assumed that the unknown parameters have discrete probability distribution which is adapted with each new observation. For each particular selection from this discrete probability distribution there corresponds a unique optimal control. These authors suggest weighting the controls by the appropriate probability values and computing a weighted measure of the approximate optimal control. The second procedure adjusts the parameters such that the control settings will improve some "subgoal" generated from the criterion function and has been referred to as performance-adaptive schemes (Saridis and Gilbert [1970], Saridis [1971], and Tsypkin [1973]). For the first procedure an attempt is made to estimate the true parameters of the system while for the second procedure it is the control law's feedback parameters that must be estimated.

Other adaptive schemes utilize second-order perturbation analysis and have been referred to in the literature as trajectory-shaping (Denham [1964], Meier [1965], and Van Der Stolp [1968]). Much of this literature is a simple modification of the linear quadratic Gaussian approach. However, instead of employing the deterministic trajectory, another trajectory which minimizes the cost criterion obtained by utilizing second-order analysis is selected as the reference trajectory.

3.4 Comparative Performance of Passive and Active Control Schemes

A large number of comparisons of some of the alternative suboptimal adaptive schemes are available but almost all of these comparisons are based on rather questionable Monte Carlo simulation designs. This includes the work of Bar-Shalom and Tse [1976], Chow [1975], MacRae [1975], Norman [1976], Prescott [1971], Sarris and Athans [1973], and Zellner [1971]. Generally this work examines only the case of a scalar-state and control variable, a quadratic criterion function, a linear plant representation, and few if any of these studies employ any of the useful results from the statistical literature on sequential experimental designs. Moreover, due to the computational complexity and resulting costs of the various methods, the number of Monte Carlo simulations carried out are generally meager. Furthermore, the data base used for comparisons among various authors differ and thus realistic comparisons of all the alternative suboptimal adaptive schemes are generally nonexistent. Nevertheless, some comparative performance results are available which are worth briefly summarizing here.

Most of the results obtained thus far conform to *a priori* expectations. For example, Prescott [1971] compares three alternative schemes: two passive adaptive schemes and the dual control scheme. The first passive scheme is a certainty equivalence approach while the second is a myopic decision rule which minimizes the criterion function only for the current period and thus completely neglects the future. He finds that the certainty equivalence approach is a reasonable procedure when uncertainty in the unknown parameter is small; specifically when the ratio of the prior's mean to its standard deviation is at least 4 in absolute value. When the ratio of the prior's mean to its standard deviation is smaller than 2, however, experimentation becomes a relevant consideration; i.e., it pays to select a decision larger in absolute value than the one which minimizes current expected loss in order to obtain additional information about the unknown parameter. Another result is that the more periods remaining in the planning horizon, the more important is experimentation.

MacRae [1972, 1975] compares three alternative schemes, viz. certainty equivalence, the sequential stochastic, and the adaptive covariance controls. For a dynamic model (i.e., current states depend upon lag states) in which the parameters of the system representation associated with the lag states are known, it is possible for the adaptive covariance scheme to generate a more conservative current policy than the sequential stochastic rule. This result for the adaptive covariance scheme simply reflects the possibility that the best policy may well be to do a very little at first to avoid the relatively large cost of uncertainty and subsequently compensate later when the effect of the policy is known with more precision. As MacRae [1972, p. 903] points out, "this paradoxical response to the introduction of learning possibilities can occur only in dynamic models and appears more likely with longer planning horizons." This result is modified, however, if the parameter summarizing the relationship between lag and current states in the system representation is unknown. In this event, parameters associated with current controls and with lag stages can be correlated and thus no general implications can be drawn about the relative magnitudes of the policy variables for the three types of schemes examined by MacRae. For this specification, although larger variances imply larger uncertainty involving the model parameters, larger covariances imply more information. In addition, there is no particular reason to presume that either the variances or covariances (both of which appear in the sequential stochastic and adaptive covariance rules) will dominate.

Several Monte Carlo simulations carried out by Tse and Bar-Shalom [1973] and their colleagues at Systems Control have evaluated the comparative performance of the TBM method with the certainty equivalence rule. For most cases examined by these authors, the TBM method gave substantial improvement over the certainty equivalence rule. In contrast, Chow [1975, pp. 267-76] reports a comparison among certainty equivalence, sequential stochastic and the Chow active adaptive scheme for both a single equation and a two-equation system representation. He finds that the various rules do not lead to numerical values of the controls which differ to any substantial degree. An explanation of these apparent conflicting results is that the model representation employed in the Chow comparisons are more precisely estimated than those examined by Tse, Bar-Shalom, and others. In particular, for the former model representations, (i) the standard errors of the estimated coefficients are small relative to the prior estimates of the unknown parameters and (ii) the sample period utilized to obtain prior estimates of the unknown parameters is long compared with the planning horizon.

More recent results involving a comparison of the TBM method, the Chow actively adaptive scheme, the adaptive covariance scheme, and such passively adaptive schemes as open loop feedback and certainty equivalence have been reported by Bar-Shalom and Tse [1976]. For the same dynamic model examined by MacRae [1972] they found that the more sophisticated actively adaptive schemes did not always perform better than the open loop feedback or even the certainty equivalence rule. Similar small differences among the performance of the various algorithms were found by Sarris and Athans [1973]; in particular they did not obtain a distinct ordering among the certainty equivalent and sequential stochastic rules.

Another comparative performance study by Norman [1976] has addressed the trade-off between computational complexity and performance. Utilizing the same model as MacRae [1972] and Bar-Shalom and Tse [1976], he compares augmented and non-augmented variations of the first and second order perturbation control schemes along with heuristic certainty equivalence and open loop feedback. His results clearly indicate that the performance of the alternative schemes is problem-specific and for some cases parameter estimation error can lead to poorer performance for the first-order scheme than the open loop feedback rule. Finally, when computational cost is explicitly considered, Norman finds some specifications for which the open loop feedback as well as the certainty equivalence appear to be the most desirable schemes.

4. EMPIRICAL APPLICATIONS AND CONCLUDING REMARKS

4.1 Empirical Applications

There have been a large number of deterministic as well as stochastic control applications in agricultural as well as general economics. Moreover, there are literally myraids of open-loop feedback applications in automated planning and control systems. Most forecasting-inventory systems utilize an open loop feedback framework to update cost and demand parameters in computing lot size formula and buffer stock levels. These and other deterministic, stochastic and passively adaptive applications are far too numerous to survey here. Although deterministic formulations often provide useful approximations to decision problems faced in agricultural and food systems, it is our view that the very nature of these problems require the explicit treatment of uncertainty. Hence, the advancement of stochastic control methods should receive increased attention from the profession.[14] In the context of agricultural decision making under uncertainty, we have also argued that learning should be given explicit consideration. For these reasons, we shall limit our brief survey to actively adaptive control schemes.[15]

The principal applications of active control schemes include the work of Pekelman and Tse [1976] on advertising, Rausser and Freebairn [1974, 1974a] on agricultural trade policy and a few studies dealing with a deliberate learning of demand functions parameters. Freebairn and Rausser [1974] have examined the inventory level to be carried from one period to another as well as pricing in wheat production and Chong and Cheng [1974] have investigated pricing strategies for the introduction of a new product. Prior to these studies, Little [1966] advanced an adaptive framework with a specialized structure to investigate optimal advertising policies. In the management science literature, this work has been extended in a number of directions including a multivariate representation by Little [1977]. In the context of commodity marketing boards, Rausser and Hochman [1978] have provided a number of applications involving the selection of output prices, quota levels, the distribution of quotas among various producers and in some instances the amount produced and sold. These applications explicitly admit non-linear risk and the possibility of inventory accumulation.

For many of the above applications, certainty equivalence sequential stochastic, adaptive covariance, M-measurement feedback, along with the closed loop approximations advanced by Tse et. al [1973] and Chow [1975] have been employed. In general, these applications illustrate the benefits to be derived from the application of actively adaptive control schemes. The schemes were found to generally lead to more (less) extreme settings of the control level in the first (last) few periods of a given planning horizon than the passive control schemes. Furthermore, they generally resulted in control settings which exceeded in some instances by substantial amounts of solutions obtained for the certainty equivalent approximations. In all of these applications, the expected gains derived from the application of active schemes exceeded all the various versions of the passive learning schemes. These results suggest that decision makers should find it beneficial to incur the cost of learning by substituting knowledge accumulation in current periods for expected gains at some later date.

4.2 Concluding Remarks

The various adaptive control schemes presented in Section 3 have been compared in terms of their utilization of information and the nature of the approximations they impose. As we have seen, existing simulation experiments do not offer any general conclusions about the relative performance of these schemes. Moreover, the recently developed approximation schemes are not based on arguments which imply improved performance, but instead they are motivated by numerical considerations. Unfortunately, no conceptual framework has been advanced on which to base an evaluation and comparison of the various schemes. Currently, the selection of an adaptive control methods for a particular application can only be determined on a trial and error basis.

A conceptual framework which is analytically tractable is required which provides a clear exposition of the interaction between learning and control of a stochastic, dynamic setting. Our principal interest is the development of such a framework should be to obtain (i) measures of the learning capability for various adaptive schemes and their relationship with performance; (ii) relationships of the structural properties of the optimal control scheme with the system structure, observation program, and performance criterion; (iii) a clear indication of how information flows through the system are influenced by the control scheme and how both are related to over-all system performance; (iv) formal measures of the trade-off between computational complexity of the various schemes and their performance; and (v) tight bounds on the performance of alternative schemes. An initial attempt with regard to (i), (ii), and (iii) has been advanced by Tse [1976]. In this work, the learning capabilities of various control laws are formalized in terms of Shannon's information measure. Furthermore, Faden and Rausser [1976] advance a framework which can be employed to formally investigate (iv) and (v).

To be sure, for the number of applications of adaptive control to increase, the development of numerically simpler schemes with high capability of learning is critical. Most of the existing schemes which are operational involve some form of numerical search. The search at times may be lengthly, due mainly to the non-convexity of the value function and the existence of multiple local extrema (Pekelman and Tse [1976]). Hence, the introduction of multiple controls would require for most active schemes a multidimensional search which may be numerically prohibitive.

In addition to the numerical issue, the development of new approaches should also consider aspects common to both private and public economic planning. The first and foremost element missing in the existing approaches is the capability of handling inequality constraints on the controls as well as the states. Although constraints on state variables, in general, can be assured to hold only in probability, their incorporation would enable us to solve the common resource allocation problem in an adaptive form where the coefficients of the return function as well as those specifying the technology could be continuously learned. Another element is concerned with whether or not the system representation is in structural form (current states depending upon other current state variables) or reduced form. All the adaptive control frameworks available in the literature assume the latter form but unfortunately many models in management science and economics are more naturally stated in their structural form. Furthermore, when these structural forms are highly nonlinear is is not possible to derive a corresponding, unique reduced form.

Another issue of importance is model discrimination. A common criticism of adaptive control is that although significant parameter uncertainty may indicate the potential value of an adaptive control scheme, it may also indicate that the model is misspecified. Clearly, identification of the model structure is essential for any quantitative application. However, in the adaptive control framework, the experimentation may be directed toward the discrimination between competing models as well as learning of their parameters. Some work has been accomplished in this area (Saridis and Stein [1968], Saridis and Dao [1972], Athans and Willner [1973], Taylor [1976]) but much remains to be done. In particular, when different model representations appear equally possible the appropriate adaptive control framework should be couched in terms of a simultaneous hypothesis testing and control problem. This is an especially difficult problem when a single model must be selected from a class of non-tested models and where a convex combination of models is not allowed.

A related issue is the impact of model misspecification, e.g., omitted variable, functional form, on the control and prediction performance of the various schemes. In the econometrics literature the reestimation of parameters is usually justified by its ability to capture some of the model misspecification and hence reduce the prediction error. However, it is not clear that control performance is also enhanced by this procedure. Here again, much in the way of additional conceptual work on adaptive control schemes would appear to be necessary.

In the applications sphere, a number of decision problems arising in food and agricultural systems readily lend themselves to the adaptive control framework. These include, as previously noted, new product development; effectiveness of new employee incentive schemes; output-maintenance relations of new machines; the production of new technology; and systems where the parameters tend to vary over time. Other potentially rewarding applications of adaptive control include inter alia (i) the selection of private and public projects where the rate of return in unknown, but the selection process repeats itself; (ii) promotion policy where a periodic decision is made on offering discounts, free samples, and the like; (iii) periodic market sample surveys to determine likely consumer responses to product specifications; (iv) public regulatory management where periodic monitoring decisions must be made to determine compliance and private sector response to regulatory controls; and (v) periodic agricultural as well as natural resource management decisions concerned with the effect of alternative input combinations and land allocations. The adaptive control framework indeed appears promising but much additional evidence should be collected before its true value as a research tool is ascertained; clearly the jury is still out.

FOOTNOTES

1. Formal procedures for utilizing sample editions to update probability distributions specifications include Bayesian methods (Zellner [1971]), least squares revision methods (Albert and Sittler [1965], Kalman type filters (Kalman [1960]), as well as others.

2. Some economists have argued against the use of control theory in economic planning. For example, Kydland and Prescott [1977] conclude that "there is no way control theory can be made applicable to economic planning when expectations are rational." Their argument is based partly on observations made by Lucas [1976]. Lucas notes that since optimal decision rules vary systematically with changes in structure of series relevant to the decision maker, any change in policy will alter the structure of these rules. Thus changes in the policy induce changes in structure which in turn necessitate reestimation in future changes in policy, and so on. This view, however, simply recognizes that the model representation for the private sector is misspecified. If the private sector behavioral rules were properly specified and governmental policy included as exogeneous variables, the later policies would not be reflected in the estimated parameters.

   The additional explanation for the Kydland and Prescott conclusion is their observation that economic planning is not a game against nature but rather a game against rational economic agents. In situations where behaviorial agents in the private sector base their decisions on governmental policy and governments base their decisions on private sector decisions, a non-zero sum game often arises. For such formulations it is well known that no equilibrium exists. While such a formulation may be reasonable in the evaluation of fiscal and monetary policy, it is less relevant in the context of agricultural policy due to the atomistic nature of many of the private sector behaviorial units. Under these circumstances, it is reasonable to assume, in the context of game theory, that reaction functions exist only for the governmental agencies. Farmers, assemblers, processors, input manufacturers and consumers would take the policy rules of government as given, but the government may be presumed to take explicit account of private sector participants decision rules in determining its strategies. In the context of Lucas [1976] and Kydland and Prescott [1977], this formulation admits a hierarchical structure in which the governmental agency is dominant.

   Of course, in an operational context, we are not arguing that an adaptive control formulation be employed to determine policies which should be directly implimented. Instead, we are suggesting such a formulation would indeed be useful to structure the decision making process and perhaps even provide the framework for a model which integrates data collection, estimation, and actual decision making; it should be viewed as a supplementary tool. The resulting decisions that are derived from this formulation are only suggestive of the sorts of policies that might be imposed. Thus, they represent only one input into the usual decision making processes of the public sector. Ideally, they represent the most formal process by which the costs and benefits of alternative decisions are evaluated prior to the selection and implementation of a particular policy.

3. This specification of the measurement equation (2) has been extended by Rausser and Howitt [1975] to include the possibility of measurement controls. That is, an additional variable that can be directly controlled in many decision making processes involves measurement precision. Under certain conditions, these authors demonstrate that the measurement control problem is separable from the behaviorial control problem which involve the optimal setting of (u_k) in equation (1).

4. In fact, the optimum u_1 is determined by the ration q/σ_o and σ_o. Furthermore, the result that both large and small q imply that u_1^* is approximately $-\alpha_o$ is easily seen by examining the optimal decision rule associated with (28). It is:

$$\alpha_o + u + \frac{\sigma_o^3 u^3 - \sigma_o^2 u (\sigma_o u^2 + q)}{(u^2 \sigma_o + q)^2} = 0$$

For both $q \to \infty$ and $q \to o$, the third term approaches zero, and hence $u_1^* = -\alpha_o$.

5. Under certain well-known conditions, stated first by Simon [1956] and Theil [1957], the C.E. solution is optimal. Under these conditions, nothing can be gained by utilizing information on future measurements and uncertainty.

6. For the general conditions under which these properties hold, see Wonham [1968] and Bar-Shalom and Tse [1974]. The word separation is used to indicate the complete independence between the procedures used for estimation of the mean value and those used for the control calculation.

7. For a time-varying parameter problem with quadratic objective function, a linear system and Guassian noise, this approach will yield a linear rule of the following form:

$$\underline{u}_k = - G(k)\underline{x}(k) + \underline{\gamma}(k)$$

where G(k) is the adaptive gain matrix, $\gamma(k)$ is the adaptive correction term and $\underline{x}(k)$ is the estimate of the state mean at time k. The adaptive gain and correction term are both functions of the estimated and predicted mean and covariances of the parameters. It is interesting to note that for a linear-quadratic system where the certainty equivalence assumption holds the optimal rule has the same form, but G(k), $\gamma(k)$ are functions of the mean values of the parameters.

8. For the remaining periods of the planning horizon [N-(k+m)], one of the passively adaptive schemes can be employed to approximate the dual control solution. Nevertheless, as with active schemes that are applied to entire planning horizon, the M-measurement active scheme approximation is quite complicated due to nonlinear dependencies of future variances and covariances on current controls. For illustrations of these nonlinearities and various numerical solution procedures, see Curry [1969], Rausser and Freebairn [1974], Early and Early [1973], and Tse and Bar-Shalom [1973].

9. Note that for this approximation, the dimension of the information state remains constant in time.

10. For a thorough exposition of this approach, see Kendrick and Kang [1976].

11. For the scalar control case, Bar-Shalom and Tse [1976] have used a linear search with quadratic interpolation.

12. It should also be noted that when the criteria function is linear in these states the procedure of MacRae does not admit any active accumulation of information.

13. The TPBV problem has been employed by Rausser and Freebairn [1973] to determine the covariance update matrices and associated Lagrangian multipliers which enter the gain matrix. Their results suggest that for rather large models, this approach is computationally feasible.

14. For interested readers, the journal Management Science will publish a special issue this Fall on deterministic and stochastic control. This issue contains a number of papers which provide a complete survey of deterministic and stochastic control applications to both management science and economics.

15. A number of interesting applications of passively adaptive control methods to individual commodity systems are available in the literature. For example, Kim, et. al. [1975] derived feedback control rules for the stabilization of international cocoa markets. Underwood [1977] as well as Pindyck [1976] have examined the payoff to the formation of cartels by deriving certainty equivalent feedback control solutions for copper and tea. A more recent application of the certainty equivalent method has been provided by Arzac and Wilkinson [1977] who examine the optimal stabilization of the U.S. livestock and feedgrain markets.

REFERENCES

1. ALBERT, A., AND SITTLER, R.W., "A Method of Computing Least Squares that Keep up with Data," SIAM J. Contr., Ser. A, Vol. 3 (1966), pp. 384-417.

2. ALSPACH, D., "Dual Control Based on Approximate A Posteriori Density Functions," IEEE Trans. Aut. Contr., Vo. AC-17 (October 1972) pp. 689-693.

3. ALSPACH, D.L., AND SORENSON, H.W., "Approximation of Density Function by a Sum of Gaussian for Nonlinear Bayesian Estimation," Proc. Symp. Nonlinear Estimation Theory and Its Application, San Diego, Calif., 1970.

4. ALSTER, J., AND BELANGER, P.R., "A Technique for Dual Adaptive Control," Automatica, Vol. 10 (1974), pp. 627-634.

5. AOKI, M., Optimization of Stochastic Systems, Academic Press, 1967.

6. ARZAC, E.R., AND M. WILKINSON, "Optimal Stabilization of a Quarterly Econometric Model of the U.S. Livestock and Feed Grain Markets: Some Preliminary Results," Graduate School of Business Administration, Columbia University, New York, 1977.

7. ATHANS, M., "The Discrete Time Linear-Quadratic - Gaussian Stochastic Control Problem," Annals of Economic and Social Measurement, Vol. 1, No. 4 (1972), pp. 449-491.

8. ATHANS, M., AND WILLNER, D., "A Practical Scheme for Adaptice Aircraft Flight Control Systems," Proc. Symposium on Parameter Estimation Techniques and Applications in Aircraft Flight Testing, NASA Flight Research Center (April 1973).

9. ATHANS, M., WISHNER, R.P., AND BERTOLINI, A., "Suboptimal State Estimation for Continuous-Time Nonlinear Systems from Discrete Noisy Measurements," IEEE Trans. Aut. Cont., Vol. AC-13 (October 1968), pp. 504-514.

10. BAR-SHALOM, Y. AND TSE, E., "Concepts and Methods in Stochastic Control," in C.T. Leondes, (ed.), Control and Dynamic Systems: Advances in Theory and Applications, Academic Press, 1975.

11. BAR-SHALOM, Y., AND TSE, E., "Dual Effect, Certainty Equivalence and Separation in Stochastic Control," IEEE Trans. Automatic Control, Vol. Ac-19 (October 1974), pp. 494-500.

12. BAR-SHALOM, Y., TSE, E., AND LARSON, R.E., "Some Recent Advances in the Development of Closed-Loop Stochastic Control and Resource Allocation Algorithms," Proc. IFAC Symposium on Stochastic Control, Budapest, Hungary, 1974.

13. BELLMAN, R., Adaptive Control Process: A Guided Tour, Princeton, N.J.: Princeton Univ. Press, 1961.

14. BRANDOW, G.E., "Policy for Commercial Agriculture, 1945-71" in Survey of Agricultural Economics Literature, L. Martin, Editor, University of Minnesota Press, 1976.

15. BUCY, R.S., and SENNES, K.D., "Realization of Optimum Discrete-Time Nonlinear Estimators," Proc. Symp. Nonlinear Estimation Theory and Its Application, San Diego, Calif., 1970.

16. CHONG, C.Y., AND CHENG, D.C., "Multistage Pricing Under Uncertain Demand," Annals of Economic and Social Measurement, Vol. 4, No. 2 (Spring 1975), pp. 311-324.

17. CHOW, G.C., Analysis and Control of Dynamic Economic Systems, New York: Wiley, 1975.

18. CHOW, G.C., "An Approach to the Feedback Control of Nonlinear Econometric Systems," Annals of Economic and Social Measurement, Vol. 5, No. 3 (Summer 1976), pp. 297-310.

19. CURRY, R.E., "A New Algorithm for Suboptimal Stochastic Control," IEEE Trans. Auto. Control, Vol AC-14 (October 1969), pp. 533-536.

20. DENHAM, W., "Choosing the Nominal Path for a Dynamic System with Random Forcing Function to Optimize Statistical Performance," Div. Eng. Appl. Phys. Harvard Univ., Cambridge, Mass., Tech. Report 449, 1964.

21. DESHPANDE, J.G., AND UPADHYAY, T.N. AND LAINIOTIS, D.G., "Adaptive Control of Linear Stochastic Systems," Automatica, Vol. 9 (1972), pp. 107-115.

22. DREYFUS, S.E., Dynamic Programming and the Calculus of Variation (Mathematics in Science and Engineering, (Vol. 21), New York: Academic Press 1965.

23. EARLY, R.H., AND EARLY B.N., "On the Relative Performance of the Optimal Control System with M Measurements," IEEE Trans. an Automatic Control, Vol. AC-17, 1973.

24. FELDBAUM, A.A., Optimal Control System, Academic Press, 1965.

25. FITZROY, P.T., "An Adaptive Model for Promotional Decision Making," Special Report, Marketing Science Institute, Philadelphia (April, 1967).

26. FOX, K., SENQUPTA, J.K., AND THORBECK, The Theory of Quantitative Economic Policy, Amsterdam: North-Holland, 1966.

27. GOREUX, L.M., KENDRICK, D.A., AND KIM, H., "Feedback Control Rules for Cocoa Market Stabilization" in W.C. Labys (ed.) Quantitative Models of Commodity Markets, Cambridge; Ballinger Book Publishing Co., 1975.

28. GORMAN, D., AND ZABORDSZKY, J., "Stochastic Optimal Control of Continuous Time Systems with Unknown Gain," IEEE Trans. Automatic Control, Vol. AC-13 (1968), pp. 630-638.

29. JAZWINSKI, A., Stochastic Processes and Filtering Theory, New York: Academic, 1970.

30. JUDGE, G.G., DAY, R., JOHNSON, S.R., RAUSSER, G.C. AND MARTIN, L., Editors, Survey of Agricultural Economics: Quantitative Methods, Volume II, University of Minnesota Press, 1977.

31. KALMAN, R.E., "A New Approach to Linear Filtering and Prediction Problems," ASME J. Basic Eng., Vol. 82 (1960), pp. 35-45.

32. KENDRICK, D., AND KANG, BO HYUN, "An Economists' Guide to Wide Sense Dual Control," Working Paper 75-1, Department of Economics, The University of Texas at Austin, 1976.

33. KU, R., AND ATHANS, M., "On the Adaptive Control of Linear Systems Using the Open Loop Feedback Optimal Approach," IEEE Trans. Automatic Contr., Vol. AC-18 (1973), pp. 489-493.

34. KYDLAND, F.E., AND PRESCOTT, E.C., "Rules Rather than Discretion: The Inconsistency of Optimal Plans," Journal of Political Economy, Vol. 85, May, 1977, pp. 473-491.

35. LITTLE, J.D.C., "A Model of Adaptive Control of Promotional Spending," Operations Research, Vol. 14, No. 6 (1966), p. 1075-1082.

36. LITTLE, J.D.C., "A Multivariate Adaptive Control Model," IEEE Trans. Automatic Control, Forthcoming, 1977.

37. LUCAS, R.E., Jr., "Econometric Policy Evaluation: A Critique," in The Phillips Curve and Labor Markets, edited by K. Brunner and A.H. Meltzer. Amsterdam: North-Holland, 1976.

38. MACRAE, E.C., "An Adaptive Learning Rule for Multiperiod Decision Problems," Econometrica, Vol. 5-6 (1975), pp. 893-906.

39. MACRAE, E.C., "Linear Decision with Experimentation," Annals of Economic and Social Measurement," Vol. 1 (1972), pp. 427-447.

40. MEHRA, R.A., "Approaches to Adaptive Filtering," IEEE Trans. Aut. Contr., Vol. AC-17 (October 1972), pp. 693-698.

41. MEIER, L., "Adaptive Control and the Combined Optimization Problem," Stanford Res. Inst., Menlo Park, Calif., Contract NAS 12-59, Memo 6. SRI Project 5578, February 10, 1966.

42. MEIER, L., "Combined Optimal Control and Estimation," in Proceedings, 3rd Annual Allerton Conference, Circuit and System Theory, 1965.

43. MURPHY, W.J., "Optimal Stochastic Control of Discrete Linear Systems with Unknown Gain," IEEE Trans. Automatic Control, Vol. AC-13 (1968), pp. 338-344.

44. NORMAN, A.L., "First Order Dual Control," Annals of Economic and Social Measurement, Vol. 5, No. 3 (Summer 1976), pp. 311-322.

45. PEKELMAN, D., AND TSE, E., "Experimentation and Control in Advertisement: An Adaptive Control Approach," Working Paper 76-04-01, Decision Science Department, University of Pennsylvania, Philadelphia, Pa.

46. PINDYCK, R.S., "Gains to Producers from the Cartelization of Exhaustive Resources," M.I.T. Sloan School Working Papers, May 1976.

47. PRESCOTT, E.C., "Adaptive Decision Rules for Macroeconomic Planning," Western Economic Journal, Vol. 9 (1971), pp. 369-37.

48. PRESCOTT, E.C., "The Multi-period Control Problem under Uncertainty," Econometrica, Vol. 40 (1972) pp. 1043-1058.

49. RAIFFA H., AND SCHLAIFER R., Applied Statistical Decision Theory, Boston: Harvard Business School Press, 1961.

50. RAUSSER, G.C., "Stabilization of International Commodity Markets: A Stochastic Control Formulation," Management Science, forthcoming 1977.

51. RAUSSER, G.C., AND FREEBAIRN, J.W., "Approximate Adaptive Control Solutions to U.S. Beef Trade Policy," Annals of Economics and Measurement, Vol. 3 (January 1974), pp. 177-203.

52. RAUSSER, G.C., AND FREEBAIRN, J.W., "Estimation of Policy Preference Functions: An Application to U.S. Beef Import Policy," Review of Economics and Statistics, Vol. 56 (November 1974a) pp. 437-449.

53. RAUSSER, G.C. AND FREEBAIRN, J.W., Adaptive Control Methods and U.S. Beef Import Quotas, Center for Mathematical Studies in Business and Economics, University of Chicago, Report No. 7337, (June, 1973).

54. RAUSSER G.C., AND HOCHMAN, E., Dynamic Agricultural Systems: Economic Prediction and Control, Elsevier-North Holland Book Publishing Co. (Dynamic Economic Series), 1978.

55. RAUSSER, G.C., AND HOWITT, R., "Stochastic Control of Environmental Externalities," Annals of Economic and Social Measurement, Vol. 4, No. 2 (Spring 1975), pp. 271-292.

56. RAUSSER, G.C. AND C. RIBOUD, Agricultural Policy: Efficiency Versus Equity, Working Paper Nol 136, Harvard University, 1977.

57. RAUSSER, G.C. AND D. PEKELMAN, "Management Science Implications of Adaptive Control Methods," Cybernetics, forthcoming, 1977.

58. SARIDIS, G.N., "On a Class of Performance Adaptive Self-Organizing Control Systems," in K.S. Fu, (ed.), Pattern Recognition and Machine Learning, Plenum Press, 1971.

59. SARIDIS, G.N., AND DAO, T.K., "A Learning Approach to the Parameter-Adaptive Self-Organizing Control Problem," Automatica, Vol. 9 (September 1972), pp. 589-598.

60. SARIDIS, G.N., AND GILBERT, H.D., "Self-Organizing Approach to the Stochastic Fuel Regulator Problem," IEEE Trans. On Systems Sciences and Cybernetics, (July 1970), pp. 186-191.

61. SARIDIS, G.N., AND LOBBIA, R.N., "Parameter Identification and Control of Linear Discrete-Time Systems," IEEE Trans. Automatic Control, Vol. AC-17 (February 1972), pp. 52-60.

62. SARIDIS, G.N., AND STEIN, G., "An Approach to the Parameter Adaptive Control Problem, "TR-EE68-35, Purdue University, School of Electrical Engineering, December, 1968.

63 SARRIS, A.H., AND ATHANS, M., "Optimal Adaptive Control Methods for Structurally Varying Systems," Working Paper #24, National Bureau of Economic Research, Cambridge, Mass., December 1973.

64. SIMON, H.H., "Dynamic Programming under Uncertainty with a Quadratic Function," Econometrica, Vol. 24 (1956), pp. 74-81.

65. SORENSON, H.W., "Kalman Filtering Techniques," in C.T. Leondes, (ed.), Advances in Control Systems, Vol. 3, Academic Press, 1966.

66. SWORDER, D., Optimal Adaptive Control Systems, New York: Academic Press, 1966.

67. TAYLOR, J.B., "Methods of Efficient Parameter Estimation in Control Problems," Annals of Economic and Social Measurement, Vol. 5, No. 3 (1976), pp. 339-348.

68. THEIL, H., "A Note on Certainty Equivalence in Dynamic Planning," Econometrica, (April 1957), pp. 346-349.

69. TSE, E., "Sequential Decision and Stochastic Control, " Unpublished Manuscript, Department of Engineering-Economic Systems, Stanford University, 1976.

70. TSE, E., AND ATHANS, M., "Adaptive Stochastic Control for a Class of Linear Systems," IEEE Trans. Automatic Contr., Vol. AC-17, (February 1972), pp. 38-52.

71. TSE, E., AND BAR-SHALOM, Y., "An Actively Adaptive Control for Discrete-Time Systems with Random Parameters," IEEE Trans. Automatic Control, Vol. AC-18 (April 1973), pp. 109-117.

72. TSE, E, BAR-SHALOM, AND MEIER, L., "Wide-Sense Adaptive Dual Control of Stochastic Nonlinear System," IEEE Trans. Automatic Control, Vol. AC-18 (April 1973), pp. 98-108.

73. TSYPKIN, Y.Z., Foundations of the Theory of Learning Systems, Academic Press, 1973.

74. UNDERWOOD, J., "Optimal Rules for Cartel Managers with Empirical Applications for the Copper and Tea Markets," Annals of Economic and Social Measurement, Vol. 6, (Spring 1977).

75. VAN DER STOEP, D.R., "Trajectory Shaping for the Minimization of State Variable Estimation Errors," IEEE Trans. Automatic Contr. (June 1968), pp. 284-286.

76. WIESLANDER, J., AND WITTENMARK, "An Approach to Adaptive Control Using Real Time Identification," Automatica, Vol. 7 (May 1971).

77. WITSENHAUSEN, H.S., "Separation of Estimation and Control for Discrete-Time Systems," Proceedings of the IEEE, Vol. 59, No. 11, Nov. 1971.

78. WONHAM, W.M., "On the Separation Theorem of Stochastic Control," SIAM J. Control, Vol. 6, No. 2 (1968), pp. 312-326.

79. ZELLNER, A., An Introduction to Bayesian Inference in Econometrics, New York: John Wiley and Sons, Inc., 1971.

49 p.