MULTIPERIOD OPTIMIZATION IN ECONOMIC SYSTEMS WITH UNKNOWN PARAMETERS

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF ECONOMICS

AND THE COMMITTEE ON THE GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

by

John Brian Taylor

March 1973

ACKNOWLEDGMENTS

I wish to express my thanks to Professor Theodore W. Anderson, my principal dissertation advisor, for his numerous helpful suggestions and criticisms. This study owes much to his advice ranging from expert guidance around mathematical pitfalls in the early stages to detailed editorial suggestions on mathematical style in later stages. In addition I wish to thank Professor Takeshi Amemiya for helpful comments on earlier drafts and for general encouragement in my econometric studies.

I also wish to acknowledge the help of Professor Hayne Leland who encouraged me to relate my results to other methods of economic optimization and suggested ways to do so. Professor Louis Gordon of the Statistics Department was also of assistance in pointing out some of the mathematical subtleties of probability theory and in doing so prevented some mistakes on my part.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Investigations of optimization in economic systems usually assume that economic agents have complete information about the parameters of the model used to describe the system. However, any economic model contains parameters about which there is only incomplete information. Macroeconomic models used for stabilization or planning purposes contain parameters whose true values are unknown to the planner. Microeconomic models of the firm or consumer involve parameters whose true values are unknown to the entrepreneur or consumer whose behavior they describe. If these models are to be utilized, it is necessary to have a theory which explains how this incomplete information affects optimization strategies.

In situations where the model is used in only a single period there has been much research concerning various statistical and economic policy issues both at the macro level and at the micro level. The work of Fisher (1962), Brainard (1967), Basu (1972), and Leland (1972), for example, deals with some of these issues in the single period problem when there is uncertainty in the parameters of the model.

However, in situations where the model is to be used for more than one period, there has been relatively little research. In multi-period problems, present period decisions affect not only present period performance, but also the amount of information one can learn about the unknown parameters. For example, a firm facing an unknown demand curve may experiment with different prices in order to learn more about demand. But such experimentation is costly if it requires deviating from the profit maximizing price. Similar possibilities for experimentation might be available to a macroeconomic planner to obtain information for more accurate planning in the future.

The purpose of this study is to investigate methods of optimization in such multiperiod control problems with unknown parameters. Particular emphasis is placed on the statistical estimation aspects, which determine how one learns about the unknown parameters through experience as the process evolves over time. Since the main difference between the single period problem and the multiperiod problem is this estimation aspect, it is of interest to investigate how it affects control behavior.

1.1 A Review of the Literature

As will be discussed in Chapter II multiperiod control is a problem in sequential statistical decision making about which there is a great deal of literature starting with the important works of Wald (1947) and Arrow, Blackwell, and Girshick (1949). The latter paper was the first to use the method of backward induction, a method which has become fundamental to dynamic programming in general. There are many problems of this type which are scattered through the statistics, engineering, and operations research literature.

The first attempts to apply these methods to the multiperiod control problem seem to have been in the engineering literature. In a series of four papers, Fel'dbaum (1960) applied the methods of Wald to general systems of equations using Bayesian methods, calling the approach dual control. This approach was also taken by Aoki (1967) in a book which summarizes much of the engineering literature on the subject until that time.

Economists have been concerned with multiperiod control under uncertainty for quite some time, but the additional elements of learning and estimation have been considered only recently. Simon (1956) and Theil (1957) showed how in a very special case one could use the principle of certainty equivalence to solve such problems, but neither author considered estimation aspects. Prescott (1967) discussed the problem from a Bayesian viewpoint, proposed a heuristic procedure for solving it, and applied the solution to a small macroeconomic model. Zellner (1971) considered the two period problem and some possible approximations.

In a recent paper Prescott (1972) examined a linear model with one unknown parameter and characterized how experimentation is reflected in the control decision. He also performed numerical integration to calculate the Bayes rule for this special case. We discuss his results in more detail in Chapter II.

All these studies have approached the multiperiod control problem from a Bayesian point of view. This approach is rather straightforward to set up and solutions can be written down, in principle.

However, the calculation of specific control rules or even the characterization of their properties has proved quite difficult. In addition, the results are subject to the prior distribution assumptions, which are made partly on the basis of convenience. The approach of this study is not restricted to the Bayesian viewpoint.

1.2 Plan of Study

Before proceeding it will be useful to briefly review the plan of study. In Chapter II we describe the sequential statistical decision model which will serve as the framework of analysis for this study. The Bayes method is described in detail here and is used to demonstrate the trade-off between estimation and single period performance. (Those readers who are familiar with the Bayes approach to this problem may want to skim these sections.) Several control rules other than the Bayes optimal control rule are defined at the end of the chapter.

An investigation of the properties of these control rules is contained in Chapter III. We show that, under certain assumptions, various control rules converge with probability 1 to the value which would be used if the parameters were known with certainty. Since the data which is used for estimation purposes is controlled, it must be considered random, a complication which requires more general methods than are usually necessary to prove consistency in econometric investigations. We also examine the asymptotic distribution of the control rules. These distributions can be used for making confidence statements about the parameter estimates as well as the control itself. In addition, we use these results as measures of control performance in Chapter IV. The results of this chapter show that even some of the simple control rules have certain desirable properties and converge to their true value rather rapidly.

In Chapter IV we consider the possibility of using criteria from the theory of estimation as criteria for the control problem. Such criteria are free from prior distributional assumptions on the parameters and do not require specific distributions for the random disturbance terms. They therefore are attractive alternatives to the Bayes approach to the control problem. The results of Chapter III are then used to show that some simple control rules behave quite nicely and that for long horizon problems there is little to gain from experimentation. Since the results hold only asymptotically we do not know how accurate they are in problems of short time horizon. It is hoped that this method of analysis might prove useful in other control problems and serve as a complement to the Bayesian viewpoint which has dominated analyses of these problems.

In Chapter V we show that, in addition to normative implications, these results have some important implications for the theory of economic behavior. We consider a model of a firm which faces an unknown market demand curve, define equilibrium for such a firm, and show that under certain price adjustment assumptions the equilibrium is stable with probability one. We also consider an alternative model which leads to different conclusions and compare the assumptions of the two approaches.

CHAPTER II

THE MODEL DESCRIPTION AND CONTROL RULES

2.1 Introduction

This chapter describes in detail the framework which we use in analyzing multiperiod control problems in models with unknown parameters. The framework is restricted to the linear models which we consider in this study, but can be generalized to include other problems. We begin by describing the sequential statistical decision model and then describe various approaches to finding control rules. The Bayesian approach is described in detail here and is used to motivate the intuitive considerations which lie behind the problem. Finally, on the basis of the development of the chapter, we define several possible control rules. We investigate a number of these rules in Chapter III with an aim at developing some new criteria for studying control.

2.2 The Sequential Statistical Decision Model

The previous analyses of the linear control problem mentioned in the introduction have been from a strictly Bayesian viewpoint, where expected loss with respect to some prior distribution is minimized. As a consequence, prior distributions on the parameters and the updating mechanisms have been included as part of the structure of the decision model. For the purpose of calculating Bayes control rules, this is a satisfactory procedure, but an unfortunate by-product is confusion between the statement of the problem and the method of solution. Since we do not restrict ourselves to Bayesian methods in this study, we first set up the complete model in sequential statistical decision theoretic terms. Other methods for selecting control rules can then be clearly seen.

The first element in the decision model is a description of the possible states of the world. In the case of the control problem we assume that there is a linear relation between an endogenous variable x_t and a control variable u_t which is subject to an unobservable random disturbance term ϵ_t with zero mean and finite variance σ^2. That is,

$$(2.1) \qquad x_t = \beta_1 + \beta_2 u_t + \epsilon_t ,$$

with ϵ_t independently and identically distributed and having distribution $F(\cdot)$ for all t. In the most general case we consider, β_1, β_2 and the distribution of ϵ_t are unknown, so that the possible states of the world are described by (2.1) where $(\beta_1, \beta_2) \in R^2$ and $F \in \{$all distribution functions with zero mean and finite variance$\}$.

The second element of the decision model is a description of the available control rules which can be used to control the endogenous variable in equation (2.1). We would like a control rule to utilize the information which accumulates over time in the form of observations on x_t and u_t. Therefore, we require that the set of available controls consist of all sequences $\{u_t\}$, the elements of which are chosen sequentially on the basis of past observations. More specifically each element of $\{u_t\}$ is a function of all random variables observed prior to time t ; that is,

$$u_t = u_t(x_1, \ldots, x_{t-1}; u_1, \ldots, u_{t-1}) \quad . \tag{2.2}$$

Thus, a particular control rule can be thought of as a set of instructions which specifies the control action to take at each point in time for all possible developments of the process until that point in time.

The third element of the decision model is a loss function which describes the loss incurred when a particular control rule is used in a given state of the world. In general, loss is defined as a function of the sequences $\{x_t\}$ and $\{u_t\}$; however, for most of this discussion we will deal with a more specific type of loss function which we introduce immediately. We assume that there is some desired level a for the endogenous variable x_t and that this desired level remains fixed for all t . The penalty associated with deviations of x_t from a is assumed to be the quadratic $(x_t - a)^2$. In a finite horizon model, with time horizon T , the loss function is assumed to be the sum of these squared deviations:

$$L = \sum_{t=1}^{T} (x_t - a)^2 \quad . \tag{2.3a}$$

In an infinite horizon model the sum of these losses might not converge, so we could introduce a discount factor, $0 < \rho < 1$ and set

$$L = \lim_{T \to \infty} \sum_{t=1}^{T} \rho^t (x_t - a)^2 \quad . \tag{2.3b}$$

In some infinite horizon problems it may not be appropriate to discount the future (for example in stabilizing the rate of inflation in an economy). In such cases a more appropriate loss function might be an average loss over all time periods. That is,

$$L = \lim_{T \to \infty} \frac{\sum_{t=1}^{T} (x_t - a)^2}{T} \quad . \tag{2.3c}$$

However, such an average loss function would not be appropriate if the penalty associated with deviations of x_t from a converges to zero, for then the above loss function would be identically zero. In such situations one could divide the sum of the losses by a function which diverges more slowly than T. In Section 4.4 of this study the penalty for deviations of x_t from a is assumed to be the regret obtained by subtracting the expected loss if β were known from actual loss; this regret is $[(x_t - a)^2 - \sigma_t^2]$. As will be shown this function converges to zero with probability one (as will the time average), so that a loss function such as (2.3c) is not appropriate. We therefore consider the loss function

$$L = \lim_{T \to \infty} \frac{\sum_{t=1}^{T} [(x_t - a)^2 - \sigma_t^2]}{\log T} \tag{2.3d}$$

which we will later show does not converge to zero. There has been relatively little research done with loss functions such as (2.3c) or (2.3d), although they seem especially useful in problems where discounting seems inappropriate.

Where one is concerned with bringing a variable like the rate of inflation to some desirable level, quadratic loss functions are common. If the model represented the demand curve of a firm with price u_t and quantity sold x_t then, under fixed production costs, the loss in each period would be $-x_t u_t$. We will refer to this alternative type of loss function at various points in the analysis.

Given these three elements of the decision model for the control problem (the set of possible states of the world, the set of available control rules, and the loss function), the objective is to choose a control rule to minimize expected loss, where the expectation is taken with respect to the distribution of the random disturbances $\{\varepsilon_t\}$. In general, this expected loss, or risk R , will depend on the unknown parameters. With the model (2.1) and the above loss function we have that $R = R(\beta_1, \beta_2, \sigma^2)$.

However, the fact that the risk depends on the unknown parameters means that we cannot expect a control rule which minimizes risk for all values of $(\beta_1, \beta_2, \sigma^2)$ to exist except in special cases. Therefore, the decision problem as stated so far is not complete. This is, of course, the usual situation in statistical decision problems and there are several possible methods that can be used to arrive at a reasonable control rule.

One general method is to restrict the class of available control rules to those satisfying a certain desirable property. Out of this restricted class it then might be possible to find a control rule which minimizes $R(\beta_1, \beta_2, \sigma^2)$ for all values. Two methods of restriction which are frequently employed are <u>unbiasedness</u> and <u>invariance</u>. Even if these were desirable in the control problem, they would be difficult to apply because of the complicated structure of $R(\beta_1, \beta_2, \sigma^2)$.

Another method of restriction, which is used in the theory of estimation, is to restrict the class of control rules to those which are consistent and asymptotically normal, and then find the control rules out of this class which have the smallest asymptotic variance. This latter method has special appeal in the control problem where one is concerned with the behavior of an entire sequence and where the risk is difficult to evaluate. Further, in situations where the risk may not be finite for any control rule, the variance of the asymptotic distribution is a viable alternative. In problems with a long time horizon this method would be especially useful and we discuss it in detail in Chapter IV after deriving the necessary results in Chapter III.

A second general method is to decide on a principle which leads to an ordering of the available decision rules. Two such principles are minimax and Bayes. Since the minimax principle usually requires calculation via the Bayes rule we will only consider the latter here.

The Bayes method assumes that prior knowledge of the unknown parameters can be characterized by a prior distribution function. One can then compute the expectation of the risk $R(\beta_1, \beta_2, \sigma^2)$ with respect to this prior distribution and arrive at the Bayes risk r. The Bayes rule is then to choose that control rule which minimizes the Bayes risk. Such a control rule will be called the Bayes optimal control rule. As mentioned above the Bayes method has been the favorite method of previous investigations, and has suggested several control rules which are either numerical or analytic approximations to the Bayes optimal control rule. We discuss this approach in some special cases under some particular prior distribution assumptions below.

2.3 The Bayes Optimal Control Rule

The advantage of the Bayesian approach to the multiperiod control problem is that it provides a convenient way to describe how information accumulates from one time period to the next and, therefore, permits the use of the backward induction method of dynamic programming. We discuss this procedure first in a model with known intercept (β_1 known) and, secondly, in a model with both unknown intercept and unknown slope.

2.3.1 The Case of One Unknown Parameter

Without loss of generality[1] we assume that $\beta_1 = 0$ and, therefore, drop the subscript on $\beta_2 = \beta$. Thus, the model (2.1) becomes

---

[1] If the known intercept were $\alpha \neq 0$ then, by redefining the endogenous variable $x_t^* = x_t - \alpha$ and the target $a^* = a - \alpha$, the model could be reduced to the zero intercept case of equation (2.4).

(2.4) $$x_t = \beta u_t + \epsilon_t \quad .$$

We assume that the prior distribution function of β is normal $N(b_0, \sigma_0^2)$ and that the distribution of ϵ_t is also normal $N(0, \sigma^2)$. Thus we assume that the variance of the distribution of the disturbance term is known and equal to 1.

Under such distribution assumptions, it can be shown[2] that the posterior distribution of β at any time t is also normal $N(b_t, \sigma_t^2)$ where

(2.5) $$b_t = \frac{\dfrac{b_{t-1}}{\sigma_{t-1}^2} + \dfrac{x_t u_t}{\sigma^2}}{\dfrac{1}{\sigma_{t-1}^2} + \dfrac{u_t^2}{\sigma^2}} ,$$

and

(2.6) $$\frac{1}{\sigma_t^2} = \frac{1}{\sigma_{t-1}^2} + \frac{u_t^2}{\sigma^2} \quad .$$

These recursive relations show how information in the form of observations updates the prior distributions into posterior distributions.

---

[2]See Raiffa and Schlaifer (1961), p. 337 for this calculation.

We will assume that the problem has a finite time horizon and that the loss function can be written

$$(2.7) \qquad L = \sum_{t=1}^{T} (x_t - a)^2 .$$

The Bayes optimal control rule for this problem is given by that control rule which minimizes the Bayes risk, where the Bayes risk is found by integrating (2.7) with respect to the distribution of the error terms and the prior distributions of the unknown parameter β . The backward induction method of dynamic programming can be used.

In the last period we minimize the Bayes risk given all observations until that period. Letting $Z_t = (u_1, \ldots, u_t; x_1, \ldots, x_t)$ and substituting from (2.4) into the last term of (2.7) we compute

$$(2.8) \qquad E[(x_t - a)^2 | Z_t] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (\beta u_T + \epsilon_T - a)^2 dP(\epsilon_T) dP(\beta | Z_{T-1})$$

$$= \int_{-\infty}^{+\infty} [(\beta u_T - a)^2 + \sigma^2] dP(\beta | Z_{T-1})$$

$$= (b_{T-1}^2 + \sigma_{T-1}^2) u_T^2 - 2ab_{T-1} u_T + a^2 + \sigma^2 ,$$

since the conditional distribution $P(\beta | Z_{T-1})$ is $N(b_{T-1}, \sigma_{T-1}^2)$. Minimizing (2.8) with respect to u_T results in

$$(2.9) \qquad u_T^* = \frac{ab_{T-1}}{b_{T-1}^2 + \sigma_{T-1}^2} ,$$

which would be the Bayes optimal control rule in a one period problem. But for the multiperiod problem we must compute the rest of the sequence $\{u_t\}$.

Substituting (2.9) into (2.8) we get the minimizing value of the Bayes risk in the last period:

$$(2.10) \qquad r_T^* = \frac{\sigma_{T-1}^2 a^2}{b_{T-1}^2 + \sigma_{T-1}^2} + \sigma^2 \quad .$$

In the next to last period we minimize the Bayes risk given all observations until that period. That is, we compute

$$(2.11) \qquad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [(\beta u_{T-1} + \epsilon_{T-1} - a)^2 + r_T^*]^2 dP(\epsilon_{T-1}) dP(\beta | Z_{T-2}) \; .$$

The integration of the squared error term involves no difficulties as it is exactly analogous to (2.8). But the integration of r_T^* is not so easy, as can be seen by substituting from (2.5) into (2.10) to get

$$(2.12) \qquad r_T^* = \frac{a^2}{\left(\frac{b_{T-2}^2}{\sigma_{T-2}^2} + \beta u_{T-1}^2 + u_{T-1}\epsilon_{T-1}\right)^2 \Big/ \left(\frac{1}{\sigma_{T-2}^2} + u_{T-1}^2\right) + \sigma^2} \quad .$$

We know of no way to integrate analytically this expression multiplied by the normal density. Thus even in this extremely simple problem it seems that analytic expressions for Bayes optimal control rules are not feasible. Further, if there are more than two time periods in the problem we must continue this backward induction procedure with increasingly complicated integrations. Such computation problems are not uncommon to dynamic programming formulations and the usual recourse is to numerical methods or other approximations. This is one reason why we consider alternative approaches in Chapters III and IV.

Before proceeding in the analysis several points should be made. The expectation of (2.12) will be a function of u_{T-1}. If we could ignore r_T^* in (2.4), then the minimizing value of u_{T-1} would look exactly like u_T; and in general, if we could ignore future risk, the control rule would be

$$(2.13) \qquad u_{t+1} = \frac{ab_t}{b_t^2 + \sigma_t^2} ,$$

for all t. But, since the expectation of r_T^* will be a function of u_{T-1}, as can be seen from (2.12), we cannot ignore r_T^* when minimizing and in general we cannot ignore the future.

The reason that future risk depends on the present control value in this problem even though there is no formal mechanism relating the two in the model, is that present action influences the amount of future information about the unknown parameters. For example (2.6) shows that the variance σ_{T-1}^2 of the posterior distribution will be smaller the larger is u_{T-1}^2. The expectation of r_T^* represents the expected value of this information in the next period, and is therefore also a function of u_{T-1}.

Intuitively this result makes sense. In a model with known intercept, we get more information by taking observations as far away as possible from the known intercept. However, there is a limit to such experimentation with u_{T-1} because we are also concerned with present period performance. That is, an excessively large value for u_{T-1}, for experimentation, would result in an excessive mean square error loss in period T-1 . Thus there is a trade-off between learning for the future and present period performance. Prescott (1972) proved that the Bayes optimal rule will always be larger in absolute value than the single period rule under the distributional assumptions of this section. This result confirms these intuitive conjectures.

The important implication of this discussion is that the multi-period control problem is quite different than the one period control problem. The choice of control influences the amount of information which can be obtained about the unknown parameters. This feature occurs even in models such as (2.4) which do not contain lagged dependent variables and are thus essentially static. The ability to use the control for learning purposes has made a dynamic model out of what at first appears to be a static model.

2.3.2 The Case of Two Unknown Parameters

When the intercept in the linear model is also unknown the complexity of the Bayes procedure greatly increases. In addition, our intuition as to how the optimal control rule might reflect experimentation also breaks down.

In model (2.1) we now assume that the prior distribution function of the unknown parameters β_1 and β_2 is bivariate normal; that is

$$(2.14)\qquad \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \sim N\left[\begin{pmatrix} b_{10} \\ b_{20} \end{pmatrix}, \begin{pmatrix} \sigma_{110} & \sigma_{120} \\ \sigma_{210} & \sigma_{220} \end{pmatrix}\right],$$

and that the distributions of ε_t is also normal $N(0, \sigma^2)$. Under these distribution assumptions it can be shown that the posterior distribution will also be normal with mean

$$(2.15)\qquad \begin{pmatrix} b_{1t+1} \\ b_{2t+1} \end{pmatrix} = \left[\begin{pmatrix} \sigma_{11t} & \sigma_{12t} \\ \sigma_{21t} & \sigma_{22t} \end{pmatrix}^{-1} + \frac{1}{\sigma^2}\begin{pmatrix} 1 & u_{t+1} \\ u_{t+1} & u_{t+1}^2 \end{pmatrix}\right]^{-1}$$

$$\times \left[\begin{pmatrix} \sigma_{11t} & \sigma_{12t} \\ \sigma_{21t} & \sigma_{22t} \end{pmatrix}^{-1}\begin{pmatrix} b_{1t} \\ b_{2t} \end{pmatrix} + \frac{1}{\sigma^2}\begin{pmatrix} x_{t+1} \\ u_{t+1}x_{t+1} \end{pmatrix}\right]$$

and covariance

$$(2.16)\qquad \begin{pmatrix} \sigma_{11t+1} & \sigma_{12t+1} \\ \sigma_{21t+1} & \sigma_{22t+1} \end{pmatrix}^{-1} = \left[\begin{pmatrix} \sigma_{11t} & \sigma_{21t} \\ \sigma_{12t} & \sigma_{22t} \end{pmatrix}^{-1} + \frac{1}{\sigma^2}\begin{pmatrix} 1 & u_{t+1} \\ u_{t+1} & u_{t+1}^2 \end{pmatrix}\right]^{-1}.$$

We assume a finite horizon model with the loss function defined in Equation (2.7).

The Bayes optimal control rule for this problem can be computed in principle by the same method of backward induction described above. In the last period we minimize the Bayes risk given all observations until that period. This Bayes risk is computed as

$$(2.17) \qquad \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (\beta_1 + \beta_2 u_T + \varepsilon_T - a)^2 \, dP(\varepsilon_T) \, dP(\beta_1, \beta_2 | Z_{T-1})$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [(\beta_1 + \beta_2 u_T - a)^2 + \sigma^2] \, dP(\beta_1, \beta_2 | Z_{T-1})$$

$$= (b_{1T-1} + b_{2T-1} u_T - a)^2 + \sigma_{22T-1}(u_T + \frac{\sigma_{12T-1}}{\sigma_{22T-1}})^2$$

$$+ \frac{\sigma_{11T-1}\,\sigma_{22T-1} - (\sigma_{12T-1})^2}{\sigma_{22T-1}} + \sigma^2$$

which, when minimized with respect to u_T, gives

$$(2.18) \qquad u_T^* = \frac{(a - b_{1T-1})b_{2T-1} - \sigma_{12T-1}}{\sigma_{22T-1} + b_{2T-1}^2} .$$

This value can then be substituted into (2.17) to obtain the minimized value of the risk in the last period. We can then proceed by backwards induction to the next to last period, but as in the last section, this results in an integration which seems to have no closed form. As yet there has been no progress in either numerical or analytic approximations to the problem, and since our main interest is not in deriving the Bayes control rule we will not continue the analysis here.

In the problem with known intercept, we found that experimentation could be carried out by using larger control values, in absolute value, than one otherwise would. In this case experimentation is not so obvious. Alternating between large and small values of the control would seem to be a good strategy for learning about the unknown parameters β_1 and β_2 . On the other hand, we are not basically interested in these two parameters, but only in the point where the linear relation $x_t = \beta_1 + \beta_2 u_t$ intersects the line $x_t = a$. It is not obvious that the best way to obtain information about this point is to obtain the most information about the individual parameters. Perhaps values close to the point would give more information. This would suggest a much different control rule.

2.4 Other Control Rules

The above section has defined the Bayes optimal control rule for two different control problems. Other control rules which may be chosen for various reasons are discussed in this section. In the model with known intercept a particularly easy to calculate control rule is the sequence defined by $u_1 \neq 0$ but otherwise arbitrary, and

$$u_{t+1} = \frac{a}{\hat{\beta}_t}, \quad t = 1, 2, \ldots, \tag{2.19}$$

where $\hat{\beta}_t$ is the least squares estimate of β . This is the value of the control which would be used if one treated β as known with certainty and equal to the least squares estimate. This has been called the certainty equivalence rule. In other problems, where the unknown parameters are additive, certainty equivalence rules have been shown to be Bayes optimal rules. In this problem, of course, (2.19) is not equal to the Bayes control rule. We call this rule the least squares certainty equivalence rule. It is of particular interest to investigate the properties of this rule since we expect that it is frequently used in practical applications.

A related control rule would be preferred to (2.19) if there were some prior knowledge about the unknown parameter β. This prior knowledge may be due to some observations which have been made before the control problem starts. Such a control rule is defined by the sequence

$$(2.20) \qquad u_{t+1} = \frac{a}{b_t}, \quad t = 0, 1, \ldots,$$

where b_t is the mean of the posterior distribution of β at time t. If one were estimating β with Bayes procedures and quadratic loss, then b_t would be the Bayes estimate for β. For this reason we call this rule the Bayesian certainty equivalence control rule and hope that the apparent contradiction is not confusing.

A third rule which takes the uncertainty about the knowledge of β into account, but not experimentation, is the Bayesian myopic control rule defined by the sequence

$$(2.21) \qquad u_{t+1} = \frac{ab_t}{b_t^2 + \sigma_t^2}, \quad t = 0, 1, \ldots .$$

This is the result of minimizing the squared error loss at each time period as if the problem had no future.

Analogous control rules can be defined for the model with unknown intercept; corresponding to (2.19), (2.20) and (2.21) respectively these are:

$$u_{t+1} = \frac{a - \hat{\beta}_{1t}}{\hat{\beta}_{2t}}, \quad t = 1, 2, \ldots, \tag{2.22}$$

$$u_{t+1} = \frac{a - b_{1t}}{b_{2t}}, \quad t = 0, 1, \ldots, \tag{2.23}$$

$$u_{t+1} = \frac{(a - b_{1t})b_{2t} - \sigma_{12t}}{b_{2t}^2 + \sigma_{22t}}, \quad t = 0, 1, \ldots . \tag{2.24}$$

Each of these control rules are in the set of available control rules defined in the description of the decision model. They are functions of the observed data and describe what control value to use at any time under any eventuality.

2.5 Concluding Remarks on the Bayesian Method

In the context of the statistical decision model, the Bayes principle for ordering control rules was discussed in this chapter. The procedure has three possible difficulties: (1) the risk may be impossible to calculate even with existing numerical methods, (2) the risk may not exist for some loss functions, and (3) the resulting control rule depends on the prior distribution of the parameters.

The main purpose of the next two chapters is to develop control criteria which do not have these difficulties. We study the behavior of various control rules under less specific distributional assumptions on the error term and without any prior distribution assumptions on the unknown parameters.

CHAPTER III

ASYMPTOTIC PROPERTIES OF THE CONTROLS AND PARAMETER ESTIMATES

3.1 Introduction

In this chapter we investigate the asymptotic behavior of two of the control rules defined and discussed in Chapter II. Because the control rules are functions of the estimates of the unknown parameters, the study of their asymptotic behavior is similar to that of statistical estimates. However, the problem is complicated by the fact that each control value becomes a data point which is used to estimate the control in the next period. That is, the use of controls necessarily implies choosing the data for estimation. In terms of the regression model this means that the regressors must be treated as random variables whose behavior is determined by previous estimates of the parameters. The problem is therefore similar to the autoregressive model [see Anderson (1959)] except that the structure relating one regressor to the next is more complicated. Because of this more complicated structure we will find it necessary to use methods which are more general than those that have been used in the autoregressive model. The methods include nonprobabilistic lemmas about sequences of numbers and the use of martingales. Because of this greater generality the methods might be useful in other studies where predetermined variables are complicated functions of other random variables.

In Section 3.2 we consider the model where the intercept is known and the slope is unknown. We prove that the least squares certainty equivalence and the Bayesian certainty equivalence control rules converge with probability one to the true value. Convergence to the true value is certainly a criterion which a good control rule must satisfy. But it is a rather weak criterion in that it does not indicate how fast a particular control rule converges. Speed of convergence is particularly important in models with unknown parameters, since it may be possible to experiment with controls in order to get better estimates and thus faster convergence. To answer these questions, in the second part of Section 3.2 we derive the asymptotic distribution of these two control rules and the corresponding parameter estimates. The normalization on these distributions then measures the speed of convergence. These asymptotic distributions can also be used for testing hypotheses and making confidence statements about the unknown parameters.

In Section 3.3 we briefly consider the more general model where the intercept as well as the slope is unknown. This case is more subtle than in the model with known intercept. The difficulty is due to the fact that, as the control converges to a particular point, the information about the individual parameters gets smaller. Thus, if convergence were too fast, one might obtain inconsistent estimates of parameters and of the control. Further, even if the control converged to the true value, the estimates of the individual parameters might not be consistent. A rigorous mathematical derivation of the asymptotic properties, (on the level of Section 3.1) remains to be done in this case. Section 3.2 provides a brief discussion of some of the mathematical difficulties and possible approaches.

For completeness we might also consider the case where the intercept is unknown and the slope is known. However, the convergence proofs are rather straightforward in this case and follow directly from the strong law of large numbers.

3.2 The Model with Unknown Slope and Known Intercept

3.2.1 The Model and Assumptions

We first consider the model with an unknown intercept;[1] that is,

$$x_t = \beta u_t + \epsilon_t \ , \tag{3.1}$$

where u_t is used to control x_t about some desired level a, where β is an unknown parameter and where $\{\epsilon_t\}$ is an independent sequence with zero mean and finite variance σ^2. Then the least squares certainty equivalence control rule is u_1 fixed and nonzero and

$$u_{t+1} = \frac{a}{\hat{\beta}_t} \ , \ t = 1, 2, \ldots, \tag{3.2}$$

and the Bayesian certainty equivalence control rule is

[1] See footnote 1 in Section 2.3.1.

(3.3) $$u_{t+1} = \frac{a}{b_t}, \quad t = 0, 1, \ldots,$$

where

$$\hat{\beta}_t = \frac{\sum_{i=1}^{t} u_i x_i}{\sum_{i=1}^{t} u_i^2}$$

is the least squares estimate, and where

$$b_t = \frac{\frac{b_0}{\sigma_0^2} + \frac{1}{\sigma^2} \sum_{i=1}^{t} u_i x_i}{\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2} \sum_{i=1}^{t} u_i^2}$$

and

$$\frac{1}{\sigma_t^2} = \frac{1}{\sigma_0^2} + \frac{1}{\sigma^2} \sum_{i=1}^{t} u_i^2$$

are the mean and variance of the posterior distribution, which is normal when the prior distribution is $N(b_0, \sigma_0^2)$ and the error term is distributed $N(0, \sigma^2)$. The rationale behind these control rules was discussed in the last chapter.

The theorems which follow are not based on the distribution assumptions of the Bayesian method. These assumptions are only used to suggest the Bayesian certainty equivalence control. The following analysis assumes only that the error terms are independently and identically distributed with zero mean and finite variance. Subject to these conditions the distribution may have any form.

3.2.2 Convergence of Control Rules

In this section we show that under suitable conditions the two control rules defined above converge to the value a/β with probability 1. This is the value which would be used if the unknown parameter were known with certainty. We choose to prove convergence with probability 1 rather than convergence in probability (which is usually enough in econometric investigations) for three reasons. First, since we are interested in bringing the endogenous variable to its desired level for all time periods, we would hope that the control converges to the true value and stays there with high probability. This is guaranteed by convergence with probability 1, but not by convergence in probability. The latter says only that, for any sufficiently large fixed t, the probability that the control is near its true value is arbitrarily close to one. Convergence in probability is enough in most econometric investigations because one is usually referring to some fixed sample.

Secondly, for comparison with some related sequential statistical decision problems, it is necessary to prove strong convergence. For example, some rules in the two-armed bandit problem have a positive probability of converging to the wrong value [Rothschild (1971)], while in stochastic approximation, most rules do converge with probability one to the true value [Wetherill (1966)]. To compare the results of this paper with such problems, it is necessary to examine strong convergence.

A third reason, related to the first, is that the use of strong convergence allows one to use some non-probabilistic results for arbitrary sequences of numbers. Once a sequence of random variables is shown to have a certain property with probability 1, we then can ignore all sample points where the property does not occur and apply nonprobabilistic results to the remaining points. This technique is quite useful for the control problem of this paper where the structure of the random sequences is quite complicated.

To show convergence we first prove three preliminary lemmas, of which the first two are nonprobabilistic.

LEMMA 1: _Let_ $\{z_t\}$ _be an arbitrary sequence of numbers such that_ $z_1 \neq 0$. _If_

$$s_t = \sum_{i=1}^{t} \left(\frac{z_i^2}{\sum_{j=1}^{t} z_j^2} \right)^2 , \tag{3.4}$$

then $s_t < 2/z_1^2$ _for every_ t .

PROOF: For each t we can maximize s_t with respect to z_t , z_{t-1}, ..., z_2 in turn to obtain an upper bound. For $t = 2$ we can maximize with respect to z_2 to show that

$$s_2 = \frac{1}{z_1^2} + \frac{z_2^2}{(z_1^2 + z_2^2)^2} \leq \frac{1}{z_1^2} + \frac{1}{4z_1^2} = \frac{1}{z_1^2}(1 + \frac{1}{a_2}) , \tag{3.5}$$

where $a_2 = 4$. For $t = 3$, by maximizing with respect to z_3 and then z_2 we have

$$(3.6) \qquad s_3 = \frac{1}{z_1^2} + \frac{z_2^2}{(z_1^2 + z_2^2)^2} + \frac{z_3^2}{(z_1^2 + z_2^2 + z_3^2)^2}$$

$$\le \frac{1}{z_1^2} + \frac{z_2^2}{(z_1^2 + z_2^2)^2} + \frac{1}{4(z_1^2 + z_2^2)}$$

$$\le \frac{1}{z_1^2}\left(1 + \frac{1}{a_3}\right) ,$$

where $a_3 = \left(1 + \frac{a_2 - 1}{a_2 + 1}\right)^2$. Similarly, for any t we have

$$(3.7) \qquad s_t \le \frac{1}{z_1^2}\left(1 + \frac{1}{a_t}\right) ,$$

where

$$a_t = \left(1 + \frac{a_{t-1} - 1}{a_{t-1} + 1}\right)^2 .$$

Since $a_t > 1$ for all t , we have

$$(3.8) \qquad s_t < \frac{2}{z_1^2} .$$

LEMMA 2: Let $\{z_t\}$ be a sequence of numbers and let $\{a_t\}$ be an increasing sequence of positive numbers such that $\Sigma_{i=1}^t z_i/a_i$ converges,

(i) If $a_t \to \infty$, then $\lim_{t \to \infty} \frac{1}{a_t} \sum_{i=1}^{t} z_i = 0$,

(ii) If $a_t \to M < \infty$, then $\lim_{t \to \infty} \frac{1}{a_t} \sum_{i=1}^{t} z_i$ exists.

PROOF: Part (i) is Kronecker's lemma. [See Feller (1966), p. 239.] We need only consider part (ii). Define $s_0 = 0$ and let

$$s_t = \sum_{i=1}^{t} z_i/a_i \ , \qquad t = 1, 2, \ldots, \tag{3.9}$$

then

$$z_t = a_t(s_t - s_{t-1}) \ , \tag{3.10}$$

and therefore

$$\begin{aligned} \frac{1}{a_t} \sum_{i=1}^{t} z_i &= \frac{1}{a_t} \sum_{i=1}^{t} a_i(s_i - s_{i-1}) \\ &= s_t - \frac{1}{a_t} \sum_{i=1}^{t-1} (a_{i+1} - a_i)s_i \ . \end{aligned} \tag{3.11}$$

Now by assumption, s_t converges to s , say, so that to complete the proof of the lemma we must show that the second term on the right hand side of (3.11) converges.

For an arbitrary $\epsilon > 0$ choose t_0 so that, for all $t > t_0$, $|s_t - s| < \epsilon$. Such a t_0 exists by the convergence assumption. We then have

$$(3.12) \qquad \frac{1}{a_t}\sum_{i=1}^{t-1}(a_{i+1} - a_i)s_i$$

$$= \frac{1}{a_t}(a_t - a_1)s + \frac{1}{a_t}\sum_{i=1}^{t_0-1}(a_{i+1} - a_i)(s_i - s)$$

$$+ \frac{1}{a_t}\sum_{i=t_0}^{t-1}(a_{i+1} - a_i)(s_i - s) \quad .$$

Now, because a_t converges and t_0 is fixed, the first two terms on the right-hand side converge. Further

$$(3.13) \qquad \left|\frac{1}{a_t}\sum_{i=t_0}^{t-1}(a_{i+1} - a_i)(s_i - s)\right| < \frac{1}{a_t}(a_t - a_{t_0})\,\epsilon < \epsilon \ ,$$

and, since ϵ is arbitrary, the third term in (3.12) is arbitrarily small. Thus $\Sigma_{i=1}^{t} z_i/a_t$ converges.

The following lemma is probabilistic and uses the martingale convergence theorem. [See Feller (1966), p. 242.]

LEMMA 3: Let $\{\epsilon_i\}$ be an independent sequence of random variables with $E\epsilon_i = 0$ and $E\epsilon_i^2 = \sigma^2 < \infty$ and let $\{u_i\}$ be a sequence of random variables with u_1 fixed and nonzero and ϵ_i independent of $\{u_i, u_{i-1}, \ldots, u_1, \epsilon_{i-1}, \ldots, \epsilon_1\}$, $i = 2, 3, \ldots$. Then

$$(3.14) \qquad s_t = \sum_{i=1}^{t} \frac{u_i \epsilon_i}{\sum_{j=1}^{i} u_j^2}$$

converges with probability 1.

PROOF: We have that

$$(3.15) \qquad E\left(\frac{u_i \epsilon_i}{\sum_{j=1}^{i} u_j^2} \,\middle|\, s_{i-1}, \ldots, s_1\right)$$

$$= E\left(\frac{u_i}{\sum_{j=1}^{i} u_j^2} \,\middle|\, s_{i-1}, \ldots, s_1\right) E(\epsilon_i | s_{i-1}, \ldots, s_1)$$

$$= E\left(\frac{u_i}{\sum_{j=1}^{i} u_j^2} \,\middle|\, s_{i-1}, \ldots, s_1\right) \cdot 0$$

$$= 0$$

since $|u_i|/\Sigma_{j=1}^{i} u_j^2 \leq 1/|u_1|$. Thus $\{s_t\}$ is a martingale and to use the martingale convergence theorem it must be shown that Es_t^2 remains bounded for all t . From the independence assumptions we have

(3.16) $$E \sum_{i=1}^{t} \left(\frac{u_i \epsilon_i}{\sum_{j=1}^{i} u_j^2} \right)^2 = \sigma^2 E \left(\frac{u_i}{\sum_{j=1}^{i} u_j^2} \right)^2 \leq \sigma^2 \frac{2}{u_1^2} ,$$

where the last inequality follows from LEMMA 1 with $z_1 = u_1$. Thus the variance remains bounded for all t and by the martingale convergence theorem s_t converges with probability 1.

The following theorem contains the main convergence results about the multiperiod control rules. The proof involves showing that, with probability 1, each control rule does not stop obtaining information about the unknown parameter.

THEOREM 1: _In the model_ $x_t = \beta u_t + \epsilon_t$, _if_ $\{\epsilon_t\}$ _is an independent sequence of random variables with_ $E\epsilon_t = 0$ _and_ $E\epsilon_t^2 = \sigma^2 < \infty$ _and_ $\beta \neq 0$, _then (i) the least squares certainty equivalence control rule converges to_ a/β _with probability 1, and (ii) if_ $b_0 \neq 0$ _and_ $\sigma_0^2 \neq 0$ _then the Bayesian certainty equivalence control rule converges to_ a/β _with probability 1_.

PROOF: (i) The least squares certainty equivalence control rule can be written

(3.17) $$u_{t+1} = \frac{a}{\beta + \frac{\sum_{i=1}^{t} u_i \epsilon_i}{\sum_{i=1}^{t} u_i^2}} .$$

We first must establish that $\Sigma_{i=1}^{t} u_i^2 \to \infty$ with probability 1. Let ω be any sample point in the sample space Ω. Then we have from LEMMA 3 that

$$(3.18) \qquad P\left[\omega \,\middle|\, \sum_{i=1}^{t} \frac{u_i(\omega)\varepsilon_i(\omega)}{\sum_{j=1}^{i} u_j^2(\omega)} \text{ converges}\right] = 1 \ .$$

Thus we can apply LEMMA 2, parts (i) and (ii), at each sample point with $z_i = u_i(\omega)\varepsilon_i(\omega)$ and $a_i = \Sigma_{j=1}^{i} u_j^2(\omega)$ to obtain

$$(3.19) \qquad P\left[\omega \,\middle|\, \frac{\sum_{i=1}^{t} u_i(\omega)\varepsilon_i(\omega)}{\sum_{i=1}^{t} u_i^2(\omega)} \text{ converges}\right] = 1 \ .$$

But this implies that

$$(3.20) \qquad P\left[\omega \,\middle|\, \lim_{t \to \infty} a\left(\beta + \frac{\sum_{i=1}^{t} u_i(\omega)\varepsilon_i(\omega)}{\sum_{i=1}^{t} u_i^2(\omega)}\right)^{-1} \neq 0\right] = 1$$

and, from (3.17)

$$(3.21) \qquad P[\omega | \lim_{t \to \infty} u_{t+1}(\omega) \neq 0] = 1 \ ,$$

and therefore we have that

$$(3.22)\qquad P[\omega|\sum_{i=1}^{t} u_i^2(\omega) \text{ diverges}] = 1 .$$

Having proved that $\Sigma_{i=1}^{t} u_i^2 \to \infty$ with probability 1, we can now apply LEMMA 2(i) at every sample point to obtain

$$(3.23)\qquad P\left[\omega\left|\frac{\sum_{i=1}^{t} u_i(\omega)\epsilon_i(\omega)}{\sum_{i=1}^{t} u_i^2(\omega)} \to 0\right.\right] = 1 ,$$

and from (3.17) this implies that

$$(3.24)\qquad u_{t+1} \to \frac{a}{\beta}$$

with probability 1.

(ii) The argument for the Bayesian certainty equivalence control rule is similar, except that we must insure that the weights on the prior parameters converge to zero with probability one. With Bayesian estimates we have

(3.25)
$$u_{t+1} = a\left[\frac{\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2}\sum_{i=1}^{t} u_i^2}{\frac{b_0}{\sigma_0^2} + \frac{1}{\sigma^2}\left(\beta \sum_{i=1}^{t} u_i^2 + \sum_{i=1}^{t} u_i \epsilon_i\right)}\right]$$

$$= \frac{a\left(\frac{\sigma^2}{\sigma_0^2}\frac{1}{\sum_{i=1}^{t} u_i^2} + 1\right)}{\frac{b_0}{\sigma_0^2}\frac{\sigma^2}{\sum_{i=1}^{t} u_i^2} + \beta + \frac{\sum_{i=1}^{t} u_i \epsilon_i}{\sum_{i=1}^{t} u_i^2}} .$$

Now, since

(3.26)
$$\frac{\sigma^2}{\sigma_0^2}\frac{1}{\sum_{i=1}^{t} u_i^2} + 1$$

is nonzero, we can use the argument of equations (3.18) and (3.19) to show that

(3.27)
$$P[\omega | \lim_{t \to \infty} u_{t+1}(\omega) \neq 0] = 1 ,$$

and therefore

(3.28)
$$P[\omega | \sum_{i=1}^{t} u_i^2(\omega) \text{ diverges}] = 1 .$$

Thus, from LEMMA 2(i) applied at every sample point,

$$(3.29) \qquad P\left[\omega \,\middle|\, \frac{\sum_{i=1}^{t} u_i(\omega)\epsilon_i(\omega)}{\sum_{i=1}^{t} u_i^2(\omega)} \to 0\right] = 1 ,$$

and also

$$(3.30) \qquad P\left[\omega \,\middle|\, \frac{1}{\sum_{i=1}^{t} u_i^2(\omega)} \to 0\right] = 1 \quad .$$

From equation (3.25) this implies that

$$(3.31) \qquad u_t \to \frac{a}{\beta}$$

with probability 1.

This completes the proof of THEOREM 1 from which we have the following corollary which states that both the least squares estimate and the Bayesian estimate of β are strongly consistent.

COROLLARY 1: Under the assumptions of THEOREM 1

(i) $\hat{\beta}_t \to \beta$ with probability 1, and

(ii) $b_t \to \beta$ and $\sigma_t^2 \to 0$ with probability 1.

PROOF: Once we have established that $\Sigma_{i=1}^{t} u_i^2 \to \infty$ with probability 1, the corollary follows immediately from LEMMA 2(i) as described in the proof of THEOREM 1.

3.2.3 The Asymptotic Distribution of Control Rules

Additional information about the behavior of multiperiod control rules can be obtained by examining their asymptotic distributions. To obtain these distributions we first derive the asymptotic distributions of the estimates of the unknown parameter β . We begin by proving a preliminary lemma.[2]

LEMMA 4: Let $\{\nu_i\}$ be a sequence of random variables such that $\nu_i \to 0$ with probability 1 and let $\{\epsilon_i\}$ be an independent sequence of random variables with $E\epsilon_i = 0$ and $E\epsilon_i^2 = \sigma^2 < \infty$ and ϵ_i independent of $\{\epsilon_{i-1}, \ldots, \epsilon_1, \nu_i, \nu_{i-1}, \ldots, \nu_1\}$, $i = 2, 3, \ldots$. Then $\Sigma_{i=1}^{t} \nu_i \epsilon_i / \sqrt{t} \xrightarrow{p} 0$.

PROOF: We use a method of truncation. Define ν_i' and ν_i'' as

---

[2] In the following lemma and theorems we use the notation "$\xrightarrow{p}$" for converges in probability and "$\xrightarrow{d}$" for converges in distribution. With the exception of LEMMA 4 we use only the weak consistency property of the control rules and parameter estimates to derive the asymptotic distributions.

(3.32) $$\nu_i' = \nu_i \,, \quad \nu_i'' = 0 \quad \text{if} \quad |\nu_i| < C \,,$$

$$\nu_i' = 0 \,, \quad \nu_i'' = \nu_i \quad \text{if} \quad |\nu_i| \geq C \,,$$

then $\nu_i = \nu_i' + \nu_i''$ and

(3.33) $$\frac{1}{\sqrt{t}} \sum_{i=1}^{t} \nu_i \varepsilon_i = \frac{1}{\sqrt{t}} \sum_{i=1}^{t} \nu_i' \varepsilon_i + \frac{1}{\sqrt{t}} \sum_{i=1}^{t} \nu_i'' \varepsilon_i \quad .$$

Now, since ν_i' is bounded and since $\nu_i' \to 0$ with probability 1 we have that $E(\nu_i')^2 \to 0$. Therefore

(3.34) $$E\left(\frac{\sum_{i=1}^{t} \nu_i' \varepsilon_i}{\sqrt{t}}\right)^2 = \frac{\sigma^2 \sum_{i=1}^{t} E(\nu_i')^2}{t} \to 0 \,,$$

and by Chebyshev's inequality $\Sigma_{i=1}^{t} \nu_i' \varepsilon_i / \sqrt{t} \xrightarrow{P} 0$.

It remains to consider $\Sigma_{i=1}^{t} \nu_i'' \varepsilon_i / \sqrt{t}$. From the definition of ν_i'' we have

(3.35) $$P[\nu_i'' \neq 0] = P[|\nu_i| \geq C] \quad .$$

But since $\nu_i \to 0$ with probability 1

(3.36) $$P[\omega \,|\, |\nu_i(\omega)| \geq C \text{ infinitely often}] = 0 \,.$$

Therefore,

(3.37) $$P[\omega | \nu''_i(\omega) \neq 0 \text{ infinitely often}] = 0 ,$$

so that

(3.38) $$P[\omega | \sum_{i=1}^{\infty} \nu''_i(\omega)\epsilon_i(\omega) \text{ has a finite number of nonzero terms}] = 1 ,$$

and therefore

(3.39) $$P\left[\omega \left| \frac{\sum_{i=1}^{t} \nu''_i(\omega)\epsilon_i(\omega)}{\sqrt{t}} \rightarrow 0 \right.\right] = 1 .$$

From (3.33) we therefore have that $\Sigma_{i=1}^{t} \nu_i\epsilon_i/\sqrt{t} \xrightarrow{p} 0$.

THEOREM 2: <u>Under the assumptions of THEOREM 1,</u>

(i) $$\sqrt{t}(\hat{\beta}_t - \beta) \xrightarrow{d} N(0, \frac{\beta^2}{a^2}\sigma^2)$$

and

(ii) $$\sqrt{t}(b_t - \beta) \xrightarrow{d} N(0, \frac{\beta^2}{a^2}\sigma^2) .$$

PROOF: (i) To find the limiting distribution of $\sqrt{t}(\hat{\beta}_t - \beta)$, we have from the definition of the least squares estimate,

(3.40)
$$\sqrt{t}(\hat{\beta}_t - \beta) = \frac{t}{\sum_{i=1}^{t} u_i^2} \frac{\sum_{i=1}^{t} u_i \epsilon_i}{\sqrt{t}}$$

$$= \frac{t}{\sum_{i=1}^{t} u_i^2} \left[\frac{\sum_{i=1}^{t} (u_i - \frac{a}{\beta})\epsilon_i}{\sqrt{t}} + \frac{a}{\beta} \frac{\sum_{i=1}^{t} \epsilon_i}{\sqrt{t}} \right] .$$

From LEMMA 4 the first term in brackets converges to zero in probability with $\nu_i = u_i - \frac{a}{\beta}$. In addition, from THEOREM 1 we have that $u_i^2 \to (a/\beta)$ with probability 1, so that[3]

(3.41)
$$\frac{\sum_{i=1}^{t} u_i^2}{t} \to (\frac{a}{\beta})^2$$

with probability 1. Therefore, the difference between the right hand side of equation (3.40) and

(3.42)
$$\frac{\beta}{a} \frac{\sum_{i=1}^{t} \epsilon_i}{\sqrt{t}}$$

converges in probability to zero. By the central limit theorem (3.42)

---

[3]If a sequence converges then the arithmetic mean of the sequence also converges to the same point. [See Knopp (1956), p. 35.] We apply this result at every sample point to obtain the result of equation (3.41).

converges in distribution to $N(0, \beta^2\sigma^2/a^2)$.

(ii) To find the limiting distribution of $\sqrt{t}(b_t - \beta)$ we have, from the definition of b_t

(3.43) $$\sqrt{t}(b_t - \beta) = \frac{\sqrt{t}\left(\dfrac{(b_0 - \beta)\sigma^2}{\sigma_0^2 \sum_{i=1}^{t} u_i^2} + \dfrac{\sum_{i=1}^{t} u_i \epsilon_i}{\sum_{i=1}^{t} u_i^2}\right)}{\dfrac{\sigma^2}{\sigma_0^2 \sum_{i=1}^{t} u_i^2} + 1} .$$

Now, since $u_i \to a/\beta$ with probability 1, we have

(3.44) $$\frac{\sqrt{t}}{\sum_{i=1}^{t} u_i^2} \to 0$$

with probability 1, using the result of equation (3.41). Therefore equation (3.43) converges in probability to equation (3.40) and we can apply the same argument as in part (i) to show that $\sqrt{t}(b_t - \beta)$ has the same limiting distribution as $\sqrt{t}(\hat{\beta}_t - \beta)$.

The results of THEOREM 2 can now be used to derive the asymptotic distribution of the control rules themselves in the following theorem.

THEOREM 3: <u>Under the assumptions of THEOREM 1, if $\{u_t\}$ is defined as either (i) the least squares certainty equivalence control rule, or (ii) the Bayesian certainty equivalence control rule, then</u>

(3.45) $$\sqrt{t}(u_t - \frac{a}{\beta}) \xrightarrow{d} N(0, \frac{\sigma^2}{\beta^2}) \quad .$$

PROOF: (i) The limiting distribution of $\sqrt{t}(u_t - a/\beta)$ in the least squares case follows from

(3.46) $$\sqrt{t}(\frac{a}{\hat{\beta}_t} - \frac{a}{\beta}) = \frac{a}{\hat{\beta}_t \beta}\sqrt{t}(\beta - \hat{\beta}_t) \quad .$$

From COROLLARY 1(i) $\hat{\beta}_t \xrightarrow{p} \beta$, so that the difference between the right hand side of equation (3.46) and $\sqrt{t}(\beta - \hat{\beta}_t)a/\beta^2$ converges to zero in probability. The first part of the THEOREM then follows from

(3.47) $$\sqrt{t}(\beta - \hat{\beta}_t)a/\beta^2 \xrightarrow{d} N(0, \frac{\sigma^2}{\beta^2}) \; ,$$

which follows directly from THEOREM 2(i).

(ii) Similarly in the case of Bayesian certainty equivalence control we have

(3.48) $$\sqrt{t}(\frac{a}{b_t} - \frac{a}{\beta}) = \frac{a}{b_t \beta}\sqrt{t}(b_t - \beta) \quad ,$$

and from COROLLARY 1(ii), $b_t \xrightarrow{p} \beta$, so that the difference between the right hand side of equation (3.48) and $\sqrt{t}(\beta - b_t)a/\beta^2$ converges to zero in probability. From THEOREM 2(ii) we have

$$\sqrt{t}(\beta - b_t)a/\beta^2 \xrightarrow{d} N(0, \frac{\sigma^2}{\beta^2}) \tag{3.49}$$

which completes the proof of the THEOREM.

It is interesting to note that the two control rules have the same asymptotic distribution. This is due to the fact that in the second rule the weights on the prior parameters converge to zero faster than $\sqrt{t}$ (as expressed by the fact that $\Sigma_{i=1}^{t} u_i^2 = O(t)$ with probability 1). Note also that the desired level a does not affect the asymptotic variance. Further the asymptotic variance depends inversely on the square of the slope.

Another way to look at these asymptotic distributions is that

$$\sqrt{t}(\beta u_t - a) \xrightarrow{d} N(0, \sigma^2) \quad . \tag{3.50}$$

That is, the variance of $(\beta u_t - a)$ decreases to zero approximately as σ^2/t . Whether this speed of convergence is as fast as one could hope for, even with other rules using experimentation, is discussed below.

3.3 The Model with Unknown Slope and Unknown Intercept

In this section we discuss control rules in the more general regression model

$$x_t = \beta_1 + \beta_2 u_t + \epsilon_t \ , \qquad t = 1, 2, \ldots, \tag{3.51}$$

where β_1 and β_2 are unknown, the assumptions on ϵ_t are the same as in Section 3.2 and the desired level of x_t remains fixed at a .

To show that a particular control rule converges to the time value $(a - \beta_1)/\beta_2$ with probability 1, is considerably more difficult in this model with two unknown parameters. As in the known intercept case the regressor must be treated as a random variable whose behavior is determined by previous estimates of parameters. But an additional complication is that, in order to consistently estimate two unknown parameters, the observations on the regressors must be sufficiently spread out. However, if the control rule converges to a constant (as we would like to show), then the observations do not have as much spread as they would without convergence. To establish that these observations have sufficient spread to obtain enough information about the true control value, it is therefore necessary to show that the speed of convergence is not too rapid. Along with some mathematical technicalities, establishing this speed of convergence is the essential difficulty in this model.

To illustrate these points we will consider the least squares certainty equivalence control rule defined by: u_1 and u_2 fixed and distinct and

$$u_{t+1} = \frac{a - \hat{\beta}_{1t}}{\hat{\beta}_{2t}} , \quad t = 2, 3, \ldots, \tag{3.52}$$

where $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$ are the least squares estimates of β_1 and β_2 defined by

(3.53)
$$\begin{pmatrix} \hat{\beta}_{1t} \\ \hat{\beta}_{2t} \end{pmatrix} = \begin{pmatrix} t & \sum_{i=1}^{t} u_i \\ \sum_{i=1}^{t} u_i & \sum_{i=1}^{t} u_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{t} x_i \\ \sum_{i=1}^{t} u_i x_i \end{pmatrix}$$

Substituting from (3.53) into (3.52) we obtain

(3.54)
$$u_{t+1} = \frac{a - \beta_1 + \bar{u}_t \dfrac{\sum_{i=1}^{t} (u_i - \bar{u}_t)\epsilon_i}{\sum_{i=1}^{t} (u_i - \bar{u}_t)^2} + \dfrac{\sum_{i=1}^{t} \epsilon_i}{t}}{\beta_2 + \dfrac{\sum_{i=1}^{t} (u_i - \bar{u}_t)\epsilon_i}{\sum_{i=1}^{t} (u_i - \bar{u}_t)^2}} .$$

Since $\Sigma_{i=1}^{t} \epsilon_i/t$ converges to zero with probability 1 by the strong law of large numbers, we could show that u_t converges to $(a - \beta_1)/\beta_2$ if

(3.55)
$$\frac{\sum_{i=1}^{t} (u_i - \bar{u}_t)\epsilon_i}{\sum_{i=1}^{t} (u_i - \bar{u}_t)^2}$$

converges to zero with probability 1.

This latter expression is similar to the second term in the denominator of (3.17) except that u_i is replaced by $(u_i - \bar{u}_t)$. Thus instead of showing that $\Sigma_{i=1}^{t} u_i^2$ diverges with probability 1, we must show that $\Sigma_{i=1}^{t}(u_i - \bar{u}_t)^2$ diverges with probability 1. This latter sum will diverge if u_t does not converge to a constant too rapidly.

An additional problem is that we cannot use methods analogous to Section 3.2.2, even if we could prove that $\Sigma_{i=1}^{t}(u_i - \bar{u}_t)^2$ diverges. In that section we consider convergence of the martingale

$$s_t = \sum_{i=1}^{t} \frac{u_i \epsilon_i}{\sum_{j=1}^{i} u_j^2} , \tag{3.56}$$

and then use LEMMA 2(i) to show that

$$\frac{\sum_{i=1}^{t} u_i \epsilon_i}{\sum_{i=1}^{t} u_i^2}$$

converges to zero with probability 1. But the expression equivalent to (3.56) in this case is

$$\sum_{i=1}^{t} \frac{(u_i - \bar{u}_t)\epsilon_i}{\sum_{j=1}^{i} (u_j - \bar{u}_i)^2} \tag{3.57}$$

which is not a martingale because ϵ_i is not independent of $\bar{u}_t$ for $i < t$. Thus a different method of proof must be used.

One possibility is to consider a vector analogy to the proof in Section 3.2.2. Letting the expression

$$(3.58)\qquad \begin{pmatrix} \hat{\beta}_{1t} - \beta_1 \\ \\ \hat{\beta}_{2t} - \beta_2 \end{pmatrix} = \begin{pmatrix} t & \sum_{i=1}^{t} u_i \\ \\ \sum_{i=1}^{t} u_i & \sum_{i=1}^{t} u_i^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{t} \epsilon_i \\ \\ \sum_{i=1}^{t} u_i \epsilon_i \end{pmatrix}$$

correspond to $\hat{\beta} - \beta = \Sigma_{i=1}^{t} u_i \epsilon_i / \Sigma_{i=1}^{t} u_i^2$, the related martingale would be

$$(3.59)\qquad \sum_{i=1}^{t} \left[\sum_{j=1}^{i} \begin{pmatrix} 1 & u_j \\ u_j & u_j^2 \end{pmatrix} \right]^{-1} \begin{pmatrix} \epsilon_i \\ u_i \epsilon_i \end{pmatrix} ,$$

This vector series can be shown to converge with probability 1 using the martingale convergence theorem. Thus to show that (3.58) converges to the zero vector, we need a vector version of LEMMA 2. Unfortunately, such a lemma is not true in the full generality of LEMMA 2. A possibility for future research would be to find conditions under which such a lemma is true and if it can be applied here.

A problem related to convergence of the control is convergence of the parameter estimates. Even if the control converges to the true value it is possible that the individual parameter estimates do not converge to their true values. Again the speed of convergence of the control is crucial.

Even if these estimates were consistent, their standard errors would be larger than in the usual regression model where the data is spread out such that $\Sigma_{i=1}^{t}(u_i - \bar{u}_t)^2/t$ does not converge to zero. That is, in the usual regression model the variance of the estimates decreases approximately at the rate $1/t$, while in this model the rate will be as $1/\Sigma_{i=1}^{t}(u_i - \bar{u}_t)^2$, which may be considerably less than $1/t$.

If these convergence results could be established, further research might then consider the asymptotic distribution of the control rule. Intuition suggests that the limiting distribution of the control minus its true value normalized by $\sqrt{t}$ would be normal with mean zero and variance σ^2/β_2^2 . To see this consider the identity

$$(3.60)\quad \sqrt{t}\left(u_{t+1} - \frac{a - \beta_1}{\beta_2}\right) = \frac{1}{\hat{\beta}_{2t}}\left[\sqrt{t}\left(\bar{u}_t - \frac{a - \beta_1}{\beta_2}\right)(\hat{\beta}_{2t} - \beta_2) - \frac{\sum_{i=1}^{t} \epsilon_i}{\sqrt{t}}\right] .$$

The conjectured limiting distribution would follow if the first term in brackets converges to zero in probability. If the slope is consistently estimated and if the rate of convergence of the mean is faster than of the control itself, then the result would follow.

3.4 Concluding Remarks on the Theorems and Methods of Proof

In this chapter we have discussed the properties of some possible control rules in the linear control model. The results have been independent of any prior distributions and are therefore more general than the usual Bayesian approach. In the model with known intercept we have shown that the various decision rules converge to the true value with probability 1. We made use of the fact that even though the regressors are random, we could obtain full information about the unknown slope with probability 1. The intuitive idea was to show that with any sequence of controls, which follow the suggested decision rule, the sum $\Sigma_{i=1}^{t} u_i^2$ diverges with probability 1.

Throughout these convergence proofs we used methods which are not usually employed in econometric analysis; namely strong consistency and martingales. While the strong convergence property is certainly more desirable than the weak one, especially in dynamic problems, we also found these methods particularly well adapted to this problem. In econometric analysis, if the regressors are random, they are usually assumed to follow some relatively simple process (such as an autoregressive scheme). This structure can then be used in the proof of consistency. However, in the control problem the structure relating one regressor to the next is quite complicated, and after only a small number of steps the analytic representation is impossible. We therefore found it necessary to find a method which was more general and did not depend on the specific structure of the controls. This method is summarized in the preliminary lemmas to THEOREM 1. Because of the generality of these procedures it is hoped that these methods might have use in other econometric problems where the regressors are random with a complicated structure.

In addition to proving convergence of these control rules, we were also able to obtain the asymptotic distribution of the controls and the parameter estimates in the model with known intercept. These distributions will be used in the next chapter as criteria of control which are alternatives to the usual Bayesian methods. In addition, they can be used for making approximate tests of significance and confidence statements. Besides the usual central limit theorem, we used a method of truncation in these proofs. Although convergence in probability is sufficient in deriving asymptotic distributions, the method of truncation requires that the control converge with probability 1. Therefore, we used the strong consistency property in proving asymptotic normality.

In the model with both unknown intercept and unknown slope we have not proved convergence or asymptotic normality, and this would certainly be the prime focus of future research.

CHAPTER IV

ASYMPTOTIC EFFICIENCY AND CRITERIA OF CONTROL

4.1 Introduction

In this chapter we consider how the asymptotic normality results of Chapter III might be used as criteria for judging the effectiveness of the control rules, as well as for suggesting whether there exist other rules which might do better. Once a particular control rule has been decided upon, its behavior over time will depend on the data generated by the random disturbance term. That is, each element of a control rule (whether it is Bayes, minimax, or other) is a function of all past observations on the endogenous variable as well as of all past controls, and thus ultimately is a function of the random disturbance term. Thus the control rule itself is a random variable and whether it is a good control can be judged by considering the distribution of the values which it will assume at any point in time.

The situation is similar to problems in the theory of estimation where the sampling distribution of an estimate is investigated. In that theory an estimate is considered good if its sampling distribution is concentrated, in some sense, about the true parameter which is being estimated. The usual measure of concentration of unbiased estimates is the variance (corresponding to mean square error). An estimate which has the smallest variance out of a class of unbiased estimates is said to be efficient. In problems where the exact sampling distribution is difficult or impossible to determine and where there is a large sample, one might be able to find the asymptotic distribution of the estimate and then extend the concept of efficiency to asymptotic efficiency. The usual approach in this case is to call an estimate asymptotically efficient, out of a class of estimates which are consistent and have an asymptotic normal distribution, if it has the smallest asymptotic normal variance. (Consistency is also used as a criterion of estimation. If an estimate is good, then it is consistent.)

Since asymptotic efficiency and consistency have proved successful as criteria of estimation,[1] and since control of a system with unknown parameters necessitates estimating those parameters, it seems worthwhile to study how these criteria of estimation might be used as criteria of control. Such a study is the aim of this chapter. However, the control problem has aspects which the estimation problem does not; for example, concern with performance over time rather than at one point in time. The next section discusses some of these differences with an aim at determining what it means for a control rule to be efficient.

After this preliminary discussion we then show, in Section 4.3, in what sense the three controls in the model with known intercept are asymptotically efficient. We show that, out of a fairly wide class, the least squares control rule and the Bayesian control rule lead to

---

[1]See L. J. Savage (1954) for a discussion of the decision theoretic justification of consistency and asymptotic efficiency.

estimates which are asymptotically efficient. We also address ourselves to the question of whether there is anything to gain from experimenting in early stages. In the process of considering this, we show that a rule especially designed for experimentation cannot do any better than the more simple least squares certainty equivalence or Bayesian certainty equivalence rules. This result is surprising in that the Bayesian analysis suggested that the longer the time horizon the more there is to gain by experimenting in early periods, and our asymptotic theory really relates to an infinite time horizon.

This study of criteria is done entirely in the model with known intercept as the asymptotic distribution in the more complicated case of unknown intercept is not known. In Section 4.5 we briefly consider the implications if the asymptotic distribution conjectured in Section 3.4 is the true one.

4.2 Estimation versus Control

In this section, except for a few passing comments, we take as given the rationale for using asymptotic efficiency as a criterion of estimation. Therefore, to evaluate the concept of efficiency in the control problem, we need only discuss the difference between estimation and control.

One essential difference between the control problem with unknown parameters and the pure estimation problem is that in the former we are concerned with performance over many periods while in the latter we are concerned with obtaining a single estimate. This difference is captured by the corresponding loss functions: in the control problem we wish to minimize $(\beta u_t - a)^2$ in all periods and thus might write the loss function as $\Sigma_{t=1}^{T}(\beta u_t - a)^2$, whereas in the estimation problem we find a $\hat{\beta}$ to minimize $(\hat{\beta} - \beta)^2$, say. But whatever the form of the loss function we would like to keep x_t close to the desired level for all t . It seems that the concept of asymptotic efficiency is readily applicable to this type of situation. In fact, when we say that an estimate (or a control in this case), is asymptotically efficient, we really mean that the sequence is asymptotically efficient.[2] That is, we are referring to a property which holds for all t (sufficiently large). More precisely, if $\sqrt{t}(\beta u_t^{(1)} - a)$ has a limiting distribution $N(0, \sigma_{(1)}^2)$, while $\sqrt{t}(\beta u_t^{(2)} - a)$ has a limiting distribution $N(0, \sigma_{(2)}^2)$ with $\sigma_{(2)}^2 > \sigma_{(1)}^2$, then for every $\delta > 0$

$$(4.1) \qquad P[|\beta u_t^{(1)} - a| < \delta/\sqrt{t}] > P[|\beta u_t^{(2)} - a| < \delta/\sqrt{t}]$$

for <u>all</u> t sufficiently large. Since we are concerned with properties of the sequence and not just elements of the sequence, it seems that the concept of asymptotic efficiency is quite useful in the control problem. To emphasize this point, in estimation we make an <u>estimate</u>, not a <u>sequence of estimates</u>, so that we really do not care about what characterizes a good sequence. In control we do care about what characterizes

---

[2]The "ellipsis" of referring to a sequence of estimates as an estimate and the possible consequence is discussed in L. J. Savage (1954), p. 227.

a good sequence. Since asymptotic efficiency refers to a characteristic of a sequence, we conclude that the concept has applicability to the control problem.

Although our main concern in this section is efficiency the above comments also apply to consistency, a concept which refers to a sequence. In the introduction of Chapter III we argued that the concept of strong consistency is more useful for the control problem than that of weak consistency, since it makes a joint probability statement about all elements of the sequence. Note that in this context, inequality (4.1) is not such a joint probability concept, so that a stronger statement could be made, although we do not do so.

A second difference between control and estimation which is related to the above is the idea of a time horizon in the control problem. In applying asymptotic theory in estimation, we associate the results for $t \to \infty$ with large t, although how large t needs to be is usually uncertain. To apply asymptotic theory to the control problem the analogous association is between an infinite horizon and some finite horizon. But, as with estimation, we are uncertain as to how large the time horizon must be before the theory becomes applicable. This difficulty is inherent in any asymptotic theory and is no more a problem in control than in pure estimation.

On the other hand, in many control situations, there are good arguments for formulating the problem in terms of an infinite horizon. If there is no natural stopping point in the foreseeable future, then an infinite horizon may be assumed. For example, in the problem of stabilizing the rate of inflation in an economy, it is difficult to conceive of a terminal point, when either the economy will end or when we will no longer care about inflation. In such problems with an unknown terminal date, the concept of asymptotic efficiency is even more appealing than in estimation problems.

One objection to this idea, even in infinite horizon problems, is that a discount rate should be included in the analysis. This is necessary in some formulations where the loss function might otherwise be infinite. However, in problems like controlling inflation a discount rate does not seem appropriate, so the objection is not valid. We discuss this point further in Section 4.4 where we introduce a loss function for a finite horizon problem without a discount rate.

A third difference between the control problem and pure estimation is the question of experimental design. As discussed in Chapter II, controls in earlier periods might be used to obtain information for use in later periods. If the control problem has an infinite horizon, this aspect should reveal itself in the efficiency of the control, because the asymptotic normal variance reflects the amount of information that is obtained by the control rule. That is, if experimentation gives more information, this should show up in the variance. However, if the control problem is of short duration the benefits of experimentation might not show up in the asymptotic variance.

The last point is a difficulty which occurs in any use of the asymptotic normal variance as a measure of efficiency. In some situations one might be more interested in the limit of the variance, which may be quite different than the asymptotic normal variance. Hodges and Lehmann (1956) argue that the variance of the asymptotic distribution has relevance in situations where one is concerned more with the frequency of small deviations from the true value than with large but improbable deviations. Further, if one were concerned more with the latter, then we would question the usefulness of squared error loss when the deviations are large. In addition for some problems the limit of the variance may not exist, in which case the asymptotic normal variance would be better criterion.

4.3 Asymptotically Efficient Controls

Having discussed the usefulness of asymptotic efficiency as a criterion of control we now prove the following theorem which is a formal statement of how the control rules defined and studied in this paper are asymptotically efficient.

THEOREM 4: In the model $x_t = \beta u_t + \epsilon_t$, under the assumptions of THEOREM 1, let $\{u_t\}$ be any control rule which converges to a/β with probability 1. Then the limiting distribution of $\sqrt{t}(\hat{\beta}_t - \beta)$ and of $\sqrt{t}(b_t - \beta)$ is $N(0, (\beta^2/a^2)\sigma^2)$.

PROOF: In the proof of THEOREM 2 the only property of the least squares certainty equivalence rule and the Bayesian certainty equivalence rule which we use is convergence to the true value a/β with probability 1. This is enough to show that the first term in brackets in equation (4.4) converges in probability to zero and that $\Sigma_{i=1}^{t} u_i^2/t \xrightarrow{p} a^2/\beta^2$. Since by assumption any control rule in the class defined in this theorem has this convergence property, we obtain the same results about the limiting distributions of $\sqrt{t}(\hat{\beta}_t - \beta)$ and $\sqrt{t}(b_t - \beta)$.

The importance of this theorem is that that least squares certainty equivalence control and the Bayesian certainty equivalence control lead to parameter estimates which have as small an asymptotic variance as any other control rule in the class of rules having the property of convergence to the true value with probability 1. This class includes controls designed especially for experimentation as long as the control converges with probability 1 to a/β . The implication is that asymptotically there is nothing to gain by experimenting with controls to obtain more information about parameter estimates. In the long run as much information can be obtained by the more easily calculated control rules of this paper.

To emphasize this point consider a control rule which is designed for experimentation, being larger in absolute value than the certainty equivalence rules. That is, suppose

$$u_{t+1} = \frac{a}{\hat{\beta}_t}(1 + f(t)) , \tag{4.2}$$

where $f(t) = O(1/t)$ and is greater than zero. Then it can be easily shown that this control rule has the same asymptotic distribution as the least squares certainty equivalence rule.

The result of this section may be surprising. It says that as long as the deviation from the present period control rule for experimentation purposes is finite, then any benefit from that experimentation washes away in the limit. In other words, any additional information which comes from using a larger absolute value for the control becomes negligible in comparison to the information which is accumulated without the extra experimentation. Mathematically, this result is expressed by the fact that $\Sigma_{i=1}^{t} u_i^2/t \to (a/\beta)^2$ with probability 1, if $u_t \to a/\beta$ with probability 1 regardless of the values which the elements of the sequence $\{u_t\}$ assume.

4.4 A Criterion for Control Performance

In the last section we showed how the least squares certainty equivalence and the Bayesian certainty equivalence control rules lead to parameter estimates which are asymptotically efficient. While it certainly seems that efficient parameter estimates are necessary for good control performance, it is not clear what criterion of control performance we are implicitly assuming. To clarify this point and relate the analysis of this chapter to the decision theoretic framework, in this section we specify a loss function which is appropriate to many infinite horizon control problems. We then show that, under certain assumptions, the two certainty equivalence control rules minimize this loss function.

The criterion[3] which we propose is

---

[3]The expectation is with respect to the sequence of random variables $\{\epsilon_t\}$; we do not assume that β is a random variable as in the Bayesian approach.

$$(4.3) \qquad \lim_{T \to \infty} \frac{\sum_{t=1}^{T} [E(x_t - a)^2 - \sigma^2]}{\log T}$$

which we discussed briefly in Section 2.2. The rationale behind this criterion is that in many stabilization problems there is no natural terminal date and no reason to discount the future. The loss in each time period is defined as the difference between expected squared error loss and the minimum obtainable expected squared error loss (if β were known then $u_t = a/\beta$ and $E\epsilon_t^2 = \sigma^2$) . Thus we are actually using a regret in each period. In order to avoid the problem of unbounded loss without a discount rate we normalize the sum of these losses in each period by a suitable function of time. Since the regret in each period converges to zero, dividing by T would lead to a loss function identically equal to zero, but dividing by $\log T$ leads to a nonzero but finite quantity.

The main objection to this criterion is that there is no discounting of the future. However, a careful perusal of the literature on control theory in economics, indicates that discount rates are frequently used for the mathematical convenience of a finite loss rather than for economic reasons. In his original study, Ramsey (1928) goes so far as to state that,

> ... it is assumed that we do not discount later enjoyments in comparison with earlier ones, a practice which is ethically indefensible and arises merely from the weakness of our imagination.

Certainly in many problems a discount rate is unwarranted and the criterion of (4.3) seems to be a viable alternative in such cases.

We now show that the certainty equivalence rules minimize this criterion out of a certain class of rules. The results of Sections 4.2 and 4.3 use the fact that the controls suitably normalized have a limiting distribution which is normal. One way to relate these distributions to the limit of the expected loss is to assume that the control rules are bounded; then we can use the fact that for bounded random variables the variance of the limiting distribution is equal to the limit of the normalized variance. For the remainder of this section we therefore assume that the control rules are bounded.

Under this boundedness assumption we have from the asymptotic distribution (3.50) that, with the least squares certainty equivalence or the Bayesian certainty equivalence control rules,

(4.4)
$$\lim_{t \to \infty} t[E(x_t - a)^2 - \sigma^2]$$

$$= \lim_{t \to \infty} tE(\beta u_t - a)^2$$

$$= \sigma^2 \quad ,$$

which implies[4] that

---

[4]The following lemma is used. [See Knopp (1956), p. 32.] Let p_t be a sequence of positive numbers with $\Sigma_{t=1}^{T} p_t$ diverging to infinity and suppose that a sequence $z_t \to z$, then $\Sigma_{t=1}^{T} p_t z_t / \Sigma_{t=1}^{T} p_t \to z$.

$$\text{(4.5)} \qquad \frac{\sum_{t=1}^{T} [E(x_t - a)^2 - \sigma^2]}{\log T}$$

$$= \frac{\sum_{t=1}^{T} \frac{1}{t}}{\log T} \; \frac{\sum_{t=1}^{T} [E(x_t - a)^2 - \sigma^2]}{\sum_{t=1}^{T} \frac{1}{t}}$$

$$\rightarrow \sigma^2 \; .$$

Now let us consider any other bounded control rule which is a linear function of either the least squares certainty equivalence or the Bayesian certainty equivalence rules and which converge to a/β . By an argument similar to that at the end of Section 4.3 we can show that with any such rule, $\sqrt{t}(\beta u_t - a)$ has a normal limiting distribution with variance at least as large as σ^2 . Thus for this class of control rules we have

$$\text{(4.6)} \qquad \min \lim_{T \to \infty} \frac{\sum_{t=1}^{T} [E(x_t - a)^2 - \sigma^2]}{\log T} = \sigma^2 \; .$$

Both the least squares certainty equivalence and the Bayesian certainty equivalence rules attain this minimum.

4.5 The Case of Unknown Intercept: Discussion

Since the results of Chapter III do not include an asymptotic distribution in the model where both the slope and intercept are unknown, we cannot make any judgments on control policy in terms of efficiency. However, if the conjecture of equation (3.60) is correct then we have a rather strong result. Since the asymptotic variance which results from (3.60) is σ^2/β_2^2 , the same as when we know the intercept, then the controls in this model must also be efficient. That is, we obtain as much information about the control value in a model with two unknown parameters as in the model with one unknown parameter. This is especially interesting in that it has generally been thought that experimentation could be quite beneficial in this model.

As a final indication of how experimenting is not beneficial (if the result of (3.60) is true), we consider the asymptotic variance of the estimate for $(a - \hat{\beta}_{1t})/\hat{\beta}_{2t}$ when controls are used purely for experiments; that is, when u_t is non-stochastic and does not converge. Therefore, $\Sigma_{i=1}^{t}(u_i - \bar{u}_t)^2/t$ does not converge to zero. In that case, the asymptotic variance can be computed easily from the fixed regressor case; it is

$$\lim_{t \to \infty} \frac{\sigma^2}{\beta_2^2} \frac{\beta_2^2 t \sum_{i=1}^{t} u_i^2 + 2\beta_1\beta_2 t \sum_{i=1}^{t} u_i + t^2\beta_1^2}{\beta_2^2 t \sum_{i=1}^{t} (u_i - \bar{u}_t)^2}$$

which can be shown to be greater than σ^2/β_2^2 for any sequence $\{u_t\}$ which does not converge.

The preliminary indication of these few remarks is that we obtain most information about a point on a regression line not by obtaining the most information about the slope and intercept but by trying values near that point. However, a rigorous statement of these results is not possible until the asymptotic distribution has been mathematically derived for this model.

4.6 Conclusion

This chapter has used the asymptotic normality results derived in Chapter III as criteria of control in the model with known intercept. Unlike the Bayesian methods, which have been used to suggest various control rules, these methods do not depend on any particular specification of the prior distribution. In this sense the results are more general than the Bayesian analysis which previous studies have used. On the other hand, the results are asymptotic and are therefore more relevant to problems with long time horizons.

The important results of the investigation are that (1) the least squares certainty equivalence, and the Bayesian certainty equivalence control rules lead to efficient parameter estimates and (2) the benefits of using a control, which obtains extra information by experimenting, become negligible asymptotically. Since the calculation of optimal rules which allow for experimentation has proven so difficult, or impossible in more complicated cases, the suggestion of this analysis is that the computationally easier rules be used instead, especially in problems with a long time horizon.

Which of the two easier rules to use cannot be examined by this analysis as they are equally efficient. However, since the Bayesian certainty equivalence rule can incorporate data which may have accumulated prior to the formulation of the problem, it would be preferred in such situations.

CHAPTER V

THEORIES OF FIRM PRICING UNDER UNKNOWN DEMAND

5.1 Introduction

In addition to normative implications for economic policy, the convergence results derived in Chapter III have some implications for economic theory. In this chapter we illustrate these implications by considering a simple model of firm pricing behavior. In doing so we will also compare the control problem of this study with a related sequential statistical decision model, the two-armed bandit problem [Robbins (1952), Bellman (1956)].

Conventional theories of imperfect competition assume that firms have complete knowledge of the parameters in their demand function. This assumption is also made in recent research on firms facing uncertain demand, in which a random disturbance term is added to an otherwise deterministic demand function [see, for example, Leland (1972)]. These theories do not consider how firms obtain knowledge about the parameters in their demand function, and have no explanation of how firms adjust price on the basis of new information. The models discussed in this study suggest a way to include such considerations in the theory of the firm.

5.2 A Structural Model of Firm Pricing

Consider a profit maximizing firm facing a market demand curve for a single product. We assume that the firm knows the structure of demand, namely that there is a linear relation between the price charged and the quantity purchased, but that the slope of this linear relation is unknown. Further it is assumed that observed contradictions to this demand structure will be thought of as shocks or unexplainable errors in the relationship between price and quantity. Under such assumption the demand curve can be represented by

$$q_t = \alpha + \beta p_t + \varepsilon_t , \quad t = 1, 2, \ldots, \tag{5.1}$$

where p_t is price charged, q_t is quantity sold, ε_t is a random disturbance term, β is the unknown slope of the demand curve, and α is a known intercept. The firm can obtain information about β over time by observing the quantity sold at various prices.

We assume that the firm has only fixed costs so that profits in any time period can be represented as

$$\Pi_t = p_t q_t - C \qquad t = 1, 2, \ldots, \tag{5.2}$$

where C is a constant. The objective of the firm is to maximize profits over time. This objective might be represented by the sum of discounted profits or average profits over time, but we will not be concerned with specifying the exact criterion. We are more interested in how the existence of an unknown parameter affects equilibrium and price adjustment.

With unknown demand an appropriate definition of firm equilibrium becomes difficult. Since new information will change the firm's beliefs about the unknown parameters it is likely that it will also cause a change in price. Thus the information available to the firm must be included in the definition of an equilibrium. Hayek (1937) noted this problem in the definition of equilibrium:

> Any change in the relevant knowledge of a person disrupts the equilibrium between his actions.

One definition of equilibrium which has these characteristics is: an equilibrium price is the price the firm would charge with perfect information about the unknown parameters. For the profit maximizing firm in this discussion there is only one such equilibrium price, namely $p = -\alpha/2\beta$. Note that this does not mean that once this price is charged there is no incentive to change price. This lack of incentive to change price occurs only when the firm has acquired enough information.

Having defined this equilibrium it is of interest to know whether it is stable and whether there are any price adjustment schemes which will converge to this equilibrium price. One possible price adjustment scheme is given by the least squares certainty equivalence control rule; namely

$$(5.3) \qquad p_{t+1} = -\frac{\alpha}{2\hat{\beta}_t} \quad ,$$

where $\hat{\beta}_t$ is the least squares estimate of β. This is the price that the firm would charge if it maximized profits by treating β as known with certainty and equal to the least squares estimate. Such a pricing scheme generates a sequence of prices over time. The use of least squares estimates may be questionable as a description of behavior, but it seems like the most harmless assumption which incorporates learning on the basis of observation on quantity and price. An alternative price adjustment scheme might be

$$(5.4) \qquad p_{t+1} = -\frac{\alpha}{2b_t} ,$$

where b_t is the Bayes estimate of β.

A slight modification of the arguments of Chapter III can be used to show that both these price adjustment mechanisms converge with probability one to the equilibrium price. That is, the equilibrium is stable.

With these price adjustment mechanisms, price is a function of all past prices and quantities sold. In disequilibrium prices are changed as the firm learns about the true value of the unknown parameter. Such a behavioral description of price dynamics seems more realistic than ad hoc mechanisms in which today's price is a function of prices in the last few periods only.

5.3 A Two-Armed Bandit Model of Firm Pricing

Rothschild (1971) has considered a different model to describe firm behavior under unknown demand. Say that there are two (this can be generalized) possible prices, p_1 and p_2 which the firm may charge for its product, and that the probabilities that a consumer will buy the product at each price are Π_1 and Π_2. The unknown demand for this firm is represented by the fact that Π_1 and Π_2 are not known to the firm. If the firm wishes to maximize profits, then the above definition of equilibrium implies that the price which leads to highest expected profits is the equilibrium price.

This problem is exactly equivalent to the two-armed bandit problem where a gambler is faced with the prospect of playing two slot machines whose probability of pay-off is unknown. Rothschild extends an argument of Bellman (1956) to show that, if the firm follows an optimal pricing policy, then with positive probability, it will charge the price with lower expected profits infinitely often. That is, there is not convergence with probability 1 to the equilibrium price. Rothschild then shows that this result gives an endogenous reason for price variability in equilibrium.

The reason for the lack of convergence is that, with positive probability, a bad string of luck will lead the firm to believe that the wrong price should be charged rather than the true price. Once this wrong price is used, it is not possible to obtain information about the other price. Thus the wrong price might be charged forever.

5.4 A Comparison of the Two Models

The two models introduced above to describe firm behavior under unknown demand have led to entirely different conclusions. In the structural model the price adjustment mechanism converges with probability one to the true equilibrium price while in the two-armed bandit model the price adjustment mechanism has a positive probability of converging to the wrong price. The main reason for this difference seems to be that in the model with known structure the firm can obtain information about the entire demand relation at any price it charges, while in the two-armed bandit model the firm obtains information about demand only at the one price that is being charged. Without a structural relation between price and quantity it is possible for the two-armed bandit firm to remain ignorant about demand at prices other than the ones that are charged. Thus the assumption that the firm has some knowledge of the structure seems crucial to get the convergence result. Further research might be concerned with "how much structure" is sufficient to lead to convergence with probability one.

Since the results of the two models are different it is interesting to ask which is the better representation of firm behavior. It seems that most firms have some notion as to how their price affects the quantity purchased (e.g., downward sloping demand), and that observed contradictions to this notion will be called shocks or unexplainable errors. In such situations a structural model seems more appropriate. On the other hand, the two-armed bandit approach would be more useful in situations where there is little or no structure (e.g., technological innovations). In any case, it seems that the two approaches complement each other in suggesting ways to handle the problem of unknown parameters in the theory of economic behavior.

CHAPTER VI

SUMMARY AND CONCLUSIONS

This study has investigated the problem of controlling an endogenous variable in a linear model over time when the parameters of the model are unknown and must be estimated. Such a model is representative of a large variety of economic models both at the macroeconomic level and the microeconomic level, so that the results should suggest other applications than those discussed in this paper.

In summary, we have proved that two different control rules converge with probability 1 to their true value. This was done in a model with one unknown parameter, where we demonstrated that enough information would be accumulated by each control rule to consistently estimate the single unknown parameter. Because the structure relating control values at different points in time is quite complicated, we had to develop methods of proof which are general enough to avoid using this structure. It is hoped that these methods might be useful in other econometric investigations where the regressors are random variables with a complicated structure. In addition to these convergence results we also proved that the control rules were asymptotically normal. The normalization of this distribution is $\sqrt{t}$, so that the variance of the control about its true value decreases as $1/t$.

Since asymptotic efficiency has proved to be a useful criterion in the theory of estimation we suggested applying this method in the control problem. We discussed the differences between control and estimation in order to show the advantages and disadvantages of such an approach. THEOREM 4 was then used to show that the three simple control rules are asymptotically efficient. Further, we demonstrated how the value of experimentation is negligible in the long run by showing that a control designed for experimentation has the same asymptotic variance as a nonexperimenting control. A performance criterion of control was specified to clarify the assumptions behind these conclusions. This criterion is most appropriate in problems with a long time horizon and no discounting.

Finally in the last chapter we showed how these results have implications for the theory of economic behavior under uncertainty. The model provides a description of such behavior which is a useful alternative to other approaches.

The unifying conclusion from this analysis is that the simple "quick and easy" control rules perform quite well relative to those rules which are designed for experimentation.

APPENDIX

PROCEDURES FOR APPROXIMATING BAYES OPTIMAL CONTROL RULES

As was pointed out in Chapter II there is no exact analytic solution for the Bayes optimal control rule, even in the model with only one unknown parameter, for any time horizon greater than one time period. Therefore, it is necessary to develop approximation procedures if this rule is to be employed. There are several types of approximations which might be used. One is a numerical approximation where particular values of the parameters are chosen, and numerical integration and minimization of numerical functions are used to solve the problem. Prescott (1972) used such a method and calculated the Bayes optimal control rule for various prior distributions and various time horizons. The results of that analysis further confirm the intuitive suggestions of Section 2.3. However, in models of any more generality than (2.4) such procedures become quite difficult to employ in practice because of the large number of numerical integrations and minimizations which must be performed.

Another approach is to approximate the future risk with a function which can be integrated analytically. For example, consider the two-period version of the problem introduced in Section 2.3.1. We can approximate the risk r_2^* as follows. If we set $z = b_1/\sigma_1$, then the expectation of r_2^* in (2.10) becomes

(A.1) $$G(u_1) = \int_{-\infty}^{+\infty} \frac{a^2}{z^2 + 1} \, d\Phi(z; \mu, \tau^2) ,$$

where $\Phi(z; \mu, \tau^2)$ is the normal distribution with mean

(A.2) $$\mu = b_0 \sqrt{\frac{1}{\sigma_0^2} + u_1^2}$$

and variance

(A.3) $$\tau^2 = \sigma_0^2 u_1^2 .$$

One approximation to $a^2/(1 + z^2)$ is a multiple of the normal density function with the height, location, and area under the two curves the same. Such an approximation is

(A.4) $$A(z) = a^2 e^{-z^2/\pi} ,$$

and can be integrated analytically with respect to the normal distribution Φ to obtain an approximation of the future risk as a function of u_1 . This can be added to the present risk in period one and minimized with respect to u_1 . The results of this method were compared with the numerical integration results; for the parameters selected the results were very similar. As our main purpose here is an exposition of possible techniques, we do not present a detailed study comparing the two methods.

Another technique is to investigate the properties of (A.1) directly and approximate it with a function which has the same properties. It can be shown that (A.1) has the following properties:

(A.5) $$G(0) = \frac{a^2\sigma_0^2}{\sigma_0^2 + b_0^2} \quad ,$$

(A.6) $$G'(0) = 0 \; ,$$

(A.7) $$G''(0) = -2a^2\sigma_0^4 \left[\frac{(b_0^2 - \sigma_0^2)^2}{(b_0^2 + \sigma_0^2)^2} \right] \; ,$$

and $G(u_1)$ is a decreasing positive function of $|u_1|$ which approaches zero as $|u_1|$ gets large. A relatively simple function which has all these properties is

(A.8) $$\tilde{G}(u_1) = \frac{a^2}{u_1^2 \left[\frac{b_0^2 - \sigma_0^2}{b_0^2 + \sigma_0^2} \right]^2 + \frac{b_0^2}{\sigma_0^2} + 1} \; .$$

This function added to the risk in period one can be minimized with respect to u_1 to obtain an approximation to the Bayes optimal control rule. Comparing this method with the first two approximations shows that this is a rougher approximation. Its advantage is that it can be readily generalized to more than one control and in addition gives a simple measure of how much future risk can be reduced by experimenting with u_1 .

This brief discussion of approximation procedures has aimed at demonstrating what can be done when the risk functions are impossible to calculate analytically. A more serious problem which frequently arises in these multiperiod control models is that for a given loss function the risk may not be finite. For example, suppose that (as in Chapter V)

(A.9) $$x_t = \alpha + \beta u_t + \epsilon_t \quad ,$$

represents the demand curve of a monopolist who wishes to maximize profits over a two period time horizon, where α is known with certainty and β has the usual prior distribution. Then the loss function is $L = -\Sigma_{t=1}^{2} x_t u_t$. The optimal price in the last period is $u_2 = -\alpha/2b_2$ with corresponding risk $r_2^* = \alpha^2/4b_2$. When calculating the optimal price in the first period, we must take the expected value of r_2^* . But since b_2 is normally distributed this expectation does not exist.

Such difficulties are usually avoided in theoretical analyses by assuming from the start that all expectations are finite. However, if these problems are to be used in practice one must have alternative procedures. The main approach of this study, which is developed in Chapters III and IV, does not rely on calculating expected loss in future periods and, therefore, does not have this difficulty.

REFERENCES

1. Anderson, T. W. (1959), On asymptotic distributions of estimates of parameters of stochastic difference equations, Ann. Math. Stat., 30, 676-687.

2. Aoki, Masanao (1967), Optimization of Stochastic Systems, Academic Press Inc., New York.

3. Arrow, K. J., D. Blackwell, M. A. Girshick (1949), Bayes and minimax solutions of sequential decision problems, Econometrica, 17, 213-244.

4. Basu, A. (1973), Economic regulation under parameter uncertainty. Unpublished Ph.D. dissertation, Economics Department, Stanford University.

5. Bellman, R. (1956), A problem in the sequential design of experiments, Sankhya, 16, 221-229.

6. Brainard, W. (1967), Uncertainty and the effectiveness of policy, American Economic Review, 57, 411-425.

7. Cyert, R. M. and M. H. DeGroot (1970), Bayesian analysis and duopoly theory, Journal of Political Economy, 78, 1168-1184.

8. Dvoretzky, A., J. Kiefer, and J. Wolfowitz (1952), The inventory problem II. Case of unknown distributions of demand, Econometrica, 20, 451-466.

9. Fel'dbaum, A. A. (1960), Dual control theory I, Automatic and Remote Control, 21, 874-880.

10. Feller, W. (1966), An Introduction to Probability Theory and Its Applications, Volume II (Second Edition), John Wiley and Sons, Inc., New York.

11. Fisher, W. (1962), Estimation in the linear decision model, International Economic Review, 3, 1-29.

12. Hayek, F. A. (1937), Economics and knowledge, Economica, 4, 33-54.

13. Hodges, J. L. and E. L. Lehmann (1956), Two approximations to the Robbins-Monroe process, Proc. Third Berkeley Symposium on Math. Stat. and Prob., 1, 95-104.

14. Knopp, K. (1956), Infinite Sequences and Series, Dover Publications, Inc., New York.

15. Krickeberg, K. (1965), Probability Theory, Addison Wesley Publishing Company, Inc., Redding, Massachusetts.

16. Leland, H. (1972), The theory of the firm facing uncertain demand, American Economic Review, 62, 278-291.

17. Prescott, E. C. (1967), Adaptive decision rules for macroeconomic planning. Unpublished Ph.D. dissertation, GSIA, Carnegie-Mellon University.

18. Prescott, E. C. (1972), The multiperiod control problem under uncertainty, Econometrica, forthcoming.

19. Raiffa, H. and Schlaifer, R. (1961), Applied Statistical Decision Theory, The M.I.T. Press, Cambridge.

20. Ramsey, F. (1928), A mathematical theory of saving, Economic Journal, 38, 543-559.

21. Rao, C. R. (1965), Linear Statistical Inference and Its Applications, John Wiley and Sons, Inc., New York.

22. Robbins, H. (1952), Some aspects of the sequential design of experiments, Bulletin of American Mathematical Society, 58, 527-535.

23. Rothschild, M. (1971), A two-armed bandit theory of market pricing, Working Paper No. 10, M.S.S.B. Workshop, University of California, Berkeley, 1971.

24. Savage, L. J. (1954), The Foundations of Statistics, John Wiley and Sons, Inc., New York.

25. Simon, H. A. (1956), Dynamic programming under uncertainty with a quadratic criterion function, Econometrica, 24, 74-81.

26. Theil, H. (1957), A note on certainty equivalence in dynamic programming, Econometrica, 25, 346-349.

27. Wald, A. (1947), Sequential Analysis, John Wiley and Sons, Inc., New York.

28. Wetherill, G. B. (1966), Sequential Methods in Statistics, Methuen and Co., Ltd., London.

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