

Federal Reserve Bank of Minneapolis
Research Department Working Paper

ON THE ACCURACY OF LINEAR QUADRATIC
APPROXIMATIONS: AN EXAMPLE*

Lawrence J. Christiano

Federal Reserve Bank of Minneapolis

Working Paper 303

March 1986

*I have benefited from conversations with Ian Bain, John Bryant, David Card, Ron Dye, Ed Prescott, Tom Sargent and Mike Stutzer.

The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. The material contained is of a preliminary nature, is circulated to stimulate discussion, and is not to be quoted without permission of the author.

Abstract

This paper investigates--in the context of a simple example--the accuracy of an econometric technique recently proposed by Kydland and Prescott. We consider a hypothetical econometrician who has a large sample of data, which is known to be generated as a solution to an infinite horizon, stochastic optimization problem. The form of the optimization problem is known to the econometrician. However, the values of some of the parameters need to be estimated. The optimization problem--presented in a recent paper by Long and Plosser--is not linear quadratic. Nevertheless, its closed form solution is known, although not to the hypothetical econometrician of this paper. The econometrician uses Kydland and Prescott's method to estimate the unknown structural parameters. Kydland and Prescott's approach involves replacing the given stochastic optimization problem by another which approximates it. The approximate problem is a element of the class of linear quadratic problems, whose solution is well-known--even to the hypothetical econometrician of this paper. After examining the probability limits of the econometrician's estimators under "reasonable" specifications of model parameters, we conclude that the Kydland and Prescott method works well in the example considered. It is left to future research to determine the extent to which the results obtained for the example in this paper applies to a broader class of models.

1. Introduction and Summary

Much progress has been made in developing tools for econometrically estimating the structural parameters of agents' dynamic decision rules when it is assumed that these represent solutions to dynamic, stochastic optimum problems. Examples are the work of Chow (1979), Hansen and Sargent (1980), Kennan (1980), Sargent (1978), and Long and Plosser (1982). With the exception of Long and Plosser, advances in estimation technique have come at the cost of assuming quadratic optimum problems and linear transition laws (LQ problems).¹

This paper represents a first step towards evaluating the quality of an estimation method which we refer to as the "LQ approximation method." The method can be used when the analyst assumes a stochastic optimization problem that is not LQ and for which calculating the exact optimal decision rule is too costly or intractable. The method involves replacing the assumed non-quadratic problem with another--referred to here as the "LQ approximation problem"--which is LQ and which can be solved easily.

The LQ approximation method is the following. It is assumed that the problem can, by appropriate substitution, be converted into an infinite horizon stochastic calculus of variations problem with a return function which is not necessarily quadratic.² The random variables of the problem are replaced by their expected value and the steady state value of the state vector of the resulting nonstochastic optimum problem is then calculated. (From here on, by "steady state values" we mean the unconditional expectation of the stochastic term and the steady state of the state variables just described.) Next, the second order Taylor series expansion of the return function about the steady state values of the variables is computed. The result is a stochastic LQ

problem which has a closed form solution that is straightforward to calculate. The closed form solution to the LQ problem, together with a statistical model for the uncontrollable stochastic terms, approximates the exact reduced form of the model. The approximate reduced form may be combined with the available sample data to compute a likelihood function, which is to be maximized with respect to the unknown structural parameters of the problem. Applications of the LQ approximation method may be found in Kydland and Prescott (1980, 1982). Related approximations are surveyed in Chow (1975) and Mariano (1978).

The intuition underlying the LQ method seems to be the following. The certainty version of a nonquadratic problem and the certainty version of its LQ approximation have the same steady states by construction.³ Because of this, a sufficiently small deviation ("shock") in initial conditions from the steady state generates similar return paths to the steady state. (We assume stability.) The expectation underlying the LQ method is that in the uncertainty case, if the variances on the random shocks are small, then the dynamic behavior about the unconditional mean of the two systems will also be roughly similar.⁴ The latter, one expects, will lead to estimators having good properties. We formalize this idea below and show in Proposition 1 that this expectation is fulfilled in the example considered in the paper.

We consider the accuracy of the LQ approximation method in the context of a real model of the business cycle. The model is a scalar version of a model presented and studied extensively in Long and Plosser (1982). Long and Plosser's model was chosen for this study for two reasons. First, although the model is stochastic and nonquadratic, the exact closed form solution is known. Second, the model satisfies sufficient conditions for the solution to the LQ approximation problem to be the first order Taylor's series expansion about the

exact solution evaluated at the steady state (Theorem 1). As a result, the algebraic form of the mapping from the structural parameters of the model to the parameters of the stochastic difference equation ("reduced form parameters") that solves the LQ approximation problem is clarified and shown to be quite simple. This, together with the fact that the model is just identified, enables us to determine the associated inverse mapping. The latter is what is required to calculate the probability limit of the econometrician's estimators. The measure of approximation error that is used is the ratio of the probability limit of the econometrician's estimator of a parameter to its true value. There are just three free parameters in the model.

What we find is that, for the example considered, the error of approximation is negligible for reasonable values of the error variances.

The plan of the paper is as follows. In section 2 the LQ approximation method is described precisely. In section 3 the version of Long and Plosser's model that is studied is presented. In section 4 the probability limit of the econometrician's estimator is calculated, given the model of section 3. In section 5 numerical measures of approximation error under alternative parameterizations are tabulated and discussed. Conclusions appear in section 6.

2. Preliminaries

This section begins by defining a linear quadratic approximation to a given optimum problem. The approximate problem is referred to as the "LQ approximation problem." A result is then presented which provides conditions under which the first order Taylor series expansion about the stationary feedback rule that solves the given problem solves the LQ approximation problem. This result greatly simplifies the calculations in section 3. The "LQ approximation method"--the econometric method that is the object of study in this paper--is defined at the end of this section.

2.1 The Exact Problem

Consider the problem of choosing a sequence of contingent plans for setting $\{z_{t+j}\}_{j=0}^{\infty}$ to maximize

$$(2.1) \quad E \left[\sum_{j=0}^{\infty} \beta^j u(z_{t+j}, z_{t+j-1}, v_{t+j}) \mid v_t, z_{t-1} \right], \quad 0 < \beta < 1,$$

subject to $(z_{t+j}, z_{t+j-1}, v_{t+j}) \in T$ for all $j \geq 0$ and z_{t-1} and v_t fixed. Here, z_t and v_t are vectors of dimension n and m , respectively, with v_t being a vector of random variables, and $E(\cdot)$ denotes the mathematical expectation operator. The return function, $u(\cdot, \cdot, \cdot)$, is strictly concave in its first two arguments. The set $T \subset \mathbb{R}^{(2n+m)}$ is convex in its first two dimensions. By this we mean that $(z'_1, z_1, v_1), (z'_2, z_2, v_1) \in T$ implies $(\lambda z'_1 + (1-\lambda)z'_2, \lambda z_1 + (1-\lambda)z_2, \lambda_1) \in T$, where $0 \leq \lambda \leq 1$.

We assume that $\{v_t\}$ is covariance stationary and has the following representation

$$(2.2) \quad v_t = \rho v_{t-1} + \mu + \varepsilon_t, \quad E\varepsilon_t = 0, \quad E\varepsilon_t \varepsilon_{t-\tau}^T = \begin{cases} V_\varepsilon & \tau = 0 \\ 0 & \tau \neq 0 \end{cases}$$

where ρ is a square matrix with eigenvalues inside the unit circle, μ is a vector of constants and V_ε is positive semi definite. The process $\{\varepsilon_t\}$ is serially independent and independent of $\{v_{t-s}, s \geq 1\}$. Since we require only that $\{v_t\}$ be a vector process of finite dimension, the setup in (2.1) can accommodate random variables with vector autoregressive representations of arbitrary, finite, order (see Sargent [1979, pp. 272-73] or Chow [1975, pp. 49-50]).

Suppose that the sequence of feasible contingent plans that maximize (2.1) exists, is stationary and unique and can be written in the following form:

$$(2.3) \quad z_{t+j} = f(z_{t+j-1}, v_{t+j}, V_\varepsilon), \quad j = 0, 1, \dots$$

Let ε^{t+i} , $i \geq 1$, denote the vector random variable $\{\varepsilon_{t+1}, \varepsilon_{t+2}, \dots, \varepsilon_{t+i}\}$. Note from (2.2) that $v_{t+i} = v_{t+i}(\varepsilon^{t+i}, v_t)$. By recursive substitution, (2.3) can be used to express (z_{t+j}, z_{t+j-1}) , $j \geq 0$ as a function of ε^{t+j} , v_t , z_{t-1} and v_ε . Write this

$$(2.3a) \quad (z_{t+j}, z_{t+j-1}) = Z(\varepsilon^{t+j}, z_{t-1}, v_t, v_\varepsilon, j) \quad j \geq 0,$$

with $Z(\varepsilon^{t+j}, z_{t-1}, v_t, v_\varepsilon, j) \equiv (f(z_{t-1}, v_t, v_\varepsilon), z_{t-1})$ when $j = 0$.

Substituting the above expression into (2.1), we obtain $v(z_{t-1}, v_t, v_\varepsilon)$, the value of the optimal plan,

$$(2.4) \quad v(z_{t-1}, v_t, v_\varepsilon) = E \left[\sum_{j=0}^{\infty} \beta^j u(Z(\varepsilon^{t+j}, z_{t-1}, v_t, v_\varepsilon, j), v_{t+j}(\varepsilon^{t+j}, v_t)) \mid v_t, z_{t-1} \right],$$

where the expectation is taken relative to the distribution of the random vector $\{\varepsilon^{t+j}\}_{j=1}^{\infty}$.

2.2 The LQ Approximation Problem

Problem (2.1) may be difficult to solve when the return function, $u(\cdot, \cdot, \cdot)$, is not quadratic, as we assume. In applied work one may choose to solve the following LQ approximation problem instead. Take the limit as $N \rightarrow \infty$ of the following sequence of problems. Maximize over plans of the form $z_{t+j} = L_j^{(N)}(z_{t+j-1}, v_{t+j})$, where $L_j^{(N)}$ are linear functions, $j = 0, 1, \dots, N$, the expression

$$(2.5) \quad E \left[\sum_{j=0}^N \beta^j U(z_{t+j}, z_{t+j-1}, v_{t+j}) \mid z_{t-1}, v_t \right].$$

In (2.5), $U(\cdot, \cdot, \cdot)$ is a second order Taylor series expansion of $u(\cdot, \cdot, \cdot)$

about the points $z_t = z_s, z_{t-1} = z_s, v_t = v_s$, and is defined more precisely below. Let $\hat{L}_j^{(N)}(z_{t+j-1}, v_{t+j})$, $j = 0, 1, \dots, N$ denote the solution to the N period problem, $N < \infty$. In Lemma 6 (proved in Appendix A) it is shown that the concavity assumption on the return function guarantees that a solution exists and that

$$(2.5a) \quad \lim_{N \rightarrow \infty} \hat{L}_j^{(N)} = \hat{L}_j$$

for all j . We refer to the stationary plan $(\hat{L}(z_{t-1}, v_t), \hat{L}(z_t, v_{t+1}), \dots)$ as the solution to the LQ approximation problem.

The value of the LQ approximation problem is defined as follows. Let $J^{(N)}(z_{t-1}, v_t)$ be the value of (2.5) when the plan $z_{t+j} = \hat{L}_j^{(N)}(z_{t+j-1}, v_{t+j})$, $j = 0, 1, \dots, N$ is executed. Then

$$J(z_{t-1}, v_t) = \lim_{N \rightarrow \infty} J^{(N)}(z_{t-1}, v_t),$$

is the value of the LQ approximation problem. In Lemma 6 in Appendix A it is shown that the concavity assumption on the return function in (2.1) guarantees the existence of J .

Calculating the solution to the LQ approximation problem (e.g., the functions \hat{L} and J) is straightforward. For example, in Lemma 6 in Appendix A it is shown how the LQ approximation problem can be written as a linear regulator problem. The problem of solving a linear regulator problem has been studied extensively and is well understood. See, for example, Bertsekas (1976), Chow (1975), Hansen and Sargent (1981), Kushner (1971), Kwakernaak and Sivan (1972) and Sargent (1980).

In (2.5),

$$\begin{aligned}
 (2.6) \quad U(z_{t+j}, z_{t+j-1}, v_{t+j}) &= u(z_s, z_s, v_s) \\
 &+ u_1(z_s, z_s, v_s)(z_{t+j} - z_s) + u_2(z_s, z_s, v_s)(z_{t+j-1} - z_s) \\
 &+ u_3(z_s, z_s, v_s)(v_{t+j} - v_s) \\
 &+ \frac{1}{2}(z_{t+j} - z_s)^T u_{11}(z_s, z_s, v_s)(z_{t+j} - z_s) \\
 &+ \frac{1}{2}(z_{t+j-1} - z_s)^T u_{22}(z_s, z_s, v_s)(z_{t+j-1} - z_s) \\
 &+ \frac{1}{2}(v_{t+j} - v_s)^T u_{33}(z_s, z_s, v_s)(v_{t+j} - v_s) \\
 &+ (z_{t+j} - z_s)^T u_{12}(z_s, z_s, v_s)(z_{t+j-1} - z_s) \\
 &+ (z_{t+j} - z_s)^T u_{13}(z_s, z_s, v_s)(v_{t+j} - v_s) \\
 &+ (z_{t+j-1} - z_s)^T u_{23}(z_s, z_s, v_s)(v_{t+j} - v_s).
 \end{aligned}$$

In (2.6), $v_s = E v_t$, and z_s is the steady state value (we assume it exists, is independent of z_{t-1} and v_s , and that $u(z_s, z_s, v_s)$ is well defined) of the following problem:

$$(2.7) \quad v(z_t, v_s, 0) = \max_{\text{feasible } \{z_{t+j}\}_{j=0}^{\infty}} \sum_{j=0}^{\infty} \beta^j u(z_{t+j}, z_{t+j-1}, v_s),$$

which is deterministic. Frequently z_s can be computed for (2.7) even when it is not known how to calculate the values of $\{z_{t+j}\}_{j=0}^{\infty}$ that solve (2.7).

When z_s cannot be calculated from (2.7), then the LQ approximation method we consider breaks down.

In (2.6),

$$u_i(z_s, z_s, v_s) \equiv \frac{\partial u(x_1, x_2, x_3)}{\partial x_i^T} \left| \begin{array}{l} x_1 = z_s \\ x_2 = z_s \\ x_3 = v_s \end{array} \right.$$

$$u_{ij}(z_s, z_s, v_s) \equiv \frac{\partial^2 u(x_1, x_2, x_3)}{\partial x_i \partial x_j^T} \left| \begin{array}{l} x_1 = z_s \\ x_2 = z_s \\ x_3 = v_s \end{array} \right.$$

We use the following convention with respect to differentiation. When y is a vector of length n and x a vector of length m , then

$$\frac{\partial y}{\partial x^T} \equiv \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \dots & \frac{\partial y_1}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \dots & \frac{\partial y_n}{\partial x_m} \end{bmatrix}$$

When $n = 1$, then

$$\frac{\partial^2 y}{\partial x \partial x^T} \equiv \begin{bmatrix} \frac{\partial^2 y}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 y}{\partial x_1 \partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 y}{\partial x_m \partial x_1} & \dots & \frac{\partial^2 y}{\partial x_m \partial x_m} \end{bmatrix}$$

2.3 A Theorem

Suppose that the function defined in (2.3) is differentiable at least once in each argument. Define

$$(2.8) \quad g(z_{t-1}, v_t) = f(z_s, v_s, V_\varepsilon) + f_1(z_s, v_s, V_\varepsilon)(z_{t-1} - z_s) \\ + f_2(z_s, v_s, V_\varepsilon)(v_t - v_s).$$

Also, define

$$(2.9) \quad V(z_{t-1}, v_t) = v(z_s, v_s, V_\varepsilon) + v_1(z_s, v_s, V_\varepsilon)(z_{t-1} - z_s) \\ + v_2(z_s, v_s, V_\varepsilon)(v_t - v_s) \\ + 1/2(z_{t-1} - z_s)^T v_{11}(z_s, v_s, V_\varepsilon)(z_{t-1} - z_s) \\ + (z_{t-1} - z_s)^T v_{12}(z_s, v_s, V_\varepsilon)(v_t - v_s) \\ + 1/2(v_t - v_s)^T v_{22}(z_s, v_s, V_\varepsilon)(v_t - v_s).$$

Evidently, the functions $U(\cdot, \cdot, \cdot)$, $g(\cdot, \cdot)$ and $V(\cdot, \cdot)$ are functions of z_s and v_s . In addition, the latter two functions are functions of V_ε . The dependence is not made explicit in order to simplify the notation and in the expectation that this will not cause confusion.

We are in a position to state the following theorem, the proof of which is given in Appendix A.

Theorem 1 If,

- (i) the solution to (2.1) exists, is unique, is of the form given in (2.3) and $(f(z_{t-1}, v_t, V_\varepsilon), z_{t-1})$ is interior to the feasible set T ,
- (ii) the stationary state, z_s , of (2.7) is finite, independent of z_{t-1} , the initial condition, and $u(z_s, z_s, v_s)$ is well defined,
- (iii) the value function $v(\cdot, \cdot, \cdot)$ is differentiable at least twice in the first two arguments,

- (iv) the function $f(\cdot, \cdot, \cdot)$ given in (2.3) has at least one derivative in the first two arguments,
- (v) $v(z_{t-1}, v_t, V_\varepsilon) = \tilde{v}(z_{t-1}, v_t) + \ell(V_\varepsilon)$, where $\tilde{v}(z_{t-1}, v_t)$ is functionally independent of the elements of V_ε ,
- (vi) the return function $u(\cdot, \cdot, \cdot)$ in (2.1) is strictly concave in the first two arguments,
- (vii) the set T is convex in its first two dimensions.

Then,

- (viii) $g = \hat{L}$, where g is defined in (2.8), and the stationary plan $(\hat{L}(z_{t-1}, v_t), \hat{L}(z_t, v_{t+1}), \dots)$ is the solution to the LQ approximation problem. (\hat{L} is defined in (2.5a).)

Summary of Proof

First we introduce some new notation. Let V be the set of quadratic functions in z_{t-1} and v_t negative semidefinite in z_{t-1} . Define the operator $U: V \rightarrow V$ as follows:

$$(2.9) \quad U(q)(z_{t-1}, v_t) = \sup_{\substack{z_t \in \mathbb{R}^n \\ z_t \in R}} \{U(z_t, z_{t-1}, v_t) + \beta \text{Eq}(z_t, \rho v_t + \mu + \varepsilon_{t+1})\},$$

where $q \in V$. The space V is closed under the U operator because of

- (vi). Define the function \tilde{V} as follows:

$$(2.10) \quad \tilde{V}(z_{t-1}, v_t) = V(z_{t-1}, v_t) + \frac{1}{2} \left(\frac{\beta}{1-\beta} \right) \text{tr}[v_{22}(z_s, v_s, V_\varepsilon) V_\varepsilon] + \ell(Q) - \ell(V_\varepsilon),$$

where $V(\cdot, \cdot)$ is defined in (2.9). Following is an outline of the proof to Theorem 1.

- (a) The function \tilde{V} in (2.10) is shown to be a fixed point of U in V , i.e., $U(\tilde{V}) = \tilde{V}$ and $\tilde{V} \notin V$.

- (b) It is shown that the function $g(\cdot, \cdot)$ defined in (2.8) achieves the sup in (2.9) for $q = \tilde{V}$. That is,

$$\begin{aligned} & U(g(z_{t-1}, v_t), z_{t-1}, v_t) + \text{BEV}(g(z_{t-1}, v_t), \rho v_t + \mu + \varepsilon_{t+1}) \\ & > U(z_t, z_{t-1}, v_t) + \text{BEV}(z_t, \rho v_t + \mu + \varepsilon_{t+1}), \end{aligned}$$

for all $z_t \in \mathbb{R}^n$.

- (c) The optimal value for problem (2.5) is a function $J \in \mathbf{V}$. Moreover, J is the unique element in \mathbf{V} such that $U(J) = J$. Consequently,

$$(2.12) \quad J = \tilde{V}.$$

- (d) The optimal linear feedback law for (2.5) exists and is a stationary plan $(\hat{L}(z_{t-1}, v_t), \hat{L}(z_t, v_{t+1}), \dots)$. The linear function \hat{L} is the unique function for which the sup in (2.9) is achieved for $q = J$. That is,

$$\begin{aligned} & U(\hat{L}(z_{t-1}, v_t), z_{t-1}, v_t) + \text{BEJ}(\hat{L}(z_{t-1}, v_t), \rho v_t + \mu + \varepsilon_{t+1}) \\ & > U(z_t, z_{t-1}, v_t) + \text{BEJ}(z_t, \rho v_t + \mu + \varepsilon_{t+1}) \end{aligned}$$

for all $z_t \in \mathbb{R}^n$. Taking (2.11) and (2.12) into account, it follows that $g = \hat{L}$.

Proof: (See Appendix A.)

Following is a brief discussion of the assumptions of the theorem. Condition (i) simplifies the proof by guaranteeing that first order necessary conditions for (2.3) to solve (2.1) are met as a strict equality for $j = 0$.

To illustrate the effect of (ii), it excludes the following problem, which satisfies all the other conditions of Theorem 1.

$$\begin{aligned}
& v(z_{t-1}, v_t, v_\varepsilon) \\
(2.13) \quad & = \sup E \left\{ \sum_{j=0}^{\infty} \beta^j \ln[\exp(v_{t+j})z_{t+j} - z_{t+j}^2/z_{t+j-1}] \mid z_{t-1}, v_t \right\},
\end{aligned}$$

where the sup is evaluated over contingent plans for setting $\{z_{t+j}\}_{j=0}^{\infty}$ subject to $\{z_{t+j}, z_{t+j-1}, v_{t+j}\} \in T$ for $j \geq 0$. The set T is characterized by the condition $0 < z_{t+j} < \exp(\frac{1}{2} v_{t+j})z_{t+j-1}$ for all $j \geq 0$. Also, $v_t = \mu + \varepsilon_t$ and ε_t is independently distributed with mean zero and variance v_ε . (The v_t process is a scalar version of (2.1) and (2.2) with $\rho = 0$.) It is shown in Christiano and Prescott (1982) that

$$(2.14) \quad v(z_{t-1}, v_t, v_\varepsilon) = c + \left(\frac{1}{1-\beta}\right) \ln z_{t-1} + \left[\frac{2-\beta}{2(1-\beta)}\right] v_t.$$

where

$$c = \left(\frac{1}{1-\beta}\right) \left\{ \ln 2 - \left[1 + \frac{\beta}{2(1-\beta)}\right] \ln\left(2 + \frac{\beta}{1-\beta}\right) + \frac{\beta(2-\beta)}{2(1-\beta)} \mu + \frac{\beta}{2(1-\beta)} \ln\left(\frac{\beta}{1-\beta}\right) \right\}.$$

In Christiano and Prescott (1982) it is shown that the following stationary feedback rule has value (2.14):

$$(2.15) \quad z_t = \left(\frac{\beta}{2-\beta}\right)^{1/2} \exp\left(\frac{1}{2} v_t\right) z_{t-1}.$$

In the certainty case, with $v_\varepsilon = 0$ and v_t replaced by its unconditional expectation, $v_s (= \mu)$, the solution to (2.13) is

$$(2.16) \quad z_t = \left(\frac{\beta}{2-\beta}\right)^{1/2} \exp\left(\frac{1}{2} v_s\right) z_{t-1}.$$

From (2.16), we see that if $v_s = \ln[(2-\beta)/\beta]$, then $z_s = z_{t-1}$, the initial condition, while $z_s = \infty$ for $v_s > \ln[(2-\beta)/\beta]$. When $v_s < \ln[(2-\beta)/\beta]$, then $z_s = 0$, a point at which the return function in (2.13) is not defined.

Conditions (iii) and (iv) are necessary if the Taylor series expansions in (2.9) and (2.8), respectively, are to be well defined.

Benveniste and Scheinkman (1979) present conditions under which $v(\cdot, \cdot, \cdot)$ has one derivative in z_{t-1} in the nonstochastic case with $v_t = v_s$ for all t . They do not consider the stronger condition (iii). This condition is necessary to compute the expression $\tilde{V}(\cdot, \cdot)$ in (2.10).

Condition (v) guarantees that the problem, (2.1), satisfies certainty equivalence. (See Lemma 2 in Appendix A.) This is because under (v) the feedback rule, $f(\cdot, \cdot, \cdot)$, in (2.3) is functionally unrelated to the elements of V_ε . In particular, setting $V_\varepsilon = 0$ amounts to applying the $E(\cdot)$ operator in (2.1) directly to v_{t+j} , $j > 0$. Consequently, the $f(\cdot, \cdot, V_\varepsilon)$ function that solves (2.1) may be found by solving the certainty problem that results when v_{t+j} is replaced by its conditional expectation for all $j > 0$. This is the principle of certainty equivalence discussed in, for example, Holt, et al. (1960).

The fact that certainty equivalence obtains in (2.13) (and also in the problem in Section 3) establishes that a quadratic return function is not a necessary condition for certainty equivalence to hold. Certainty equivalence is used extensively to establish the following two key facts:

$$V(z_{t-1}, v_t) = U(g(z_{t-1}, v_t), z_{t-1}, v_t) + \beta E\tilde{V}(g(z_{t-1}, v_t), \rho v_t + \mu + \varepsilon_{t+1}) \text{ and}$$

$$U_1(g(z_{t-1}, v_t), z_{t-1}, v_t) + \beta E\tilde{V}_1(g(z_{t-1}, v_t), \rho v_t + \mu + \varepsilon_{t+1}) = 0.$$

Condition (vi) is used to establish all the results described in the summary of the proof to Theorem 1 concerning the solution to the LQ approximation problem (see (c) and (d)). Conditions (vi) and (vii) are used to establish the concavity of the function v in its first argument. The latter is what guarantees that \tilde{V} is an element of the set V .

A feature of Theorem 1 which limits its general applicability is that it makes assumptions on the value function, (2.4), and on the solution, (2.3). It is precisely the difficulty of obtaining these that leads one to solve the LQ approximation problem in the first place. For this reason it would be preferable to state all the assumptions of Theorem 1 in terms of the implied restrictions on the return function, u , the feasibility set, T , and the parameters of (2.2), if these were known. In the meantime, for the purposes of this paper, the theorem as it is stated is entirely adequate. In Section 3 we apply the theorem to a problem for which the f and v functions in (2.3) and (2.4), respectively, are known. The advantage of the theorem is that it permits us to solve the LQ approximation problem posed in Section 3 by taking a simple first order Taylor series expansion. This calculation is transparent analytically, in contrast with a direct solution method such as the recursive one used to solve the linear regulator problem. The theorem makes it possible to write an analytic formula for the large sample approximation error implicit in the LQ approximation method applied to a particular problem. The formula is presented in Section 4.

2.3 The LQ Approximation Method Defined

The "LQ approximation method" is the following. The structural parameters to be estimated--call these β --are the parameters of $u(\cdot, \cdot, \cdot)$ and of (2.12). Solve the LQ approximation problem (2.4). Because the problem is quadratic, the solution is a linear time series representation for $\{z_t\}$. (If the conditions of Theorem 1 are met, then the solution is given by (2.8).) Denote the parameters of this reduced form representation by Γ . The previous discussion implies that $\Gamma = h(\beta)$ for some function h .

The parameters Γ can be estimated using standard techniques of time series analysis (see, e.g., Granger and Newbold (1977) and Nerlove, Grether and

Carvalho (1979)). Call the estimates of Γ , $\hat{\Gamma}$. Then, assuming parameter identification, the implied estimate of β , call it $\hat{\beta}$, is $\hat{\beta} = h^{-1}(\hat{\Gamma})$. (Global parameter identification is equivalent with $h^{-1}(\cdot)$ being a function.) We use the phrase "LQ approximation method" to denote this method of estimating the structural parameters, β . Unless $u(\cdot, \cdot, \cdot) = U(\cdot, \cdot, \cdot)$, the method is an example of specification error. This is because the true functional form of the reduced form representation of $\{z_t\}$ will not be the same as the one implied by the LQ approximation problem.

3. The Example

The example we consider--a scalar version of an optimum problem studied in Long and Plosser (1982)--satisfies the conditions of Theorem 1. After presenting the model, the $f(\cdot, \cdot, \cdot)$, and $q(\cdot, \cdot)$ functions (see (2.3) and (2.8), respectively) implied by the example are derived. In the second part of this section the LQ approximation method, as applied by a hypothetical econometrician to estimate some of the parameters of the Long and Plosser model, is described in detail. The specification error implicit in the LQ approximation method for the example is made explicit in the third part of this section. Finally, three expressions are presented which we shall use to characterize the econometrician's loss from using the LQ approximation method in the example.

3.1 Long and Plosser's Model

We assume that the time series of decision variables in a hypothetical economy is generated as a solution to a particular optimum problem. The economy is populated by a single representative agent ("Robinson Crusoe") who maximizes utility subject to a given production technology. The production technology is Cobb-Douglas with a multiplicative shock term. The shock process is not controllable by the agent.

The problem is to maximize over contingency plans for setting $\{L_{t+j}, c_{t+j}\}_{j=0}^{\infty}$

$$(3.1) \quad E_t \sum_{j=0}^{\infty} \beta^j \{ \ln(H - L_{t+j}) + \theta \ln c_{t+j} \} \quad 0 < \beta < 1 ,$$

subject to

$$(3.2a) \quad c_{t+j} + x_{t+j} \leq Y_{t+j}, \quad c_{t+j}, x_{t+j} > 0, \quad 0 \leq L_{t+j} \leq H$$

$$(3.2b) \quad Y_{t+j} = \lambda_{t+j} L_{t+j}^{(1-\alpha)} x_{t+j}^{\alpha} \quad 0 < \alpha < 1$$

$$(3.2c) \quad L_{t-1}, x_{t-1}, \lambda_t \text{ given .}$$

Here,

$c_t \sim$ consumption at time t ,

$x_t \sim$ investment at time t ,

$Y_t \sim$ output at time t ,

$L_t \sim$ labor supplied at time t , in hours,

$H \sim$ total hours available at time t ,

$E_t(\cdot) \sim$ mathematical expectation operator conditioned on variables dated t and earlier,

$\lambda_t \sim$ stochastic term distributed lognormally,

$\theta, \alpha, \beta \sim$ parameters of the model.

According to (3.1), on a particular date, say t , the representative agent has at her/his disposal a known (because L_{t-1} , x_{t-1} and λ_t are known at t) quantity of output, Y_t . Also available is the fixed quantity of time H . The problem at time t is to choose an optimal sequence of plans for determining leisure ($H - L_{t+j}$) and consumption (c_{t+j}) for $j = 0, \dots$ subject to the constraints (3.2) and the limit on time available, H . The problem is one of decision making under uncertainty since values of λ_{t+j} , $j > 0$ are unknown at time t .

We assume that λ_t may be written

$$(3.3a) \quad \lambda_t = \exp(v_t) ,$$

where

$$(3.3b) \quad v_t = \rho v_{t-1} + \mu + \varepsilon_t ; \quad |\rho| < 1 , \quad \mu \text{ constant, } \varepsilon_t \sim \text{iid } N(0, \sigma_\varepsilon^2) .$$

The random variable $\varepsilon_t \times 100$ is unit free and represents perturbations, in percent terms, on Y_t . That is

$$(3.3c) \quad \varepsilon_t = \ln Y_t - E_{t-1} \ln Y_t .$$

To get (3.1) into the discrete time calculus of variations format of section 2, substitute (3.2) into (3.1) yielding the problem of maximizing over contingency plans for setting $\{L_{t+j}, x_{t+j}\}_{j=0}^{\infty}$,

$$(3.4) \quad E_t \sum_{j=0}^{\infty} \beta^j \{ \ln(H - L_{t+j}) + \theta \ln(\exp(v_{t+j}) L_{t+j}^{(1-\alpha)} x_{t+j}^\alpha - x_{t+j}) \} ,$$

subject to L_{t-1}, x_{t-1}, v_t given at time t . Let

$$z_t^T = (L_t, x_t)$$

$$(3.5) \quad u(z_t, z_{t-1}, v_t) = \ln(H - L_t) + \theta \ln(\exp(v_t) L_{t-1}^{(1-\alpha)} x_{t-1}^\alpha - x_t) .$$

Bellman's equation for (3.4) is, using the above notation,

$$(3.6) \quad v(z_{t-1}, v_t) = \sup_{z_t \in T} \{ u(z_t, z_{t-1}, v_t) + \beta E_t v(z_t, v_{t+1}) \} ,$$

where the constraint set, T , is characterized in (3.2). Long and Plosser (1982) discovered the solution to the functional equation (3.6), which is

$$(3.7) \quad v(z_{t-1}, v_t) = \left(\frac{\gamma}{1 - \beta\alpha}\right)v_t + (1 - \alpha)\gamma \ln L_{t-1} + \alpha\gamma \ln x_{t-1} + K$$

where

$$K = \left(\frac{1}{1 - \beta}\right) \left\{ \ln \left(\frac{H}{1 + \beta\gamma(1 - \alpha)} \right) + \theta \ln(1 - \beta\alpha) \right. \\ \left. + \beta\gamma(1 - \alpha) \ln \left(\frac{H\beta\gamma(1 - \alpha)}{1 + \beta\gamma(1 - \alpha)} \right) + \beta\gamma\alpha \ln \beta\alpha + \frac{\beta\gamma\mu}{1 - \beta\rho} \right\}$$

$$\gamma = \frac{\theta}{1 - \beta\alpha}.$$

In Christiano and Prescott (1982) it is shown that the function v in (3.6) is the optimal value of (3.4). That (3.7) does indeed solve (3.6) is verified by simple substitution. A necessary condition that z_t^* be an element of the optimal program in (3.4) is, assuming an interior solution, that

$$(3.8) \quad u_1(z_t^*, z_{t-1}, v_t) + \beta E_t v_1(z_t^*, v_{t+1}) = 0,$$

which implies,

$$(3.9a) \quad z_t^* = \begin{bmatrix} L_t^* \\ x_t^* \end{bmatrix} = \begin{bmatrix} \left(\frac{H\beta\gamma(1 - \alpha)}{1 + \beta\gamma(1 - \alpha)} \right) \\ \beta\alpha \exp(v_t) L_{t-1}^{(1-\alpha)} x_{t-1}^\alpha \end{bmatrix}$$

or,

$$(3.9b) \quad z_t^* = f(z_{t-1}, v_t, \sigma_\varepsilon^2).$$

It is straightforward to verify that this example satisfies the conditions of Theorem 1 in section 2.3. It follows that the solution to the LQ approximation problem for this example (see (2.8)) is

$$z_t = g(z_{t-1}, v_t),$$

where,

$$(3.10) \quad g(z_{t-1}, v_t) = \begin{bmatrix} \frac{H\beta\gamma(1 - \alpha)}{1 + \beta\gamma(1 - \alpha)} \\ \alpha x_{t-1} + (1 - \alpha) \left(\frac{x_s}{L_s} \right) L_{t-1} + x_s (v_t - v_s) \end{bmatrix}.$$

Here,

$$(3.10a) \quad B = \beta\alpha \left(\frac{HBY(1-\alpha)}{1 + \beta\gamma(1-\alpha)} \right)^{1-\alpha}$$

$$(3.10b) \quad L_s = \left[\frac{B}{\beta\alpha} \right]^{1/1-\alpha}$$

$$(3.10c) \quad x_s = \left[\exp\left(\frac{\mu}{1-\rho}\right) B \right]^{1/1-\alpha}$$

$$(3.10d) \quad \gamma = \frac{\theta}{1 - \beta\alpha}.$$

3.2 The Hypothetical Experiment

The hypothetical experiment we consider is the following. The econometrician knows the true form of the structural econometric model given by (3.1), (3.2), and (3.3). Some of the structural parameters of the model are unknown and must be estimated using available data.

We assume

- AI. the econometrician only observes $\{x_t\}$, and not $\{L_t\}$, $\{Y_t\}$, or $\{v_t\}$.

The structural parameters are α , β , θ , ρ , μ , and σ_ε^2 . We also assume that

- AII. the econometrician knows the true values of ρ , θ , and β .

It will be shown below that AII guarantees that the remaining parameters, α , μ , and σ_ε^2 , are identified given observations on $\{x_t\}$.

The econometrician is assumed to be unaware that the exact solution to (3.4) is given by (3.9). Consequently, the econometrician is led to employ the LQ approximation method--defined in Section 2.3--to estimate the unknown parameters, α , μ and σ_ε^2 . A description of the method, as applied to the present example, follows.

Denote the econometrician's estimators of α , μ , and σ_ε^2 by $\hat{\alpha}_T$, $\hat{\mu}_T$, and $\hat{\sigma}_{\varepsilon,T}^2$, respectively. The estimators $\hat{\alpha}_T$ and $\hat{\mu}_T$ are chosen to minimize a least squares criterion, S , to be described in detail below. The estimator $\hat{\sigma}_{\varepsilon,T}^2$ is a simple function of the minimized value of S . $T + 2$ is the number of observations in the sample. Write $x^T = \{x_1, \dots, x_T\}$. Then the available sample is $\{x_{-1}, x_0, x^T\}$. Write

$$S = S(\alpha, \mu, \rho, \theta, \beta; x_{-1}, x_0, x^T).$$

By definition,

$$(3.11) \quad \begin{aligned} & S(\hat{\alpha}_T, \hat{\mu}_T, \rho, \theta, \beta; x_{-1}, x_0, x^T) \\ & \leq S(\alpha, \mu, \rho, \theta, \beta; x_{-1}, x_0, x^T), \end{aligned}$$

for all α such that $|\alpha| < 1$, and for $\mu \in \mathbb{R}$.

Given the known values of $x_{-1}, x_0, \dots, x_T, \rho, \theta, \beta$, and for given values of α and μ , which we denote by α^r and μ^r , S may be computed in the following sequence of four steps.

Step 1. Calculate x_s^r and L_s^r , the steady state values of x_t and L_t for $\sigma_\varepsilon^2 = 0$ and $v_s^r = \mu^r / (1 - \rho)$, given the known values of ρ, θ , and β , and given α^r and μ^r . These may be found, for example, by solving the first order necessary conditions of the certainty version of (3.4) (i.e., setting $\sigma_\varepsilon^2 = 0$).⁵

Step 2. Calculate the second order Taylor series expansion of $u(\cdot, \cdot, \cdot)$ in (2.1) about $(L_t, L_{t-1}, x_{t-1}, v_t) = (L_s^r, L_s^r, x_s^r, x_s^r, v_s^r)$, to get the function $U(\cdot, \cdot, \cdot)$ given in (2.6).

Step 3. Find the \hat{L} function that solves the quadratic approximation problem (2.5). This may be done using the recursive methods of dynamic programming (see, e.g., Chow (1975) or Hansen and Sargent (1981)), or by using classical

optimization techniques (see, e.g., Hansen and Sargent (1980)). The \hat{L} function may be written

$$(3.12) \quad \begin{bmatrix} L_t \\ x_t \end{bmatrix} = \hat{L}(L_{t-1}, x_{t-1}, v_t) \\ = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} L_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} v_t + \begin{bmatrix} c_1 \\ c_2 \end{bmatrix},$$

say, where a_{ij} , d_i , c_i , $i, j=1,2$ are scalars.

It would appear, given the calculations just outlined, that closed form expressions relating the a_{ij} 's d_i 's and c_i 's to α^r , β , θ , ρ , and μ^r would either not exist or would be exceedingly complicated. However, because the conditions of Theorem 1 are satisfied, it follows that $g = \hat{L}$, where g is defined in (3.10).⁶ That is,

$$(3.13) \quad a_{11} = 0, \quad a_{12} = 0, \quad a_{21} = (1 - \alpha^r) \left(\frac{x_s^r}{L_s^r} \right), \quad a_{22} = \alpha^r \\ d_1 = 0, \quad d_2 = x_s^r, \quad c_1 = \frac{H\beta\gamma^r(1 - \alpha^r)}{1 + \beta\gamma^r(1 - \alpha^r)}, \quad c_2 = -x_s^r v_s^r.$$

The econometrician of our hypothetical experiment, who cannot make use of Theorem 1, will nevertheless notice relations (3.13) after executing steps 1-3 several times.

Step 4. The econometrician is now in a position to calculate S . In the example, labor supply is constant (see the first line of (3.10)) and unobserved (recall AI). The econometrician is assumed to substitute this constant for L_{t-1} in the second equation in (3.10). This yields the following reduced form equation:

$$x_t = (1 - \alpha^r)x_t^r + \alpha^r x_{t-1}^r + x_s^r (v_t - v_s^r) .$$

Multiplying through by $(1 - \rho L)$,

$$(3.14) \quad (1 - \rho L)x_t = (1 - \alpha^r)(1 - \rho)x_s^r + \alpha^r(1 - \rho L)x_{t-1}^r + w_t^r,$$

say, where $w_t^r = x_s^r \epsilon_t$. Then,

$$(3.15) \quad S(\alpha^r, \mu^r, \rho, \theta, \beta; x_{-1}, x_0, x^T) \\ = \sum_{t=1}^T [(1 - \rho L)x_t - (1 - \alpha^r)(1 - \rho)x_s^r - \alpha^r(1 - \rho L)x_{t-1}^r]^2 ,$$

and

$$(3.16') \quad \sigma_{\epsilon}^{2,r} = \frac{S(\alpha^r, \mu^r, \rho, \theta, \beta; x_{-1}, x_0, x^T)}{T(x_s^r)^2} .$$

Write

$$(3.16) \quad y_t = b^0 + b^1 y_{t-1} + w_t ,$$

where

$$(3.16a) \quad b^0 = (1 - \alpha)(1 - \rho)x_s$$

$$(3.16b) \quad b^1 = \alpha$$

$$(3.16c) \quad \sigma_w^2 = (x_x)^2 \sigma_{\epsilon}^2$$

$$(3.16d) \quad y_t = (1 - \rho L)x_t .$$

Let $\beta = (\mu, \alpha, \sigma_{\epsilon}^2)$ denote the vector of structural parameters that the econometrician wishes to estimate. Let $\Gamma = (b^0, b^1, \sigma_w^2)$. Equations (3.16a)-(3.16c) represent a function, call it $h(\cdot)$, where $\Gamma = h(\beta)$.

Equations (3.11), (3.15) and (3.16) indicate that in our example, the econometrician's estimators, $\hat{\alpha}_T$, $\hat{\mu}_T$, $\hat{\sigma}_{\varepsilon, T}^2$ are calculated as follows. First, the regression indicated in (3.16) is calculated, producing estimators of the reduced form coefficients b^0 , b^1 , and σ_{ω}^2 , denoted by \hat{b}_T^0 , \hat{b}_T^1 , $\hat{\sigma}_{\omega, T}^2$, respectively.

Then the following calculations are carried out.

$$(3.17a) \quad \hat{\alpha}_T = \hat{b}_T^1$$

$$(3.17b) \quad \hat{x}_{s, T} = \hat{b}_T^0 / (1 - \hat{\alpha}_T)(1 - \rho)$$

$$(3.17c) \quad \hat{\gamma}_T = \theta / (1 - \hat{\alpha}_T \beta)$$

$$(3.17d) \quad \hat{B}_T = \beta \hat{\alpha}_T \left(\frac{H \beta \hat{\gamma}_T (1 - \hat{\alpha}_T)}{1 + \beta \hat{\gamma}_T (1 - \hat{\alpha}_T)} \right) (1 - \hat{\alpha}_T)$$

$$(3.17e) \quad \hat{\sigma}_{\varepsilon, T}^2 = \hat{\sigma}_{\omega, T}^2 / (\hat{x}_{s, T})^2$$

$$(3.17f) \quad \hat{\mu}_T = \log \left[\hat{x}_{s, T}^{(1 - \hat{\alpha}_T)} / \hat{B}_T \right] (1 - \rho) .$$

Here, (3.17a), (3.17e), and (3.17f) are estimators of the structural parameters and (3.17b)-(3.17d) are intermediate calculations. Equations (3.17) are the inverse of the $h(\cdot)$ mapping defined above. In particular,

$$(3.18) \quad \hat{\beta}_T = h^{-1}(\hat{\Gamma}_T) ,$$

in obvious notation. As long as $\hat{\alpha}_T \neq 1$, the h^{-1} mapping is well defined. This establishes that α , μ , and σ_{ε}^2 are identifiable. Had we not permitted the econometrician to know the values of at least three parameters, as in AII, the remaining parameters would not be identifiable.

3.3 Characterizing the Specification Error in the Example

The example we consider permits an explicit characterization of the specification error involved in applying the LQ approximation method. Loosely, the specification error is one of estimating a linear regression in levels of variables rather than in the logs. It is shown that the consequence of the specification error is not likely to be severe for small values of σ_ε^2 , where σ_ε^2 is defined in (3.3b).

The true data generating mechanism in our example, is taking the logarithm of the second line in (3.9a)

$$\ln x_t = \ln[\beta \alpha L_s^{(1-\alpha)}] + \alpha \ln x_{t-1} + v_t,$$

where L_s is defined in the first line of (3.9a). Multiplying both sides by $1 - \rho$,

$$\begin{aligned} \ln x_t &= (1 - \rho) \left\{ \ln[\beta \alpha L_s^{(1-\alpha)}] + \frac{\mu}{1-\rho} \right\} + (\alpha + \rho) \ln x_{t-1} - \alpha \rho \ln x_{t-2} + \varepsilon_t \\ &= (1 - \rho) \left\{ \ln[\beta \alpha \exp\left(\frac{\mu}{1-\rho}\right) L_s^{(1-\alpha)}] \right\} + (\alpha + \rho) \ln x_{t-1} - \alpha \rho \ln x_{t-2} + \varepsilon_t. \end{aligned}$$

Finally, taking (3.10c) into account,

$$(3.19) \quad \ln x_t = (1 - \rho)(1 - \alpha) \ln x_s + (\alpha + \rho) \ln x_{t-1} - \alpha \rho \ln x_{t-2} + \varepsilon_t.$$

From the latter equation, it is evident that $\ln x_s = E \ln x_t$.

Suppose the econometrician wishes to estimate α , x_s , and σ_ε^2 . (If, as elsewhere in this paper, it is μ and not x_s that is of interest, then μ can be obtained from the known values of β , ρ , θ , H and the estimated values of x_s and α). If it were known that the true data generating mechanism were (3.19), then the econometrician could consistently estimate α , x_s , and σ_ε^2 by the following procedure.

First, estimate the coefficients $\tilde{\delta}_0$, $\tilde{\delta}_1$, $\tilde{\delta}_2$ in the following regression.

$$(3.20) \quad \ln x_t = \tilde{\delta}_0 + \tilde{\delta}_1 \ln x_{t-1} + \tilde{\delta}_2 \ln x_{t-2} + u_t$$

subject to $\rho^2 - \tilde{\delta}_1 \rho + \tilde{\delta}_2 = 0$. Call the estimates of $\tilde{\delta}_0$, $\tilde{\delta}_1$, $\tilde{\delta}_2$ and σ_u^2 , $\hat{\delta}_{0,T}$, $\hat{\delta}_{1,T}$, $\hat{\delta}_{2,T}$, $\hat{\sigma}_{u,T}^2$ respectively. Then, $\hat{\alpha}_T = -\hat{\delta}_{2,T}/\rho$, or $\hat{\alpha}_T = \hat{\delta}_{1,T} - \rho$, $\hat{\alpha}_{s,T} = \exp\{\hat{\delta}_{0,T}/(1 - \hat{\alpha}_T)(1 - \rho)\}$, and $\hat{\sigma}_{\epsilon,T}^2 = \hat{\sigma}_{u,T}^2$.

Next consider what our hypothetical econometrician, who uses the LQ approximation method, does. From (3.11) and (3.15) deduce that the econometrician estimates the coefficients δ_0 , δ_1 , δ_2 in the projection of x_t on a constant and x_{t-1} , x_{t-2} :

$$(3.21) \quad x_t = \delta_0 + \delta_1 x_{t-1} + \delta_2 x_{t-2} + w_t$$

subject to $\rho^2 - \delta_1 \rho + \delta_2 = 0$. Then, $\hat{\alpha}_T = -\hat{\delta}_{2,T}/\rho$, or $\hat{\alpha}_T = \hat{\delta}_{1,T} - \rho$, $\hat{\alpha}_{s,T} = \hat{\delta}_{0,T}/(1 - \hat{\alpha}_T)(1 - \rho)$, $\hat{\sigma}_{\epsilon,T}^2 = \hat{\sigma}_{w,T}^2/(\hat{\alpha}_{s,T})^2$.

It is clear that $\text{Plim } \hat{\alpha}_T = \alpha$. However, in general, we would expect $\text{Plim } \hat{\alpha}_T \neq \alpha$. Values of $\text{Plim } \hat{\alpha}_T$, $\text{Plim } \hat{\sigma}_{\epsilon,T}^2$ and $\text{Plim } \hat{\alpha}_{s,T}$ under alternative parameterizations of the model in section 3.1 are presented in section 5.

The specification error the econometrician commits by applying the LQ approximation method in this example can be described consisely as follows. The true data generating mechanism is (3.19), which is linear in the log of x_t . On the other hand, the econometrician proceeds as if the true model were (3.19) with the logs removed and with the error variance multiplied by x_s^2 .

When σ_ϵ^2 , the variance of ϵ_t in (3.19), is small, then the consequences of the econometrician's specification error will be small. This is because, for small enough values of σ_ϵ^2 , the covariogram of x_t generated by (3.19), $\sigma_x^2(\tau)$, and that generated by x_s are similar. We show this below by looking at the series expansion of $\sigma_x^2(\tau)$.

In Appendix B it is shown that the mean, μ_x , and covariogram, $\sigma_x^2(\tau)$, τ integer, of an $\{x_t\}$ process generated according to (3.19) is given by

$$(3.22) \quad \begin{aligned} \text{Ex}_t &\equiv \mu_x = x_s \exp \left[(1/2)g(\alpha, \rho, 0)\sigma_\varepsilon^2 \right] \\ &= x_s [1 + R(\alpha, \rho, \sigma_\varepsilon^2)] \end{aligned} \quad \frac{7/}{}$$

$$(3.23) \quad \begin{aligned} \sigma_x^2(\tau) &= (\mu_x)^2 \{ \exp [g(\alpha, \rho, \tau)\sigma_\varepsilon^2] - 1 \} \\ &= g(\alpha, \rho, \tau)x_s^2\sigma_\varepsilon^2 + F(\alpha, \rho, \tau, \sigma_\varepsilon^2), \end{aligned}$$

where,

$$\begin{aligned} g(\alpha, \rho, \tau) &= \frac{\rho |\tau|}{(1-\rho^2)(1-\alpha\rho)(\rho-\alpha)} + \frac{\alpha |\tau|}{(1-\alpha^2)(1-\alpha\rho)(\alpha-\rho)} \\ R(\alpha, \rho, \sigma_\varepsilon^2) &= (1/2)g(\alpha, \rho, 0)\sigma_\varepsilon^2 + \frac{1}{2!} [1/2g(\alpha, \rho, 0)\sigma_\varepsilon^2]^2 + \dots, \end{aligned}$$

and $F(\alpha, \rho, \tau, \sigma_\varepsilon^2)$ is a function such that $(F(\alpha, \rho, \tau, \sigma_\varepsilon^2)/\sigma_\varepsilon^2) \rightarrow 0$ as $\sigma_\varepsilon^2 \rightarrow 0$. Consequently,

$$\frac{\sigma_x^2(\tau)}{\sigma_\varepsilon^2} \xrightarrow{\sigma_\varepsilon^2 \rightarrow 0} g(\alpha, \rho, \tau)x_s^2,$$

so that

$$(3.24) \quad \sigma_x^2(\tau) \approx g(\alpha, \rho, \tau)x_s^2\sigma_\varepsilon^2, \text{ for small } \sigma_\varepsilon^2.$$

Also, from (3.22)

$$(3.25) \quad \mu_x \approx x_s, \text{ for small } \sigma_\varepsilon^2.$$

It can be verified that

$$(3.26) \quad g(\alpha, \rho, \tau) - (\alpha + \rho)g(\alpha, \rho, \tau-1) + \alpha\rho g(\alpha, \rho, \tau-2) = 0,$$

for $\tau = 1, 2, 3, \dots$.

The coefficients δ_1, δ_2 in the projection equation in (3.21) are uniquely defined by the equations

$$(3.27) \quad \sigma_x^2(\tau) = \delta_1 \sigma_x^2(\tau-1) + \delta_2 \sigma_x^2(\tau-2)$$

for $\tau = 1, 2$. Thus, when σ_ε^2 is small--so that $\sigma_x^2(\tau)$ is nearly proportional to $g(\alpha, \rho, \tau)$ by (3.24)--(3.26) and (3.27) indicate that $\delta_1 \approx (\alpha + \rho)$ and $\delta_2 \approx -\alpha\rho$.

The constant term, δ_0 , in (3.21) is defined by

$$(3.28) \quad \delta_0 = \mu_x(1 - \delta_1 - \delta_2).$$

By (3.25), $\mu_x \approx x_s$ where σ_ε^2 is small, so that in this case,

$$\delta_0 \approx x_s(1 - \alpha)(1 - \rho).$$

Finally, the variance of the projection error in (3.21) can be shown, for small σ_ε^2 to be approximately $(x_s)^2 \sigma_\varepsilon^2$. In addition, as a small aside, it is a simple exercise to show that the errors in the projection equation, (3.21), will be serially correlated, and that this correlation goes to zero as $\sigma_\varepsilon^2 \rightarrow 0$.

The preceding argument, according to which the consequences of the specification error implicit in the LQ approximation method are less severe for small values of σ_ε^2 , has been nonrigorous and informal. This is remedied in sections 3.4 and 4. In the next section we develop three criteria that specify a kind of loss function for the econometrician. This will provide a sense in which we can say whether or not the consequences of specification error are "very" severe. In section 4 we show rigorously that as $\sigma_\varepsilon^2 \rightarrow 0$, the econometrician's loss goes to zero. The result is summarized in Proposition 1 in section 4.

3.4 Measures of Approximation Error

Let $\hat{M}_T = (\hat{\mu}_{y,T}, \sigma_{y,T}^2(0), \sigma_{y,T}^2(1))$ denote the mean, variance, and lag-one covariance of a sample of observations on y_t : $\{y_0, y_1, \dots, y_T\}$ (y_t is defined in (3.16d)). The reduced form parameter estimates, $\hat{\Gamma}_T$, are related to \hat{M}_T as follows:

$$(3.29a) \quad \hat{b}_T^1 = (\sigma_{y,T}^2(1)/\sigma_{y,T}^2(0)) + o(\frac{1}{T})$$

$$(3.29b) \quad \hat{b}_T^0 = (1 - \hat{b}_T^1)\hat{\mu}_{y,T} + o(\frac{1}{T})$$

$$(3.29c) \quad \hat{\sigma}_{\omega,T}^2 = [1 + (\hat{b}_T^1)^2] \sigma_{y,T}^2(0) - 2\hat{b}_T^1 \sigma_{y,T}^2(1) + o(\frac{1}{T})$$

where $o(\frac{1}{T})$ is a term which vanishes in probability. Equations (3.29) represents a function $r(\cdot)$, such that

$$(3.30) \quad \hat{\Gamma}_T = r(\hat{M}_T) .$$

Combining (3.18) and (3.30), we get

$$(3.31) \quad \hat{\beta}_T = h^{-1}(r(\hat{M}_T)) = q(\hat{M}_T) ,$$

say. Ignoring the cases where $|\hat{b}_T^1| > 1$, the $q(\cdot)$ function is continuous. It follows (see, e.g., Theil [1971, p. 361]) that

$$(3.32) \quad \hat{\beta} \equiv \text{Plim } \hat{\beta}_T = q(\text{Plim } \hat{M}_T) .$$

From here on we adopt the convention that a variable with a hat, but without a subscript T denotes the probability limit of the same variable with a T subscript.

We take the loss to the econometrician of using the LQ approximation method to be measured by the following three objects:

$$(3.33) \quad \frac{\hat{\alpha}}{\alpha}, \frac{\hat{\sigma}_\varepsilon}{\sigma_\varepsilon}, \frac{\hat{\mu}}{\mu},$$

where $\hat{\sigma}_\varepsilon = \text{Plim } \hat{\sigma}_{\varepsilon, T}$, $\hat{\sigma}_{\varepsilon, T} = \sqrt{\hat{\sigma}_{\varepsilon, T}^2}$.

According to (3.21), in order to calculate (3.23), we require $\text{Plim } \hat{M}_T$. We obtain this in the next section. Numerical values of (3.23) under alternative parameterizations of (3.1) are presented in section 5.

4. The Probability Limit of the Approximation Error Formula

In order to calculate (3.33), (3.32) indicates that we first require $\text{Plim } \hat{M}_T \equiv \hat{M} = (\hat{\mu}_y, \hat{\sigma}_y^2(0), \hat{\sigma}_y^2(1))$. We calculate these in this section. We also examine the behavior of $\hat{\beta} = (\hat{\alpha}, \hat{\mu}, \hat{\sigma}_\varepsilon^2)$ as $\sigma_\varepsilon^2 \rightarrow 0$. Results are summarized in Proposition 1 below.

Let $\mu_y, \sigma_y^2(\tau)$, τ integer, be the true mean and lagged- τ covariance of the y_t process. Then, since $y_t = (1 - \rho L)x_t$,

$$(4.1) \quad \mu_y = (1 - \rho)\mu_x$$

$$(4.2) \quad \sigma_y^2(\tau) = (1 + \rho^2)\sigma_x^2(\tau) - \rho\sigma_x^2(\tau + 1) - \rho\sigma_x^2(\tau - 1).$$

We assume that $\{x_t\}$ is generated by (3.9a). For convenience we reproduce (3.9a) here:

$$(4.3) \quad x_t = \lambda_t B x_{t-1}^\alpha, \quad B = \beta \alpha \left(\frac{H\beta\gamma(1-\alpha)}{1 + \beta\gamma(1-\alpha)} \right)^{(1-\alpha)}, \quad \gamma = \frac{\theta}{1 - \beta\alpha}, \quad 0 < \alpha < 1$$

where

$$(4.4) \quad \lambda_t = \lambda_{t-1}^\rho \exp(\mu + \varepsilon_t) \quad \varepsilon_t \sim \text{iidN}(0, \sigma_\varepsilon^2), \quad |\rho| < 1.$$

In Appendix B it is shown that,

$$(4.5) \quad \mu_x = \left[B \exp\left(\frac{\mu}{1-\rho}\right) \right]^{\frac{1}{1-\alpha}} \exp\left\{\frac{\sigma_\varepsilon^2}{2}\right\} \cdot \frac{1 + \rho\alpha}{(1-\rho^2)(1-\alpha\rho)(1-\alpha^2)}$$

and,

$$(4.6) \quad \sigma_x^2(\tau) = (\mu_x)^2 \left\{ \exp\left[\frac{\sigma_\varepsilon^2 \rho}{(1-\rho^2)(1-\alpha\rho)(\rho-\alpha)}\right]^{\rho|\tau|} \exp\left[\frac{\sigma_\varepsilon^2 \alpha}{(1-\alpha^2)(1-\alpha\rho)(\alpha-\rho)}\right]^{\alpha|\tau|} - 1 \right\},$$

for integer values of τ . Substituting (4.5) and (4.6) into (4.1)-(4.2), we obtain the functions relating μ_y and $\sigma_y^2(\tau)$ to the parameters of the model: $\alpha, \beta, \theta, \sigma_\varepsilon^2, \mu$ and ρ .

From (3.9a) we see that $\{\ln x_t\}$ is generated by a stationary first order stochastic difference equation. Consequently (see Anderson (1971, pp. 193-98)) $\{\ln x_t\}$ is ergodic in the mean and covariances. Since $\{x_t\}$ is the result of a one-to-one, differentiable transform on $\{\ln x_t\}$, $\{x_t\}$ has the same ergodic properties as $\{\ln x_t\}$.^{8/} It follows that

$$(4.6a) \quad \text{Plim } \hat{\mu}_{y,T} = \hat{\mu}_y = \mu_y$$

$$(4.6b) \quad \text{Plim } \hat{\sigma}_{y,T}^2 = \hat{\sigma}_y^2 = \sigma_y^2.$$

Taking the probability limit of (3.29) and substituting from (4.6),

$$(4.7a) \quad \hat{b}^1 = \text{Plim } \hat{b}_T^1 = \sigma_y^2(1)/\sigma_y^2(0)$$

$$(4.7b) \quad \hat{b}^0 = \text{Plim } \hat{b}_T^0 = \left(1 - \frac{\sigma_y^2(1)}{\sigma_y^2(0)}\right) \mu_y$$

$$(4.7c) \quad \hat{\sigma}_\omega^2 = \text{Plim } \hat{\sigma}_{\omega,T}^2 = \sigma_y^2(0) \left[1 - \left(\frac{\sigma_y^2(1)}{\sigma_y^2(0)}\right)^2\right].$$

In (4.7) we make use of the following two facts: if $\text{Plim } x_T = x$ and $\text{Plim } y_T = y$, then $\text{Plim } x_T y_T = xy$, and if $g(\cdot)$ is continuous, then $\text{Plim } g(x_T) = g(x)$ (see, e.g., Theil (1971, p. 371)). Substituting (4.1) and (4.2) into (4.7),

$$(4.8a) \quad \hat{b}^1 = \frac{(1 + \rho^2)\sigma_x^2(1) - \rho\sigma_x^2(2) - \rho\sigma_x^2(0)}{(1 + \rho^2)\sigma_x^2(0) - 2\rho\sigma_x^2(1)}$$

$$(4.8b) \quad \hat{b}^0 = (1 - \hat{b}^1)(1 - \rho)\mu_x$$

$$(4.8c) \quad \hat{\sigma}_\omega^2 = [(1 + \rho^2)\sigma_x^2(0) - 2\rho\sigma_x^2(1)][1 - (\hat{b}^1)^2] .$$

We consider now the behavior of the approximation error formulae, (3.33), as $\sigma_\varepsilon^2 \rightarrow 0$. Consider first $\hat{\alpha}/\alpha$. From (3.17a), (4.8a) and (4.6) we get,

$$(4.9) \quad \frac{\hat{\alpha}}{\alpha} = \frac{\hat{b}^1}{\alpha} = \frac{(1 + \rho^2)d_1(\sigma_\varepsilon^2)^\rho d_2(\sigma_\varepsilon^2)^\alpha - \rho d_1(\sigma_\varepsilon^2)^\rho d_2(\sigma_\varepsilon^2)^{\alpha^2} - \rho d_1(\sigma_\varepsilon^2) d_2(\sigma_\varepsilon^2) - (1 - \rho)^2}{\alpha[(1 + \rho^2)d_1(\sigma_\varepsilon^2) d_2(\sigma_\varepsilon^2) - 2\rho d_1(\sigma_\varepsilon^2)^\rho d_2(\sigma_\varepsilon^2)^\alpha - (1 - \rho)^2]}$$

$$= \frac{h(\sigma_\varepsilon^2)}{g(\sigma_\varepsilon^2)} ,$$

say. Here,

$$(4.9a) \quad d_1(\sigma_\varepsilon^2) = \exp\left[\frac{\sigma_\varepsilon^2 \rho}{(1 - \rho^2)(1 - \alpha\rho)(\rho - \alpha)}\right]$$

$$(4.9b) \quad d_2(\sigma_\varepsilon^2) = \exp\left[\frac{\sigma_\varepsilon^2 \alpha}{(1 - \alpha^2)(1 - \alpha\rho)(\alpha - \rho)}\right] .$$

The functions $h(\sigma_\varepsilon^2)$ and $g(\sigma_\varepsilon^2)$ are defined implicitly in (4.9). They are, respectively, the numerator and denominator in the expression to the right of the second equality in (4.9). Note that $d_1(0) = d_2(0) = 1$. Hence, $h(0) =$

$g(0) = 0$, so that $\hat{\alpha}/\alpha$ is not defined at the point $\sigma_\varepsilon^2 = 0$. However, we can obtain $\lim_{\sigma_\varepsilon^2 \rightarrow 0} \frac{\hat{\alpha}}{\alpha}$ by applying L'Hospital's rule (see Bartle (1964, p. 215)).

Accordingly,

$$\lim_{\sigma_\varepsilon^2 \rightarrow 0} \frac{\hat{\alpha}}{\alpha} = \lim_{\sigma_\varepsilon^2 \rightarrow 0} \frac{h(\sigma_\varepsilon)}{\alpha g(\sigma_\varepsilon^2)} = \frac{h'(\sigma_\varepsilon)}{\alpha g'(\sigma_\varepsilon^2)} \Bigg|_{\sigma_\varepsilon^2=0}$$

where the prime symbol denotes differentiation with respect to σ_ε^2 . Algebra yields that $(h'(0)/g'(0)) = \alpha$. Consequently,

$$(4.10) \quad \lim_{\sigma_\varepsilon^2 \rightarrow 0} \frac{\hat{\alpha}}{\alpha} = 1$$

Since $\mu_x = [B \exp(\frac{\mu}{1-\rho})]^{1-\alpha}$ at the point $\sigma_\varepsilon^2 = 0$, and is continuous there,

$$(4.11) \quad \lim_{\sigma_\varepsilon^2 \rightarrow 0} \hat{b}^0 = (1-\alpha)(1-\rho)[B \exp(\frac{\mu}{1-\rho})]^{1-\alpha}.$$

Similarly, from (4.8c),

$$(4.12) \quad \lim_{\sigma_\varepsilon^2 \rightarrow 0} \hat{\sigma}_w^2 = [(1+\rho^2) \cdot 0 - 2\rho \cdot 0][1-\alpha^2] = 0.$$

Taking the probability limit of (3.17b), letting $\sigma_\varepsilon^2 \rightarrow 0$ and substituting from (4.11) and (4.10),

$$(4.13) \quad \lim_{\sigma_\varepsilon^2 \rightarrow 0} \hat{x}_s = [B \exp(\frac{\mu}{1-\rho})]^{1-\alpha}.$$

Doing the same for (3.17c) and (3.17d) we get

$$(4.14) \quad \lim_{\sigma_\varepsilon^2 \rightarrow 0} \hat{\gamma} = \theta/(1-\alpha\beta) = \gamma$$

$$(4.15) \quad \lim_{\sigma_\varepsilon^2 \rightarrow 0} \hat{B} = \beta \alpha \left(\frac{H\beta\gamma(1-\alpha)}{1 + \beta\gamma(1-\alpha)} \right)^{(1-\alpha)} = B.$$

Substituting these results into (3.17e) and (3.17f), we get

$$(4.16) \quad \lim_{\sigma_\varepsilon^2 \rightarrow 0} \hat{\sigma}_\varepsilon^2 = \left(\lim_{\sigma_\varepsilon^2 \rightarrow 0} \hat{\sigma}_\omega^2 \right) / \left(\lim_{\sigma_\varepsilon^2 \rightarrow 0} \hat{x}_B \right)^2 = 0$$

$$(4.17) \quad \lim_{\sigma_\varepsilon^2 \rightarrow 0} \hat{\mu} = \ln \left[B \exp \left(\frac{\mu}{1-\rho} \right) B^{-1} \right] (1-\rho) = \left(\frac{\mu}{1-\rho} \right) (1-\rho) = \mu.$$

In deriving (4.10)-(4.17) use has been made of the fact that the functions involved satisfy the appropriate continuity conditions. We summarize the results obtained as follows:

Proposition 1

If,

- (i) estimators for parameters α , μ , σ_ε^2 are given in (3.17),
- (ii) $y_t = (1 - \rho L)x_t$ and the $\{x_t\}$ process is generated by (3.9a),

then

- (iii) the probability limit of the estimators for α , μ and σ_ε^2 , denoted by $\hat{\alpha}$, $\hat{\mu}$, $\hat{\sigma}_\varepsilon^2$, respectively, is given by substituting (4.5), (4.6) and (4.8) into the probability limit of the expressions in (3.17),
- (iv) $\lim_{\sigma_\varepsilon^2 \rightarrow 0} (\hat{\alpha}, \hat{\mu}, \hat{\sigma}_\varepsilon^2) = (\alpha, \mu, \sigma_\varepsilon^2)$.

Statement (iv) of Proposition 1 is consistent with the intuition mentioned in the introduction. Given sufficiently small variances in the random variables of the problem, one expects the IQ approximation method to imply

little error of inference. In interpreting Proposition 1 it is important to keep in mind the fact that it applies to a particular example (condition (ii)) and that no evidence is presented here as to robustness.

In fact, statement (iv) is neither sufficient nor even necessary for an estimator to be a "good" one. The data which the econometrician confronts usually implies a value of σ_ϵ^2 significantly different from zero. In the next section numerical values of the error criterion, (3.33), are tabulated for "reasonable" values of σ_ϵ^2 .

5. Numerical Results

In this section results of the previous section are used to tabulate values of the measurement error formulae, (3.33), under alternative parameterizations. In all cases we have set $H = 1$, $\theta = 2$, $\beta = \frac{1}{1.05}$, $\mu = 1$ and $\alpha = .25$, while $\sigma_\epsilon (= \sqrt{\sigma_\epsilon^2})$ and ρ are permitted to vary. Settings for σ_ϵ and ρ , given the fixed parameters, imply values for the standard deviation in the natural logarithm of output, Y_t , which we denote $\sigma_{\ln Y}$. Output, Y_t , is related to inputs according to (3.2b). We consider "reasonable" values for $\sigma_{\ln Y}$ to be in the range .017-.036. Taylor (1980; table 1, p. 221) reports that the standard deviation of the linearly detrended log of real, annual output falls in this range for ten industrial countries over the 1954-1976 period.

We compute $\sigma_{\ln Y}$ as follows. Taking the logarithm of (3.2b),

$$(5.1) \quad \ln Y_t = \ln L_{t-1}^{(1-\alpha)} + \alpha \ln x_{t-1} + v_t .$$

Substituting from (3.2b) into the second row of (3.9a), we find that the optimal choice of x_t is $x_t = \beta \alpha L_{t-1}^{(1-\alpha)} x_{t-1}^\alpha \lambda_t = \beta \alpha Y_t$. Substituting this and the first row of (3.9) into (5.1), we get

$$\begin{aligned}
 \ln Y_t &= \ln(B/\beta\alpha) + \alpha \ln \beta \alpha Y_{t-1} + v_t \\
 (5.2) \quad &= [\ln B - (1 - \alpha) \ln \beta \alpha] + \alpha \ln Y_{t-1} + v_t .
 \end{aligned}$$

Writing (5.2) in lag operator form,

$$(1 - \alpha L) \ln Y_t = \left[\ln B - (1 - \alpha) \ln \beta \alpha + \frac{\mu}{1 - \rho} \right] + \frac{\varepsilon_t}{1 - \rho L} ,$$

or,

$$(5.3) \quad \ln Y_t = C' + \frac{\varepsilon_t}{(1 - \alpha L)(1 - \rho L)} ,$$

where $C' = \left(\frac{1}{1 - \alpha} \right) \left[\ln B - (1 - \alpha) \ln \beta \alpha + \frac{\mu}{1 - \rho} \right]$. Given (5.3) it can be shown that^{9/}

$$(5.4) \quad \sigma_{\ln Y} = \sigma_{\varepsilon} \left[\frac{\rho}{(1 - \rho^2)(1 - \alpha\rho)(\rho - \alpha)} + \frac{\alpha}{(1 - \alpha^2)(1 - \alpha\rho)(\alpha - \rho)} \right]^{1/2} .$$

The result that emerges from Tables 1 and 2 is that for all values of ρ and σ_{ε} considered that place $\sigma_{\ln Y}$ in the "reasonable" range, the measures of approximation error equal 1 to within two significant digits. The standard deviation in $\ln Y_t$ has to reach twenty percent before estimators are 1 or 2 percent off their mark in probability. The deterioration in the accuracy of estimators appears to be nonlinear in σ_{ε} and ρ . For example, in Table 1 when $\rho = .1$, doubling σ_{ε} from .01 to .02 leads to no perceptible decline in the accuracy of the estimator, while doubling σ_{ε} from .5 to 1. leads to a substantial decline in accuracy. Similarly, fixing $\sigma_{\varepsilon} = .03$ in Table 2 there is no perceptible decline in accuracy when ρ is doubled from $-.1$ to $-.2$. Doubling ρ from $-.4$ to $-.8$, on the other hand, leads to a one percent drop in accuracy in the estimators for α and σ_{ε} . (The term "accuracy" is being used here to denote the ratio of the probability limit of an estimator to its true value.)

6. Conclusions and Suggestions for Further Research

This paper has shown that, for a particular example, the LQ approximation method is highly accurate, given the kind of variability one finds in economic data.^{10/} This result in itself gives little comfort to the analyst who uses the approximation method, since no evidence is presented as to the robustness of the result. It is hoped, however, that the approach taken in the paper can be generalized to obtain bounds on specification error given general conditions on return functions and constraint sets.

A weaker result than that of obtaining bounds on specification error is to establish that the specification error vanishes in probability as the underlying error variances become arbitrarily small. This result, for problems that satisfy the conditions of the theorems, seems quite likely to be true, although this paper provides a rigorous proof only in a special case (Proposition 1).

Most of the conditions of Theorem 1 restrict the exact value function and optimal decision rule of a problem. This serves the purposes of the paper since the value function and exact optimal decision rule are known for the example studied. When the LQ approximation method is applied in practice, however, this is done precisely because the exact value function and decision rules are unknown. Consequently, Theorem 1 would be more useful if the implications for the return function and constraints of the conditions assumed by the theorems were known.

37
TABLE 1*

σ_ϵ	$\frac{\delta_\epsilon}{\sigma_\epsilon}$	$\frac{\hat{\mu}}{\mu}$	$\frac{\hat{\alpha}}{\alpha}$	$\sigma_{\ln Y}$	$\frac{\delta_\epsilon}{\sigma_\epsilon}$	$\frac{\hat{\mu}}{\mu}$	$\frac{\hat{\alpha}}{\alpha}$	$\sigma_{\ln Y}$
$\rho = 0$								
.005	1.000	1.000	1.000	.005	1.000	1.000	1.000	.005
.010	1.000	1.000	1.000	.010	1.000	1.000	1.000	.010
.020	1.000	1.000	1.000	.021	1.000	1.000	1.000	.021
.030	1.000	1.000	1.000	.031	1.000	1.001	1.000	.031
.040	1.000	1.000	.999	.041	1.000	1.001	.999	.042
.050	1.001	1.002	.999	.052	1.001	1.002	.999	.052
.100	1.003	1.007	.996	.103	1.003	1.008	.995	.105
.200	1.012	1.030	.984	.207	1.013	1.031	.981	.209
.500	1.077	1.191	.902	.516	1.082	1.204	.884	.524
1.000	1.363	1.838	.641	1.033	1.386	1.921	.573	1.047
$\rho = .1$								
.005	1.000	1.000	1.000	.005	1.000	1.000	1.000	.006
.010	1.000	1.000	1.000	.011	1.000	1.000	1.000	.011
.020	1.000	1.000	1.000	.021	1.000	1.000	1.000	.022
.030	1.000	1.001	1.000	.032	1.000	1.001	.999	.033
.040	1.001	1.001	.999	.043	1.001	1.001	.999	.044
.050	1.001	1.002	.999	.053	1.001	1.002	.998	.055
.100	1.003	1.008	.994	.106	1.004	1.009	.993	.111
.200	1.013	1.033	.978	.213	1.016	1.037	.971	.222
.500	1.088	1.217	.865	.532	1.103	1.243	.824	.554
1.000	1.414	2.014	.503	1.064	1.489	2.238	.356	1.108
$\rho = .4$								
.005	1.000	1.000	1.000	.006	1.000	1.000	1.000	.011
.010	1.000	1.000	1.000	.012	1.000	1.000	1.000	.021
.020	1.000	1.000	1.000	.025	1.001	1.001	.999	.042
.030	1.001	1.001	.999	.037	1.002	1.001	.998	.063
.040	1.001	1.002	.998	.050	1.003	1.002	.996	.084
.050	1.001	1.003	.997	.062	1.005	1.003	.994	.105
.100	1.006	1.011	.989	.125	1.021	1.013	.977	.211
.200	1.023	1.043	.956	.249	1.086	1.054	.911	.422
.500	1.153	1.296	.730	.623	1.674	1.359	.493	1.054
1.000	1.765	3.412	.041	1.246	0**	0	0	0
$\rho = .9$								
.005	1.000	1.000	1.000	.015				
.010	1.000	1.000	1.000	.030				
.020	1.002	1.001	.999	.060				
.030	1.004	1.001	.998	.089				
.040	1.007	1.002	.996	.119				
.050	1.011	1.003	.993	.149				
.100	1.044	1.014	.974	.298				
.200	1.188	1.054	.897	.596				
.500	2.936	1.331	.458	1.489				
1.000	0	0	0	0				

(*), (**) See notes to Table 2

TABLE 2*

σ_ε	$\frac{\hat{\delta}_\varepsilon}{\sigma_\varepsilon}$	$\frac{\hat{\mu}}{\mu}$	$\frac{\hat{\alpha}}{\alpha}$	$\sigma_{\ln Y}$	$\frac{\hat{\delta}_\varepsilon}{\sigma_\varepsilon}$	$\frac{\hat{\mu}}{\mu}$	$\frac{\hat{\alpha}}{\alpha}$	$\sigma_{\ln Y}$
$\rho = -.05$								
.005	1.000	1.000	1.000	.005	1.000	1.000	1.000	.005
.010	1.000	1.000	1.000	.010	1.000	1.000	1.000	.010
.020	1.000	1.000	1.000	.020	1.000	1.000	1.000	.020
.030	1.000	1.001	1.000	.031	1.000	1.001	1.000	.030
.040	1.000	1.001	.999	.041	1.000	1.001	1.000	.040
.050	1.001	1.002	.999	.051	1.001	1.002	.999	.051
.100	1.003	1.007	.997	.102	1.003	1.007	.997	.101
.200	1.011	1.028	.987	.204	1.011	1.026	.990	.202
.500	1.073	1.179	.920	.511	1.070	1.166	.938	.506
1.000	1.344	1.761	.709	1.021	1.330	1.690	.775	1.012
$\rho = -.2$								
.005	1.000	1.000	1.000	.005	1.000	1.000	1.000	.005
.010	1.000	1.000	1.000	.010	1.000	1.000	1.000	.010
.020	1.000	1.000	1.000	.020	1.000	1.000	1.000	.020
.030	1.000	1.000	1.000	.030	1.000	1.000	1.000	.031
.040	1.000	1.001	1.000	.040	1.000	1.000	1.000	.041
.050	1.001	1.001	1.000	.050	1.001	1.001	1.001	.051
.100	1.003	1.006	.999	.100	1.003	1.002	1.003	.102
.200	1.010	1.022	.996	.201	1.011	1.010	1.011	.204
.500	1.067	1.139	.974	.501	1.074	1.066	1.066	.510
1.000	1.315	1.560	.908	1.003	1.345	1.314	1.208	1.019
$\rho = -.8$								
.005	1.000	1.000	1.000	.007	1.000	.999	1.001	.009
.010	1.000	.999	1.001	.014	1.001	.994	1.005	.019
.020	1.001	.996	1.004	.028	1.002	.977	1.018	.038
.030	1.001	.991	1.008	.042	1.005	.950	1.041	.057
.040	1.003	.984	1.014	.056	1.009	.915	1.071	.075
.050	1.004	.975	1.023	.070	1.015	.873	1.110	.094
.100	1.016	.909	1.087	.141	1.053	.627	1.388	.188
.200	1.059	.728	1.310	.281	1.167	.314	2.058	.377
.500	1.312	.481	2.066	.703	1.718	.303	2.970	.942
1.000	2.384	.871	2.426	1.405	0**	0	0	0
$\rho = -.9$								

(*) In all cases $H = 1$, $\mu = 1$, $\alpha = .25$, $\beta = 1/1.05$, $\theta = 2$. $\hat{\alpha}$, $\hat{\mu}$, $\hat{\delta}_\varepsilon$ are the probability limits of estimators of α , μ , σ_ε , respectively, which make use of the LQ approximation method. They are computed by substituting (4.7) into the probability limit of the expressions in (3.17) (e.g., dropping the T subscript on all variables). The calculations were performed in double precision on the University of Minnesota Cyber 172 computer.

(**) A zero indicates the calculations could not be completed due to computer overflow problems.

APPENDIX A

This appendix provides a proof of Theorem 1 stated in section 2 of the text. We begin with six lemmas that are required in the proof.

Lemma 1

Given the process defined in (2.2) and an arbitrary vector v_s ,

$$(i) \quad E_t(v_{t+1} - v_s) = \rho(v_t - v_s) + [\mu - (I - \rho)v_s],$$

$$(ii) \quad \begin{aligned} E_t(v_{t+1} - v_s)(v_{t+1} - v_s)^T &= \rho(v_t - v_s)(v_t - v_s)^T \rho^T \\ &+ [\mu - (I - \rho)v_s](v_t - v_s)^T \rho^T + \rho(v_t - v_s)[\mu - (I - \rho)v_s]^T \\ &+ [\mu - (I - \rho)v_s][\mu - (I - \rho)v_s]^T + v_\varepsilon, \end{aligned}$$

$$(iii) \quad (I - \rho)E v_t = \mu.$$

Proof

By (2.2), $v_{t+1} - v_s = \rho(v_t - v_s) + \mu - (I - \rho)v_s + \varepsilon_{t+1}$. Equation (i) follows because $E_t \varepsilon_{t+1} = 0$.

Substituting,

$$\begin{aligned} E_t(v_{t+1} - v_s)(v_{t+1} - v_s)^T &= E_t\{[\rho(v_t - v_s) + \mu - (I - \rho)v_s] + \varepsilon_{t+1}\} \\ &\times \{[\rho(v_t - v_s) + \mu - (I - \rho)v_s] + \varepsilon_{t+1}\}^T \\ &= [\rho(v_t - v_s) + \mu - (I - \rho)v_s][\rho(v_t - v_s) + \mu - (I - \rho)v_s]^T + v_\varepsilon, \end{aligned}$$

since $E_t \varepsilon_{t+1} \varepsilon_{t+1}^T = E \varepsilon_t \varepsilon_t^T = v_\varepsilon$ and $E_t[\rho(v_t - v_s) + \mu - (I - \rho)v_s] \varepsilon_{t+1} = 0$.

Equation (ii) then follows.

Taking unconditional expectations of (2.2) $EV_t = \rho EV_{t-1} + \mu$ since $E\varepsilon_t = 0$. But $\{v_t\}$ is stationary in the mean, so that $(I - \rho)EV_t = \mu$, which is (iii). Q.E.D.

Lemma 2 ("Certainty Equivalence")

If,

- (i) $v(z_{t-1}, v_t, V_\varepsilon) = \tilde{v}(z_{t-1}, v_t) + \ell(V_\varepsilon)$, where $\tilde{v}(z_{t-1}, v_t)$ is functionally independent of the elements of V_ε , and v is defined in (2.4),
- (ii) the solution to (2.1) exists, is unique and is of the form given in (2.3),

then,

- (iii) the function f in (2.3) is functionally independent of the elements of V_ε .

Proof

The proof is accomplished by assuming the contrary of (iii) and deriving a contradiction.

Write $z_t^* = f(z_{t-1}, v_t, V_\varepsilon)$. Let \tilde{V}_ε be a positive semi-definite matrix obtained by altering one or several of the elements in V_ε . Write $\tilde{z}_t^* = f(z_{t-1}, v_t, \tilde{V}_\varepsilon)$ and suppose that

$$(A.1) \quad z_t^* \neq \tilde{z}_t^* \quad \text{for some } z_{t-1}, v_t.$$

Condition (ii) implies that

$$u(z_t^*, z_{t-1}, v_t) + \beta E_t v(z_t^*, v_{t+1}, V_\varepsilon) > u(z_t, z_{t-1}, v_t) + \beta E_t v(z_t, v_{t+1}, V_\varepsilon)$$

for all z_t such that $(z_t, z_{t-1}, v_t) \in T$. In particular,

$$(A.2) \quad u(z_t^*, z_{t-1}, v_t) + \beta E_t v(z_t^*, v_{t+1}, V_\varepsilon) > u(\tilde{z}_t^*, z_{t-1}, v_t) + \beta E_t v(\tilde{z}_t^*, v_{t+1}, V_\varepsilon).$$

Arguing along the same lines, get

$$(A.3) \quad u(z_t^*, z_{t-1}, v_t) + \beta E_t v(z_t^*, v_{t+1}, \tilde{V}_\varepsilon) > u(z_t^*, z_{t-1}, v_t) + \beta E_t v(z_t^*, v_{t+1}, \tilde{V}_\varepsilon) .$$

Adding $\beta(\ell(V_\varepsilon) - \ell(\tilde{V}_\varepsilon))$ to both sides of (A.3) and making use of (i),

$$(A.4) \quad u(z_t^*, z_{t-1}, v_t) + \beta E_t v(z_t^*, v_{t+1}, V_\varepsilon) > u(z_t^*, z_{t-1}, v_t) + \beta E_t v(z_t^*, v_{t+1}, V_\varepsilon) .$$

Relations (A.2) and (A.4) imply

$$(A.5) \quad v(z_{t-1}, v_t, V_\varepsilon) = u(z_t^*, z_{t-1}, v_t) + \beta E_t v(z_t^*, v_{t+1}, V_\varepsilon) .$$

According to (A.5) the following plan:

$$\{f(z_{t-1}, v_t, \tilde{V}_\varepsilon), f(z_t, v_{t+1}, V_\varepsilon), f(z_{t+1}, v_{t+2}, V_\varepsilon), \dots\}$$

produces the same return as the different (see (A.1)) plan

$$\{f(z_{t-1}, v_t, V_\varepsilon), f(z_t, v_{t+1}, V_\varepsilon), f(z_{t+1}, v_{t+2}, V_\varepsilon), \dots\} .$$

This contradicts (iii), according to which the latter is the unique optimizing plan.

Q.E.D.

Lemma 3

If,

(i) $v(z_{t-1}, v_t, V_\varepsilon) = \tilde{v}(z_{t-1}, v_t) + \ell(V_\varepsilon)$, where $\tilde{v}(z_{t-1}, v_t)$ is functionally independent of the elements of V_ε ,

(ii) the solution to (2.1) exists, is of the form given in (2.3) and $(f(z_{t-1}, v_t, V_\varepsilon), z_{t-1}, v_t)$ is interior to the feasible set T ,

then

$$(iii) \quad \begin{aligned} & \beta E[v(z_t, v_{t+1}, V_\varepsilon) | v_t = v_s] \\ & = \beta v(z_s, v_s, V_\varepsilon) + (1 - \beta)[\ell(V_\varepsilon) - \ell(0)] \end{aligned}$$

$$\begin{aligned}
\text{(iv)} \quad & E[v_1(z_s, v_{t+1}, \nabla_\varepsilon) | v_t = v_s] = v_1(z_s, v_s, \nabla_\varepsilon) \\
\text{(v)} \quad & E[v_2(z_s, v_{t+1}, \nabla_\varepsilon) | v_t = v_s] \rho = v_2(z_s, v_s, \nabla_\varepsilon) \rho \\
\text{(vi)} \quad & f_1^\top(z_s, v_s, \nabla_\varepsilon) E[v_{11}(z_s, v_{t+1}, \nabla_\varepsilon) | v_t = v_s] \\
& = f_1^\top(z_s, v_s, \nabla_\varepsilon) v_{11}(z_s, v_s, \nabla_\varepsilon) \\
& \rho^\top \beta v_{22}(z_s, v_s, \nabla_\varepsilon) \rho + \beta f_2^\top(z_s, v_s, \nabla_\varepsilon) v_{12}(z_s, v_s, \nabla_\varepsilon) \rho \\
\text{(vii)} \quad & = \rho^\top \beta E[v_{22}(z_s, v_{t+1}, \nabla_\varepsilon) | v_t = v_s] \rho \\
& \quad + \beta f_2^\top(z_s, v_s, \nabla_\varepsilon)^\top E[v_{12}(z_s, v_{t+1}, \nabla_\varepsilon) | v_t = v_s] \rho \\
& \beta f_2^\top(z_s, v_s, \nabla_\varepsilon) E[v_{11}(z_s, v_{t+1}, \nabla_\varepsilon) | v_t = v_s] \\
\text{(viii)} \quad & + \rho^\top \beta E[v_{21}(z_s, v_{t+1}, \nabla_\varepsilon) | v_t = v_s] \\
& = \beta f_2^\top(z_s, v_s, \nabla_\varepsilon) v_{11}(z_s, v_s, \nabla_\varepsilon) + \rho^\top \beta v_{21}(z_s, v_s, \nabla_\varepsilon) ,
\end{aligned}$$

where $v_s = E v_t$.

Proof

Consider each formula in the order in which it appears, beginning with

(iii).

Formula (iii): By (ii)

$$\begin{aligned}
\text{(A.6)} \quad & v(z_{t-1}, v_t, \nabla_\varepsilon) = u(f(z_{t-1}, v_t, \nabla_\varepsilon), z_{t-1}, v_t) \\
& \quad + \beta E[v(f(z_{t-1}, v_t, \nabla_\varepsilon), v_{t+1}, \nabla_\varepsilon) | v_t] .
\end{aligned}$$

Under (i), (A.6) implies

$$\begin{aligned}
\tilde{v}(z_{t-1}, v_t) + \beta l(\nabla_\varepsilon) & = u(f(z_{t-1}, v_t, \nabla_\varepsilon), z_{t-1}, v_t) \\
& \quad + \beta E[\tilde{v}(f(z_{t-1}, v_t, \nabla_\varepsilon), v_{t+1}) | v_t] + \beta l(\nabla_\varepsilon) ,
\end{aligned}$$

so that, when $z_{t-1} = z_s$ and $v_t = v_s$,

$$(A.7) \quad \tilde{v}(z_s, v_s) + \ell(v_\varepsilon) = u(z_s, z_s, v_\varepsilon) + \beta E[\tilde{v}(z_s, v_{t+1}) | v_t = v_s] + \beta \ell(v_\varepsilon) .$$

In (A.7) the fact $z_s = f(z_s, v_s, 0) = f(z_s, v_s, v_\varepsilon)$, established in Lemma 2, has been used.

The argument leading to (A.7) applies for the case $v_\varepsilon = 0$ also, that is,

$$(A.8) \quad \tilde{v}(z_s, v_s) + \ell(0) = u(z_s, z_s, v_s) + \rho \tilde{v}(z_s, v_s) + \beta \ell(0) .$$

Subtracting (A.8) from (A.7),

$$\begin{aligned} \ell(v_\varepsilon) - \ell(0) &= \rho E[v(z_s, v_s, v_\varepsilon) | v_t = v_s] - \rho \tilde{v}(z_s, v_s) - \beta \ell(0) \\ &= \beta E[v(z_s, v_{t+1}, v_\varepsilon) | v_t = v_s] - \beta v(z_s, v_s, v_\varepsilon) \\ &\quad + \beta [\ell(v_\varepsilon) - \ell(0)] . \end{aligned}$$

Formula (iii) follows.

Formula (iv): Because the solution is interior to the feasible set, the following condition obtains,

$$(A.9) \quad \begin{aligned} &u_1(f(z_{t-1}, v_t, v_\varepsilon), z_{t-1}, v_t) + \beta E[v_1(f(z_{t-1}, v_t, v_\varepsilon), v_{t+1}, v_\varepsilon) | v_t] \\ &= 0 = h(z_{t-1}, v_t, v_\varepsilon) , \end{aligned}$$

say, for all z_{t-1}, v_t such that $(f(z_{t-1}, v_t), z_{t-1}, v_t)$ is interior to T . In the case $v_\varepsilon = 0$,

$$(A.10) \quad \begin{aligned} &u_1(f(z_{t-1}, v_t, 0), z_{t-1}, v_t) \\ &+ \beta v_1(f(z_{t-1}, v_t, 0), \rho v_t + u, 0) = 0 . \end{aligned}$$

Use the fact, $f(z_{t-1}, v_t, 0) = f(z_{t-1}, v_t, v_\varepsilon)$, established in Lemma 2 and subtract (A.10) from (A.9) to get

$$(A.11) \quad \begin{aligned} E[v_1(f(z_{t-1}, v_t, V_\varepsilon), v_{t+1}, V_\varepsilon) | v_t] \\ = v_1(f(z_{t-1}, v_t, V_\varepsilon), \rho v_t + \mu, V_\varepsilon), \end{aligned}$$

where the fact $v_1(\cdot, \cdot, 0) = v_1(\cdot, \cdot, V_\varepsilon)$ from (i) has been used. Formula (iv) follows upon substituting $z_{t-1} = z_s$ and $v_t = v_s$ into (A.11) and using $z_s = f(z_s, v_s, 0) = f(z_s, v_s, V_\varepsilon)$.

Formula (v): Differentiate (A.6) with respect to v_t to get

$$(A.11') \quad \begin{aligned} v_2(z_{t-1}, v_t, V_\varepsilon) &= h(z_{t-1}, v_t, V_\varepsilon) f_2(z_{t-2}, v_t, V_\varepsilon) \\ &+ SE[v_2(f(z_{t-1}, v_t, V_\varepsilon), v_{t+1}, V_\varepsilon) | v_t] \rho \\ &+ u_3(f(z_{t-1}, v_t, V_\varepsilon), z_{t-1}, v_t) \end{aligned}$$

where the function h , which is identically zero, is defined in (A.9). Thus, (A.11) implies

$$(A.12) \quad \begin{aligned} v_2(z_{t-1}, v_t, V_\varepsilon) &= SE[v_2(f(z_{t-1}, v_t, V_\varepsilon), v_{t+1}, V_\varepsilon) | v_t] \rho \\ &+ u_3(f(z_{t-1}, v_t, V_\varepsilon), z_{t-1}, v_t). \end{aligned}$$

Setting $z_{t-1} = z_s$ and $v_t = v_s$ in (A.12),

$$(A.13) \quad v_2(z_s, v_s, V_\varepsilon) = SE[v_2(z_s, v_{t+1}, V_\varepsilon) | v_t = V_\varepsilon] \rho + u_3(z_s, z_s, v_s).$$

The argument leading to (A.13) is also valid for the case $V_\varepsilon = 0$, so that,

$$(A.14) \quad v_2(z_s, v_s, 0) = \beta v_2(z_s, v_s, V_\varepsilon) \rho + u_3(z_s, z_s, v_s).$$

Subtracting (A.14) from (A.13) and taking the fact $v_2(z_s, v_s, 0) = v_2(z_s, v_s, V_\varepsilon)$ into account, formula (v) follows.

Formula (vi): Using the function h defined in (A.9),

$$(A.15) \quad \frac{\partial h(z_{t-1}, v_t, V_\varepsilon)}{\partial z_{t-1}} = f_1^T(z_{t-1}, v_t, V_\varepsilon) u_{11}(f(z_{t-1}, v_t, V_\varepsilon), z_{t-1}, v_t) \\ + \beta E[f_1^T(z_{t-1}, v_t, V_\varepsilon) v_{11}(f(z_{t-1}, v_t, V_\varepsilon), v_{t+1}, V_\varepsilon) | v_t] \\ + u_{21}(f(z_{t-1}, v_t, V_\varepsilon), z_{t-1}, v_t) = 0.$$

The same result holds for the case $V_\varepsilon = 0$. Formula (vi) follows upon setting $z_{t-1} = z_s$, $v_t = v_s$, taking into account the fact $z_s = f(z_s, v_s, V_\varepsilon)$ and comparing (A.15) with

$$\frac{\partial h(z_{t-1}, v_t, 0)}{\partial z_{t-1}} \left| \begin{array}{l} z_{t-1} = z_s \\ v_t = v_s \end{array} \right.$$

Formula (vii): Differentiate (A.12) with respect to v_t to get

$$(A.16) \quad v_{22}(z_{t-1}, v_t, V_\varepsilon) = f_2^T(z_{t-1}, v_t, V_\varepsilon) u_{13}(f(z_{t-1}, v_t, V_\varepsilon), z_{t-1}, v_t) \\ + u_{33}(f(z_{t-1}, v_t, V_\varepsilon), z_{t-1}, v_t) \\ + \rho^T \beta E[v_{22}(f(z_{t-1}, v_t, V_\varepsilon), v_{t+1}, V_\varepsilon) | v_t] \rho \\ + \beta f_2^T(z_{t-1}, v_t, V_\varepsilon) E v_{12}(f(z_{t-1}, v_t, V_\varepsilon), v_{t+1}, V_\varepsilon) | v_t] \rho.$$

Formula (vii) follows upon noting that (A.17) holds for $V_\varepsilon = 0$, that $f(z_{t-1}, v_t, V_\varepsilon) = f(z_{t-1}, v_t, 0)$, $v_{12}(\cdot, \cdot, V_\varepsilon) = \sigma_{12}(\cdot, \cdot, 0)$ and $v_{22}(\cdot, \cdot, V_\varepsilon) = v_{22}(\cdot, \cdot, 0)$.

Formula (viii): Differentiate (A.9) with respect to v_t ,

$$(A.17) \quad \frac{\partial h(z_{t-1}, v_t, v_\varepsilon)}{\partial v_t} = f_2^T(z_{t-1}, v_t, v_\varepsilon) u_{11}(f(z_{t-1}, v_t, v_\varepsilon), z_{t-1}, v_t) \\ + u_{31}(f(z_{t-1}, v_t, v_\varepsilon), z_{t-1}, v_t) \\ + \beta f_2^T(z_{t-1}, v_t, v_\varepsilon) E[v_{11}(f(z_{t-1}, v_t, v_\varepsilon), v_{t+1}, v_\varepsilon) | v_t = v_s] \\ + \beta \rho^T E[v_{21}(f(z_{t-1}, v_t, v_\varepsilon), v_{t+1}, v_\varepsilon) | v_t = v_s]] = 0.$$

Formula (viii) follows after noting that (A.17) is valid for $v_\varepsilon = 0$, making use of Lemma 2, and using the facts $v_{11}(\cdot, \cdot, v_\varepsilon) = v_{11}(\cdot, \cdot, 0)$,

$$v_{21}(\cdot, \cdot, v_\varepsilon) = v_{21}(\cdot, \cdot, 0).$$

Q.E.D.

Lemma 4

If,

- (i) the solution to (2.1) exists, is unique, is of the form given in (2.3) and $(f(z_{t-1}, v_t, v_\varepsilon), z_{t-1}, v_t)$ is interior to the feasible set, T ,
- (ii) the stationary state, z_s , of (2.6) is finite, independent of z_{t-1} , the initial condition, and $u(z_s, z_s, v_s)$ is well defined,
- (iii) the value function $v(\cdot, \cdot, \cdot)$ is differentiable at least twice in the first two arguments,
- (iv) the function $f(\cdot, \cdot, \cdot)$ given in (2.3) has at least one derivative in the first two arguments,
- (v) $v(z_{t-1}, v_t, v_\varepsilon) = \tilde{v}(z_{t-1}, v_t) + \ell(v_\varepsilon)$, where $\tilde{v}(z_{t-1}, v_t)$ is functionally independent of the elements of v_ε ,

then

$$(vi) \quad \tilde{V}(z_{t-1}, v_t) = U(g(z_{t-1}, v_t), z_{t-1}, v_t) + \beta E \tilde{V}(g(z_{t-1}, v_t), v_{t+1}),$$

$$\text{where } \tilde{V}(z_{t-1}, v_t) = v(z_{t-1}, v_t) + \frac{1}{2} \left(\frac{\beta}{1-\beta} \right) \text{tr}[v_{22}(z_s, v_s, v_\varepsilon) v_\varepsilon] + \ell(0) - \ell(v_\varepsilon)$$

and $U(\cdot, \cdot, \cdot)$, $g(\cdot, \cdot)$ and $V(\cdot, \cdot)$ are given in (2.6), (2.8) and (2.9), respectively, and,

$$(vii) \quad U_1(g(z_{t-1}, v_t), z_{t-1}, v_t) + \beta E_t \tilde{V}(g(z_{t-1}, v_t), v_{t+1}) = 0 .$$

Proof

First we obtain (vi) of the Lemma. Substituting (2.6), (2.8) and (2.9) into the right hand side of (vii), and making use of Lemma 1 and Lemma 2 we get

$$(A.18) \quad \begin{aligned} U(g(z_{t-1}, v_t), z_{t-1}, v_t) + \beta E_t \tilde{V}(g(z_{t-1}, v_t), v_{t+1}) \\ = W(z_{t-1}, v_t), \text{ say.} \end{aligned}$$

Here,

$$(A.19) \quad \begin{aligned} W(z_{t-1}, v_t) = & W_0 + W_1(z_{t-1} - z_s) + W_2(v_t - v_s) \\ & + \frac{1}{2}(z_{t-1} - z_s)^T W_3(z_{t-1} - z_s) + \frac{1}{2}(v_t - v_s)^T W_4(v_t - v_s) \\ & + (v_t - v_s)^T W_5(z_{t-1} - z_s) , \end{aligned}$$

where

$$(A.19a) \quad W_0 = u + \beta v + \frac{1}{2} \left(\frac{\beta}{1 - \beta} \right) \text{tr } v_{22} v_\varepsilon + \beta [\ell(Q) - \ell(v_\varepsilon)]$$

$$(A.19b) \quad W_1 = u_1 f_1 + u_2 + \beta v_1 f_1$$

$$(A.19c) \quad W_2 = u_1 f_2 + u_3 + \beta v_1 f_2 + \beta v_2 \rho$$

$$(A.19d) \quad W_3 = f_1^T u_{11} f_1 + u_{22} + 2f_1^T u_{12} + \beta f_1^T v_{11} f_1$$

$$(A.19e) \quad W_4 = f_2^T u_{11} f_2 + u_{33} + 2f_2^T u_{13} + \beta f_2^T v_{11} f_2 + 2\beta f_2^T v_{12} \rho + \beta \rho^T v_{22} \rho$$

$$(A.19f) \quad W_5 = f_2^T u_{11} f_1 + f_2^T u_{12} + u_{31} f_1 + u_{32} + \beta f_2^T v_{11} f_1 + \beta \rho^T v_{21} f_1 :$$

For the sake of notational simplicity, we drop the argument list of a function when it is evaluated at z_s, v_s, v_ε and it appears that doing so does not lead

to confusion. For example, $v \equiv v(z_s, v_s, v_\varepsilon)$, $u_1 \equiv u_1(z_s, z_s, v_s)$, and $f_2 \equiv f_2(z_s, v_s, v_\varepsilon)$. Also, the functions $v(\cdot, \cdot)$ in (2.9) and $w(\cdot, \cdot)$ in (A.2) are functions of z_s, v_ε and v_s , although this is not made explicit.

We accomplish the first part of the proof by establishing that

$$W(z_{t-1}, v_t) = \tilde{V}(z_{t-1}, v_t).$$

When the initial conditions, z_{t-1}, v_t , are z_s and v_s , respectively, then (A.9) becomes

$$(A.20) \quad h(z_s, v_s, v_\varepsilon) = u_1 + \beta E[v_1(z_s, v_{t+1}, v_\varepsilon) | v_t = v_s] = 0.$$

Here we have made use of the fact that $z_s = f(z_s, v_s, 0) = f(z_s, v_s, v_\varepsilon)$, from Lemma 2. Substituting formula (iv), Lemma 3 into (A.2),

$$(A.21) \quad h(z_s, v_s, v_\varepsilon) = u_1 + \beta v_1 = 0.$$

When $z_{t-1} = z_s, v_t = v_s$, then Bellman's optimality relation is

$$(A.22) \quad v = u + \beta E[v(z_s, v_{t+1}, v_\varepsilon) | v_t = v_s].$$

Adding $\frac{1}{2} \left(\frac{\beta}{1-\beta} \right) \text{tr } v_{22} v_\varepsilon + \ell(0) - \ell(v_\varepsilon)$ to both sides of (A.22) and taking formula (iii), Lemma 3 into account

$$(A.23) \quad \begin{aligned} & v + \frac{1}{2} \left(\frac{\beta}{1-\beta} \right) \text{tr } v_{22} v_\varepsilon + \ell(0) - \ell(v_\varepsilon) \\ &= u + \beta v + (1-\beta) [\ell(v_\varepsilon) - \ell(0)] \\ & \quad + \frac{1}{2} \left(\frac{\beta}{1-\beta} \right) \text{tr } v_{22} v_\varepsilon - [\ell(0) - \ell(v_\varepsilon)] = w_0, \end{aligned}$$

by (A.19).

Differentiating (A.6) with respect to z_{t-1} :

$$\begin{aligned}
 (A.24) \quad v_1(z_{t-1}, v_t, \nabla_\varepsilon) &= h(z_{t-1}, v_t, \nabla_\varepsilon) f_1(z_{t-1}, v_t, \nabla_\varepsilon) \\
 &+ u_2(f(z_{t-1}, v_t, \nabla_\varepsilon), z_{t-1}, v_t) \\
 &= u_2(f(z_{t-1}, v_t, \nabla_\varepsilon), z_{t-1}, v_t) .
 \end{aligned}$$

Evaluating (A.24) at $z_{t-1} = z_s$, $v_t = v_s$ and making use of (A.21),

$$(A.25) \quad v_1 = u_1 f_1 + \beta v_1 f_1 + u_2 = W_1 ,$$

by (A.19b).

Differentiating (A.24) with respect to z_{t-1} ,

$$\begin{aligned}
 (A.26) \quad \frac{\partial^2 v(z_{t-1}, v_t, \nabla_\varepsilon)}{\partial z_{t-1} \partial z_{t-1}^\top} &\equiv v_{11}(z_{t-1}, v_t, \nabla_\varepsilon) \\
 &= f_1(z_{t-1}, v_t, \nabla_\varepsilon)^\top u_{12}(f(z_{t-1}, v_t, \nabla_\varepsilon), z_{t-1}, v_t) \\
 &+ u_{22}(f(z_{t-1}, v_t, \nabla_\varepsilon), z_{t-1}, v_t) .
 \end{aligned}$$

Substituting $z_{t-1} = z_s$ and $v_t = v_s$ into (A.15),

$$(A.27) \quad f_1^\top u_{11} + u_{21} + \beta f_1^\top E[v_{11}(z_s, v_{t+1}, \nabla_\varepsilon) | v_t = v_s] = 0 .$$

Making use of formula (vi), Lemma 3, the latter equation is seen to reduce to

$$(A.28) \quad f_1^\top u_{11} + u_{21} + \beta f_1^\top v_{11} = 0 .$$

Substituting $z_{t-1} = z_s$, $v_t = v_s$ into (A.26), postmultiplying (A.28) by f_1 and adding,

$$(A.29) \quad v_{11} = f_1^\top u_{12} + u_{22} + f_1^\top u_{11} f_1 + u_{21} f_1 + \beta f_1^\top v_{11} f_1 .$$

It is easily verified using (A.29) and (A.19d) that for an arbitrary (conformable) vector x ,

$$(A.30) \quad \frac{1}{2} x^\top v_{11} x = x^\top W_3 x .$$

Adding (A.21) and (A.13),

$$v_2 = u_1 f_2 + \beta v_1 f_2 + u_3 + \beta E[v_2(z_s, v_{t+1}, v_\varepsilon) | v_t = v_s] \rho.$$

Applying formula (v), Lemma 3, this reduces to

$$(A.31) \quad v_2 = u_1 f_2 + \beta v_1 f_2 + u_3 + \beta v_2 \rho = W_2,$$

by (A.19c). Evaluating (A.17) at $z_{t-1} = z_s, v_t = v_s,$

$$\begin{aligned} f_2^T u_{11} + u_{31} + \beta f_2^T E[v_{11}(z_s, v_{t+1}, v_\varepsilon) | v_t = v_s] \\ + \rho^T \beta E[v_{21}(z_s, v_{t+1}, v_\varepsilon) | v_t = v_s] = 0. \end{aligned}$$

Applying formula (viii), Lemma 3, this reduces to

$$(A.32) \quad f_2^T u_{11} + u_{31} + \beta f_2^T v_{11} + \rho^T \beta v_{21} = 0.$$

Differentiating (A.12) with respect to $v_t,$

$$\begin{aligned} v_{22}(z_{t-1}, v_t, v_\varepsilon) &= f_2(z_{t-1}, v_t, v_\varepsilon)^T u_{13}(f(z_{t-1}, v_t, v_\varepsilon), z_{t-1}, v_t) \\ (A.33) \quad &+ u_{33}(f(z_{t-1}, v_t, v_\varepsilon), z_{t-1}, v_t) + \rho^T \beta E_t v_{22}(f(z_{t-1}, v_t, v_\varepsilon), v_{t+1}) \rho \\ &+ \beta f_2(z_{t-1}, v_t, v_\varepsilon)^T E_t v_{12}(f(z_{t-1}, v_t, v_\varepsilon), v_{t+1}, v_\varepsilon) \rho. \end{aligned}$$

Postmultiplying (A.32) by $f_2,$ substituting $z_{t-1} = z_s, v_t = v_s$ in (A.33), applying formula (vii), Lemma 3, and adding yields,

$$\begin{aligned} v_{22} &= f_2^T u_{13} + u_{33} + \rho^T \beta v_{22} \rho + f_2^T u_{11} f_2 + u_{31} f_2 \\ (A.34) \quad &+ \beta f_2^T v_{11} f_2 + \beta \rho^T v_{21} f_2 + \beta f_2^T v_{12} \rho. \end{aligned}$$

Given an arbitrary (conformable) vector $x,$ (A.19e) and (A.34) imply

$$(A.35) \quad x^T v_{22} x = x^T W_4 x.$$

Differentiating (A.24) with respect to v_t ,

$$(A.36) \quad \frac{\partial^2 v(z_{t-1}, v_t, v_\varepsilon)}{\partial v_t \partial z_{t-1}^T} \equiv v_{21}(z_{t-1}, v_t, v_\varepsilon) \\ = f_2(z_{t-1}, v_t, v_\varepsilon)^T u_{12}(f(z_{t-1}, v_t, v_\varepsilon), z_{t-1}, v_t) \\ + u_{32}(f(z_{t-1}, v_t, v_\varepsilon), z_{t-1}, v_t) .$$

Postmultiplying (A.32) by f_1 , substituting $z_{t-1} = z_s$, $v_t = v_s$ into (A.36) and adding,

$$(A.37) \quad v_{21} = f_2^T u_{12} + u_{32} + f_2^T \bar{u}_{11} f_1 + u_{31} f_1 + \beta f_2^T v_{11} f_1 + \beta \rho^T v_{21} f_1 = w_5 ,$$

from (A.19f).

From (A.23), (A.25), (A.30), (A.31), (A.35) and (A.37) it follows that

$$(A.38) \quad w(z_{t-1}, v_t) = \tilde{v}(z_{t-1}, v_t) ,$$

which establishes (vi).

Next, we establish (vii) of the Lemma. Differentiating (2.6) with respect to z_t ($j = 0$) and evaluating the result at $z_t = g(z_{t-1}, v_t)$,

$$(A.39) \quad U_1(g(z_{t-1}, v_t), z_{t-1}, v_t) = u_1 + (z_{t-1} - z_s)^T [f_1^T u_{11} + u_{21}] \\ + (v_t - v_s)^T [f_2^T u_{11} + u_{31}] .$$

Differentiating $v(z_t, v_{t+1})$, defined in (2.9), with respect to z_t , evaluating the result at $z_t = g(z_{t-1}, v_t)$ and applying the $E_t(\cdot)$ operator yields,

$$(A.40) \quad \begin{aligned} E_t \nabla_1(g(z_{t-1}, v_t), v_{t+1}) &= v_1 + (z_{t-1} - z_s)^T f_1^T v_{11} + (v_t - v_s)^T f_2^T v_{11} \\ &+ (v_t - v_s)^T \rho^T v_{21}. \end{aligned}$$

Adding (A.39) and the product of (A.40) and β ,

$$(A.41) \quad \begin{aligned} U_1(g(z_{t-1}, v_t), z_{t-1}, v_t) + \beta E_t \nabla_1(g(z_{t-1}, v_t), v_{t+1}) \\ = u_1 + \beta v_1 + (z_{t-1} - z_s)^T [f_1^T u_{11} + u_{21} + \beta f_1^T v_{11}] \\ + (v_t - v_s)^T [f_2^T u_{11} + u_{31} + \beta f_2^T v_{11} + \beta \rho^T v_{21}]. \end{aligned}$$

By (A.21), the constant term in (A.41) vanishes. Equations (A.28) and (A.32) indicate that the coefficient matrices on $(z_{t-1} - z_s)^T$ and $(v_t - v_s)^T$, respectively, also vanish. This establishes (vii) and, hence, the Lemma.

Q.E.D.

Lemma 5

If,

- (i) problem (2.1) has the solution, (2.3),
- (ii) the set T is convex in its first two dimensions
- (iii) the return function, $u(*, *, *)$ is concave in its first two arguments

then,

- (iv) the return function, v , defined in (2.4), is concave in its first argument.

Proof

Consider the initial conditions (z_{-1}, v_0) and (z'_{-1}, v_0) , where $z'_{-1} \neq z_{-1}$. (In this proof we set $t = 0$.) Fix ε^t for $t > 0$. By definition,

$$(A.41) \quad \begin{aligned} & (Z(\varepsilon^t, z_{-1}, v_0, v_\varepsilon, t), v_t(\varepsilon^t, v_0)) \in T \\ & (Z(\varepsilon^t, z'_{-1}, v_0, v_\varepsilon, t), v_t(\varepsilon^t, v_0)) \in T, \end{aligned}$$

for all $t > 0$, where the Z function is defined in (2.3a).

Consider $z_{-1}(\lambda) = \lambda z_{-1} + (1 - \lambda)z'_{-1}$ for $0 < \lambda < 1$. Define

$$(A.42) \quad \begin{aligned} & \tilde{Z}(\varepsilon^t, z_{-1}(\lambda), v_0, v_\varepsilon, t) \\ & = \lambda Z(\varepsilon^t, z_{-1}, v_0, v_\varepsilon, t) \\ & \quad + (1 - \lambda)Z(\varepsilon^t, z'_{-1}, v_0, v_\varepsilon, t) \end{aligned}$$

for $t > 0$. (\tilde{Z} is not necessarily optimal. It is feasible because of the convexity of T .) Because of (ii), (A.41) and (A.42) imply,

$$(A.42') \quad (\tilde{Z}(\varepsilon^t, z_{-1}(\lambda), v_0, v_\varepsilon, t), v_t(\varepsilon^t, v_0)) \in T$$

for all $t > 0$. Condition (ii) implies

$$\begin{aligned} & u(\tilde{Z}(\varepsilon^t, z_{-1}(\lambda), v_0, v_\varepsilon, t), v_t(\varepsilon^t, v_0)) \\ & > \lambda u(Z(\varepsilon^t, z_{-1}, v_0, v_\varepsilon, t), v_t(\varepsilon^t, v_0)) \\ & \quad + (1 - \lambda)u(Z(\varepsilon^t, z'_{-1}, v_0, v_\varepsilon, t), v_t(\varepsilon^t, v_0)) \end{aligned}$$

for all $t > 0$, which implies

$$\begin{aligned}
& \sum_{t=0}^{\infty} \beta^t u(\tilde{Z}(\varepsilon^t, z_{-1}(\lambda), v_0, v_\varepsilon, t), v_t(\varepsilon^t, v_0)) \\
(A.43) \quad & > \lambda \sum_{t=0}^{\infty} \beta^t u(Z(\varepsilon^t, z_{-1}, v_0, v_\varepsilon, t), v_t(\varepsilon^t, v_0)) \\
& + (1 - \lambda) \sum_{t=0}^{\infty} \beta^t u(Z(\varepsilon^t, z'_{-1}, v_0, v_\varepsilon, t), v_t(\varepsilon^t, v_0)) .
\end{aligned}$$

Because of (A.42), which indicates that \tilde{Z} is feasible,

$$(A.44) \quad v(z_{-1}(\lambda), v_0, v_\varepsilon) > E \left[\sum_{t=0}^{\infty} \beta^t u(\tilde{Z}(\varepsilon^t, z_{-1}(\lambda), v_0, v_\varepsilon, t), v_t(\varepsilon^t, v_0)) \right]$$

where the expectation is taken over the random vectors $\{\varepsilon^t\}_{t=1}^{\infty}$. Taking expectations on both sides of (A.43) and taking (A.44) into account,

$$v(z_{-1}(\lambda), v_0, v_\varepsilon) > \lambda v(z_{-1}, v_0, v_\varepsilon) + (1 - \lambda)v(z'_{-1}, v_0, v_\varepsilon) .$$

Q.E.D.

Before proceeding with Lemma 6 it is useful to introduce some new notation. Let a be a scalar, b and c be vectors of dimension n and m , respectively and let B , D and E be matrices. The latter are of dimension $n \times n$, $n \times m$, and $m \times m$, respectively. Also, B and E are symmetric. Define the following set of functions:

$$\begin{aligned}
(A.45) \quad \mathcal{V} = \{ & q(z_{t-1}, v_t) \mid q(z_{t-1}, v_t) = a + b^T z_{t-1} + c^T v_t \\
& + z_{t-1}^T B z_{t-1} + z_{t-1}^T D v_t + v_t^T E v_t, B \leq 0, z_{t-1} \in \mathbb{R}^n, v_t \in \mathbb{R}^m \}
\end{aligned}$$

where $B \leq 0$ indicates negative semidefinite. Evidently, \mathcal{V} is the set of quadratic functions in the variables z_{t-1} and v_t and negative semidefinite in z_{t-1} .

Define the operator U with domain \mathcal{V} as follows:

$$(A.46) \quad U(J)(z_{t-1}, v_t) = \sup_{z_t \in \mathbb{R}^n} \{U(z_t, z_{t-1}, v_t) + \beta E_t J(z_t, \rho v_t + \mu + \varepsilon_{t+1})\}$$

for $J \in \mathcal{V}$. The function $U(\cdot, \cdot, \cdot)$ is defined in (2.6). An element $J^* \in \mathcal{V}$ is said to be a fixed point in \mathcal{V} under U if $J^* = U(J^*)$. If the function in brackets in (A.46) is quadratic and strictly concave in z_t , then there is a unique linear function $L(z_{t-1}, v_t)$ such that the sup is achieved by $z_t = L(z_{t-1}, v_t)$. Also, the function L is the only one that satisfies the following condition:

$$(A.47) \quad U_1(L(z_{t-1}, v_t), z_{t-1}, v_t) + \beta E_t J_1(L(z_{t-1}, v_t), \rho v_t + \mu + \varepsilon_{t+1}) = 0.$$

Lemma 6

If,

- (i) the function $u(\cdot, \cdot, \cdot)$ in (2.1) is strictly concave in its first two arguments,

then,

- (ii) the optimal value of the LQ approximation problem, (2.5), is a function $J \in \mathcal{V}$,
- (iii) J is the unique fixed point in \mathcal{V} under the U operator,
- (iv) the optimal linear feedback law that solves (2.5) is stationary and is the only function that satisfies (A.47).

Proof

The proof is accomplished by expressing (2.5) as a linear regulator problem and making use of results established elsewhere for that problem.

Let $y_t = (z_{t-1}^T, v_t^T)^T$. Evidently, y_t is an $(n+m) \times 1$ vector. The function (2.6) may be written as follows,

$$(A.48) \quad U(z_{t+j}, y_{t+j}) = \{c + \delta^T y_{t+j} + \phi^T z_{t+j} + y_{t+j}^T \bar{R} y_{t+j} + z_{t+j}^T Q z_{t+j} + 2y_{t+j}^T \bar{F} z_{t+j}\},$$

where $\delta \sim (n+m) \times 1$, $\phi \sim n \times 1$, $\bar{R} \sim (n+m) \times (n+m)$,

$\bar{F} \sim (n+m) \times n$, $Q \sim n \times n$, $j > 0$ and c is a constant term.

It is easily verified that

$$(A.49) \quad \bar{R} = \frac{1}{2} \begin{bmatrix} u_{22} & u_{23} \\ u_{23}^T & u_{33} \end{bmatrix}, \quad Q = \frac{1}{2} u_{11}, \quad \bar{F} = \frac{1}{2} \begin{bmatrix} u_{21} \\ u_{31} \end{bmatrix}.$$

Here, as in the proof to Lemma 4, the argument list of a function is dropped if the function is evaluated at $z_{t-1} = z_s$, $z_t = z_s$, $v_t = v_s$. Thus, in (A.49),

$$u_{33} \equiv u_{33}(z_s, z_s, v_s).$$

Let

$$\bar{R} = \begin{bmatrix} \bar{R} & \frac{1}{2} \delta \\ \frac{1}{2} \delta^T & Q_{1 \times (n+m)} \end{bmatrix}, \quad F = \begin{bmatrix} \bar{F} \\ \frac{1}{2} \phi^T \end{bmatrix}, \quad x_t = \begin{bmatrix} y_t \\ 1 \end{bmatrix}$$

(A.50)

$$B = \begin{bmatrix} I_n \\ Q_{(m+1) \times n} \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} Q_{n \times n} & 0_{n \times m} & 0_{n \times 1} \\ 0_{m \times n} & \rho & 0_{m \times 1} \\ 0_{1 \times n} & 0_{1 \times m} & 1 \end{bmatrix}, \quad w_t = \begin{bmatrix} 0_{n \times 1} \\ \epsilon_{t+1} \\ 1 \end{bmatrix}$$

where ρ is the $m \times m$ matrix defined in (2.2). Define

$$(A.51) \quad R = \bar{R} - FQ^{-1}F^T, \quad v_t = z_t + Q^{-1}F^T x_t$$

$$A = [\bar{A} - BQ^{-1}F^T].$$

In (A.51), the invertibility of Q follows from the negative definiteness of

u_{11} , a consequence of (i). Specifically,

$$(A.52) \quad Q = \frac{1}{2} u_{11} \text{ is negative definite.}$$

Using the notation of (A.50) and (A.51), (A.48) may be written

$$(A.48') \quad U(z_{t+j}, Y_{t+j}) = \{x_{t+j}^T R x_{t+j} + v_{t+j}^T Q v_{t+j} + c\}.$$

Then, problem (2.5) may be formulated as a standard discounted linear regulator problem as follows. Take the limit as $N \rightarrow \infty$ of the following sequence of problems. Maximize, over plans of the form $v_{t+j} = -F_j^{(N)} x_{t+j}$, $j = 0, 1, \dots, N$, the expression

$$(A.53a) \quad E \sum_{j=0}^N \beta^j \{x_{t+j}^T R x_{t+j} + v_{t+j}^T Q v_{t+j} + c\}$$

subject to

$$(A.53b) \quad x_{t+j+1} = A x_{t+j} + B v_{t+j} + w_{t+j} \quad j = 0, 1, \dots, N,$$

where $\{w_{t+j}\}_{j=0}^N$ is serially independent and

$$(A.53c) \quad E w_{t+j} w_{t+j}^T = W \quad j = 0, 1, \dots, N.$$

Here,

$$(A.53d) \quad W = \begin{bmatrix} Q_{n \times n} & Q_{n \times m} & Q_{n \times 1} \\ Q_{m \times n} & V_\varepsilon & Q_{m \times 1} \\ Q_{1 \times n} & Q_{1 \times m} & Q_{1 \times 1} \end{bmatrix},$$

and V_ε is the $m \times m$ symmetric matrix defined in (2.2).

Partition the matrices A and B in (A.50) and (A.51) as follows:

$$(A.54) \quad A = \begin{bmatrix} A_{11} & A_{12} \\ n \times m & n \times (m+1) \\ A_{21} & A_{22} \\ (m+1) \times n & (m+1) \times (m+1) \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ n \times n \\ B_2 \\ (m+1) \times n \end{bmatrix}.$$

From (A.50) and (A.51) deduce that, in (A.54),

$$A_{21} = 0, \quad A_{22} = \begin{bmatrix} \rho & 0 \\ 0 & 1 \end{bmatrix}, \quad B_2 = 0, \quad B_1 = I_n.$$

It follows that the rank of the matrix $[B_1 \ : \ A_{11}B_1 \ : \ \dots \ : \ A_{11}^{n-1}B_1]$ is n , so that (A_{11}, B_1) is completely controllable. Because of this and the fact that the eigenvalues of $\beta^{1/2} A_{22}$ lie inside the unit circle, conclude that

$$(A.55) \quad (\sqrt{\beta} A, \sqrt{\beta} B) \text{ is stabilizable.}$$

(See Sargent [1980, Section II] and Kwakernaak and Sivan [1972, page 452] for a discussion of stabilizability.)

The partition of A and B in (A.54) implies the following partition of x_t ,

$$(A.56) \quad x_{t+j} = \begin{pmatrix} x_{1,t+j} \\ x_{2,t+j} \end{pmatrix}, \quad x_{1,t+j} = z_{t+j}, \quad x_{2,t+j} = \begin{pmatrix} v_{t+j} \\ 1 \end{pmatrix}$$

$j = 0, \dots$. Partition the matrix R defined in (A.51) conformably with the partition of x_t in (A.56):

$$(A.57) \quad R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}.$$

Substituting from (A.49) and (A.50) into (A.51), it may be verified that

$$(A.58) \quad R_{11} = \frac{1}{2} (u_{22} - u_{21} u_{11}^{-1} u_{21}^T) .$$

Condition (i) of the Lemma implies that R_{11} is negative definite. This is shown as follows. First,

$$(A.59) \quad \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}$$

is negative definite. But, a matrix is negative definite if, and only if, its inverse is (Proposition 59, Dhrymes [1978, page 488]). In the case of (A.58) this implies that

$$(A.60) \quad \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \psi & -u_{11}^{-1} u_{12} \theta \\ -\theta u_{21} u_{11}^{-1} & \theta \end{bmatrix} ,$$

where $\theta = (u_{22} - u_{21} u_{11}^{-1} u_{21}^T)^{-1}$, $\psi = (u_{11} - u_{12} u_{22}^{-1} u_{21}^T)^{-1}$ is negative definite. But (A.60) being negative definite implies that θ and, hence, θ^{-1} is. But $\theta^{-1} = 2R_{11}$, yielding the result sought.

Since R_{11} is negative definite, $-R_{11}$ is positive definite. This is equivalent with the existence of a nonsingular matrix S such that

$$(A.61) \quad -R_{11} = S^T S .$$

(See Dhrymes [1978, page 480]. Note, R_{11} is symmetric.)

Consider the matrix pair $(\beta^{1/2} A_{11}, S)$. This is said to be detectable if, and only if, $(\beta^{1/2} A_{11}^T, S^T)$ is stabilizable. But this follows trivially from the fact that $(\beta^{1/2} A_{11}^T, S^T)$ is completely controllable. Consequently,

$$(A.62) \quad (\beta^{1/2} A_{11}, S) \text{ is detectable .}$$

(See Sargent [1980] for a discussion of detectability.)

Let $\{F_0^{(N)}, \dots, F_N^{(N)}\}$ denote the sequence of optimal linear feedback rules for problem (A.53) and for finite N . It can be shown (see Sargent [1980]) that this sequence exists and is unique. (Note, the assumptions we make here are weaker than the assumptions that are usually made for these problems. For example, Bertsekas [1976], Chow [1975, Chapter 7], Kushner [1971] and Kwakernaak and Sivan [1972] assume that R is negative semi-definite.)

Let $F_j^{(N)}$ denote the j^{th} element ($0 < j < N$) in the sequence of optimal linear feedback rules described in the previous paragraph. Sargent (1980, pages III.30-III.31) proves that under (A.52), (A.55) and (A.62),

$$(A.63) \quad \lim_{N \rightarrow \infty} F_j^{(N)} = G$$

for all j . In (A.63),

$$(A.64) \quad G = (\beta B^T P B + Q)^{-1} \beta B^T P A ,$$

where P is the unique symmetric matrix with negative semidefinite upper left $n \times n$ block which solves the following matrix Riccati equation

$$(A.65) \quad P = \beta A^T P A + R - \beta^2 A^T P B (\beta B^T P B + Q)^{-1} B P A .$$

Thus, the solution to problem (A.53) exists, is unique and of the form

$$(A.66) \quad v_{t+j} = -G x_{t+j} \quad j > 0 .$$

Sargent shows that the value of problem (A.53) converges, as $N \rightarrow \infty$, to

$$(A.67) \quad \frac{c}{1-\beta} + x_t^T P x_t + \frac{\beta}{1-\beta} \text{tr} P_{22} v_{\varepsilon} = q^*(x_t) ,$$

say, where P_{22} is the middle block in the following partition of P :

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ n \times n & n \times m & n \times 1 \\ P_{21} & P_{22} & P_{23} \\ m \times n & m \times m & m \times 1 \\ P_{31} & P_{32} & P_{33} \\ 1 \times n & 1 \times m & 1 \times 1 \end{bmatrix} .$$

Substituting from (A.51), (A.66) may be rewritten

$$(A.68) \quad z_{t+j} = L(z_{t+j-1}, v_{t+j}) ,$$

where $L(z_{t+j-1}, v_{t+j}) = Lx_{t+j}$, $L = -(Q^{-1}F^T + G)$. Equation (A.68) gives the optimal linear feedback law for (2.5) in terms of the notation used there.

Define the following set of functions:

$$(A.69) \quad \tilde{V} = \{q(x) \mid q(x) = x^T H x + d, \text{ the upper left } n \times n \text{ block of } H \text{ is negative semidefinite, } x = (y^T, 1)^T, y \in R^{(m+n)}, d \in R, H \text{ is symmetric.}\} .$$

Define the operator \tilde{U} on the space \tilde{V} as follows. For $q \in \tilde{V}$,

$$(A.70) \quad \tilde{U}(q)(x_t) = \sup_{v_t \in R^m} \{c + x_t^T R x_t + v_t^T Q v_t + \beta E q(x_{t+1})\} ,$$

where $x_{t+1} = Ax_t + Bv_t + w_t$ and where the expectation is taken over the distribution of w_t , given in (A.53c) and (A.53d). It may be verified that for the q function defined in (A.67),

$$(A.71) \quad \tilde{U}(q^*) = q^* , \quad q^* \in \tilde{V} .$$

Sargent establishes that q^* is the only element of \tilde{V} that satisfies (A.71). (Essentially, this follows from the fact that the solution to (A.65) is unique in the class of P matrices with negative semidefinite upper left $n \times n$ block.)

Note that the sets \tilde{V} and V , defined in (A.69) and (A.45), respectively, coincide. Let J be the function q^* when x_t is replaced by $(z_{t-1}^T, v_t^T, 1)^T$. That is,

$$(A.72) \quad J(z_{t-1}, v_t) = \frac{\beta}{1-\beta} \text{tr } P_{22} v_t \varepsilon + \frac{c}{1-\beta} + P_{33} \\ + z_{t-1}^T P_{11} z_{t-1} + 2v_t^T P_{21} z_{t-1} \\ + v_t^T P_{22} v_t + 2P_{31} z_{t-1} + 2P_{32} v_t.$$

Since J on V , (ii) of the Lemma is established.

Since the distinction between the operator \tilde{U} on \tilde{V} and the operator U on V is one of notation only, the fact that q^* is the unique fixed point under \tilde{U} on \tilde{V} is equivalent with the fact that J is the unique fixed point under U on V . This establishes (iii) of the Lemma. Finally, consider (iv) of the Lemma.

Substitute q^* into the expression on the right side of (A.70) and then replace v_t by $z_t + QZ^{-1}F^T x_t$ and x_t by $(z_{t-1}^T, v_t^T, 1)^T$ to get,

$$(A.73) \quad \sup_{z_t \in R^n} \{U(z_t, z_{t-1}, v_t) + \beta E J(z_t, \rho v_t + \mu + \varepsilon_{t+1})\}.$$

The function in (A.73) is quadratic in z_t . It is strictly concave since u_{11} is negative definite and P_{11} is negative semidefinite. Consequently, the sup in (A.73) is achieved by a unique point $z_t^* \in R^n$. It can be verified that $z_t^* = L(z_{t-1}, v_t)$. This establishes (iv) of the Lemma.

Q.E.D.

Proof of Theorem 1

Define the quadratic function

$$(2.10) \quad \begin{aligned} \tilde{V}(z_{t-1}, v_t) &= V(z_{t-1}, v_t) \\ &+ \frac{1}{2} \left(\frac{\beta}{1-\beta} \right) \text{tr} [v_{22}(z_s, v_s, v_\varepsilon) v_\varepsilon] \\ &+ \ell(Q) - \ell(v_\varepsilon), \end{aligned}$$

defined in (2.10) in the text and reproduced here for convenience. In (2.10), $U(\cdot, \cdot, \cdot)$, $g(\cdot, \cdot)$ and $V(\cdot, \cdot)$ are given in (2.6), (2.8) and (2.9), respectively. Lemma 4 asserts that (i)-(v) imply

$$(A.74) \quad \begin{aligned} \tilde{V}(z_{t-1}, v_t) &= U(g(z_{t-1}, v_t), z_{t-1}, v_t) \\ &+ \beta \mathbb{E}_t \tilde{V}(g(z_{t-1}, v_t), v_{t+1}) \end{aligned}$$

and

$$(A.75) \quad U_1(g(z_{t-1}, v_t), z_{t-1}, v_t) + \beta \mathbb{E}_t \tilde{V}_1(g(z_{t-1}, v_t), v_{t+1}) = 0.$$

Lemma 5 establishes that, if (2.1) has the solution, (2.3), and (vi) and (vii) hold, then the function $v(\cdot, \cdot, \cdot)$, defined in (2.4) is concave in its first argument. Consequently, $v_{11}(z_s, z_s, v_\varepsilon)$ is negative semidefinite. It follows that the quadratic function \tilde{V} in (2.10) is negative semidefinite in its first argument. That is, $\tilde{V} \in \mathcal{V}$ (see (A.45) for \mathcal{V}).

By (vi) $U(\cdot, \cdot, \cdot)$ is negative definite in its first argument.

Consequently,

$$(A.76) \quad U(z_t, z_{t-1}, v_t) + \beta \mathbb{E}_t \tilde{V}(z_t, v_{t+1})$$

is a negative definite quadratic function in z_t . A necessary and sufficient condition for $z_t^* \in \mathbb{R}^n$, where $z_t^* = z_t^*(z_{t-1}, v_t)$, to maximize (A.76),

therefore, is that

$$(A.77) \quad U_1(z_t^*, z_{t-1}, v_t) + \mathbb{E}_t \tilde{V}_1(z_t^*, v_{t+1}) = 0 .$$

Thus, (A.75) establishes that $g(z_{t-1}, v_t)$ maximizes the function (A.76).

Relation (A.74) then indicates that \tilde{V} is a fixed point under the U mapping defined in (A.46). That is, $U(\tilde{V}) = \tilde{V}$.

According to Lemma 6, (vi) implies that the value of the solution to (2.5), call it J , is quadratic in z_{t-1} and v_t and negative semidefinite in z_{t-1} . In addition it is shown that $U(J) = J$ and that J is the unique fixed point under U in V . Conclude that $J = \tilde{V}$. That is, \tilde{V} is the value of problem (2.5).

Lemma 6 also asserts that the optimal linear feedback law that solves (2.5) is stationary, call it $L(z_{t-1}, v_t)$. In addition, it is shown that $z_t^* = L(z_{t-1}, v_t)$ is the only function guaranteeing (A.77). Conclude that $g = L$. Q.E.D.

Appendix B

In this appendix we calculate the mean and covariogram of $\{x_t\}$, where

$$(B.1a) \quad x_t = \lambda_t B x_{t-1}^\alpha, \quad B = \beta \alpha \left(\frac{H\beta\gamma(1-\alpha)}{1+\beta\gamma(1-\alpha)} \right)^{(1-\alpha)}, \quad \gamma = \frac{\theta}{1-\beta\alpha}, \quad 0 < \alpha < 1$$

and

$$(B.2b) \quad \lambda_t = \lambda_{t-1}^\rho \exp(\mu + \varepsilon_t) \quad \varepsilon_t \sim \text{iidN}(0, \sigma_\varepsilon^2), \quad |\rho| < 1.$$

Assume that $\{x_t\}$ started up at $t = t_0 = -\infty$ with the initial conditions x_{t_0} and λ_{t_0} being nonstochastic. We proceed now to calculate μ_x and $\sigma_x^2(\tau)$. Consider the $\{\lambda_t\}$ process first:

$$\lambda_{t_0+1} = \lambda_{t_0}^\rho \exp(\mu + \varepsilon_{t_0+1})$$

$$\lambda_{t_0+2} = \lambda_{t_0+1}^\rho \exp(\mu + \varepsilon_{t_0+2})$$

$$= \lambda_{t_0}^{\rho^2} \exp[(\mu + \varepsilon_{t_0+1})\rho] \exp(\mu + \varepsilon_{t_0+2})$$

⋮

$$(B.3) \quad \lambda_{t_0+i} = \lambda_{t_0}^{\rho^i} \prod_{j=0}^{i-1} \exp[(\mu + \varepsilon_{t_0+i-j})\rho^j].$$

Next, consider $\{x_t\}$. Making use of (B.3),

$$\begin{aligned}
 x_{t_0+1} &= \lambda_{t_0+1}^B x_{t_0}^\alpha \\
 &= \lambda_{t_0}^\rho \exp(\mu + \varepsilon_{t_0+1}) B x_{t_0}^\alpha \\
 x_{t_0+2} &= \lambda_{t_0}^{(\rho^2 + \alpha\rho)} \exp[(\mu + \varepsilon_{t_0+1})(\rho + \alpha)] \exp[(\mu + \varepsilon_{t_0+2})] B^{(1+\alpha)} x_{t_0}^{\alpha^2} \\
 &\vdots \\
 x_{t_0+i} &= \lambda_{t_0}^{\left(\sum_{j=0}^{i-1} \alpha^j \rho^{i-j}\right)} \prod_{r=1}^i \exp[(\mu + \varepsilon_{t_0+i+1-r}) \left(\sum_{j=0}^{r-1} \alpha^j \rho^{r-j}\right) \rho^{-1}] B^{\left(\sum_{j=0}^{i-1} \alpha^j\right)} x_{t_0}^{\alpha^i}.
 \end{aligned}$$

Substituting $t = t_0 + i$,

$$x_t = \lambda_{t_0}^{\left(\frac{\rho^{t-t_0} - \alpha^{t-t_0}}{1 - \alpha/\rho}\right)} \prod_{r=1}^{t-t_0} \exp[(\mu + \varepsilon_{t+1-r}) \left(\frac{\rho^r - \alpha^r}{\rho - \alpha}\right)] B^{\left(\frac{1 - \alpha^{t-t_0}}{1 - \alpha}\right)} x_{t_0}^{\alpha^{(t-t_0)}}.$$

Letting $t_0 \rightarrow -\infty$, we get

$$(B.4) \quad x_t = \prod_{r=1}^{\infty} \exp[(\mu + \varepsilon_{t+1-r}) \left(\frac{\rho^r - \alpha^r}{\rho - \alpha}\right)] B^{\left(\frac{1}{1-\alpha}\right)}.$$

Then, making use of the independence of $\{\varepsilon_t\}$,

$$\begin{aligned}
 (B.5) \quad \mu_x \equiv E x_t &= \prod_{r=1}^{\infty} E \exp[(\mu + \varepsilon_{t+1-r}) \left(\frac{\rho^r - \alpha^r}{\rho - \alpha}\right)] B^{\left(\frac{1}{1-\alpha}\right)} \\
 &= B^{\left(\frac{1}{1-\alpha}\right)} \prod_{r=1}^{\infty} \exp\left[\left(\frac{\rho^r - \alpha^r}{\rho - \alpha}\right) \mu + \frac{\sigma_\varepsilon^2}{2} \left(\frac{\rho^r - \alpha^r}{\rho - \alpha}\right)^2\right] \\
 &= \left[B \exp\left(\frac{\mu}{1-\rho}\right)\right]^{\left(\frac{1}{1-\alpha}\right)} \exp\left\{\frac{\varepsilon}{2} \cdot \frac{1 + \rho\alpha}{(1 - \rho^2)(1 - \alpha\rho)(1 - \alpha^2)}\right\}.
 \end{aligned}$$

The second equality in (B.5) makes use of the formula for the mean of a log-normally distributed random variable (see, e.g., Maddala (1977, p. 33)).

Next, consider $\sigma_x^2(\tau) = \text{Ex}_t x_{t-\tau} - (\text{Ex}_t)^2$ for $\tau > 0$. Using (B.4),

$$\begin{aligned} x_t x_{t-\tau} &= \frac{2}{B^{1-\alpha}} \prod_{r=1}^{\infty} \exp\left[(\mu + \varepsilon_{t+1-r}) \left(\frac{\rho^r - \alpha^r}{\rho - \alpha}\right)\right] \\ &\quad \times \prod_{\ell=1}^{\infty} \exp\left[(\mu + \varepsilon_{t+1-\tau-\ell}) \left(\frac{\rho^\ell - \alpha^\ell}{\rho - \alpha}\right)\right] \\ &= \frac{2}{B^{1-\alpha}} \prod_{r=1}^{\tau} \exp\left[(\mu + \varepsilon_{t+1-r}) \left(\frac{\rho^r - \alpha^r}{\rho - \alpha}\right)\right] \\ &\quad \times \prod_{\ell=1}^{\infty} \exp\left[(\mu + \varepsilon_{t+1-\tau-\ell}) \left(\frac{\rho^{\tau+\ell} - \alpha^{\tau+\ell}}{\rho - \alpha} + \frac{\rho^\ell - \alpha^\ell}{\rho - \alpha}\right)\right]. \end{aligned}$$

Making use of the independence of $\{\varepsilon_t\}$ and the formula for the mean of a log-normal variable, one obtains (after some algebra),

$$\text{Ex}_t x_{t-\tau} = (\mu_x)^2 \exp\left[\frac{\sigma_\varepsilon^2 \rho}{(1-\rho^2)(1-\alpha\rho)(\rho-\alpha)}\right]^{\rho^\tau} \exp\left[\frac{\sigma_\varepsilon^2 \alpha}{(1-\alpha^2)(1-\alpha\rho)(\alpha-\rho)}\right]^{\alpha^\tau}.$$

It may be verified that the above formula holds in the $\tau = 0$ case. Then for $\tau = 0, \pm 1, \pm 2, \dots$, we have,

$$\begin{aligned} &\sigma_x^2(\tau) \\ (B.6) \quad &= (\mu_x)^2 \left\{ \exp\left[\frac{\sigma_\varepsilon^2 \rho}{(1-\rho^2)(1-\alpha\rho)(\rho-\alpha)}\right]^{\rho^{|\tau|}} \exp\left[\frac{\sigma_\varepsilon^2 \alpha}{(1-\alpha^2)(1-\alpha\rho)(\alpha-\rho)}\right]^{\alpha^{|\tau|}} - 1 \right\} \end{aligned}$$

FOOTNOTES

¹Related work that also does not rely on the linear quadratic assumption is found in Hansen and Singleton (1982). Using results in Hansen (1982), they propose an estimation method that exploits the orthogonality properties that a solution to an optimum problem must satisfy. Their approach does not require that an optimum problem be linear quadratic and can be used to estimate the parameters of private agents' objective functions even when it is not known how to calculate the decision rule that solves the objective function.

²Here, the phrase "return function" denotes the summand in the calculus of variations problem.

³Here, by the "certainty version of a problem" is meant the problem in which random variables are replaced by their unconditional expectation.

⁴The "steady state" as it has been defined, and the unconditional mean of the stochastic process in the state variable are not in general the same. This has been pointed out in Merton (1975) and in Danthine and Donaldson (1981). See footnote 6 below.

⁵Differentiating the certainty version ($\sigma_\epsilon^2 = 0$) of (3.4) with respect to L_{t+j} and x_{t+j} yields, after setting the results to zero:

$$\frac{1}{H - L_{t+j}} = \frac{(1 - \alpha)\beta\theta y_{t+j+1}/L_{t+j}}{Y_{t+j+1} - x_{t+j+1}}$$

$$\frac{\theta}{Y_{t+j} - x_{t+j}} = \frac{\beta\theta\alpha x_{t+j+1}/x_{t+j}}{Y_{t+j+1} - x_{t+j+1}},$$

where $Y_{t+k} = \exp\left(\frac{\mu}{1-\rho}\right)L_{t+k-1}^{(1-\alpha)\alpha} x_{t+k-1}$, $k = j, j+1$. Setting $x_{t+j} = x_{t+j+1} = x_s^r$ and $L_{t+j+1} = L_{t+j} = L_s^r$ yields (3.10b) and (3.10c). x_s^r and L_s^r are obtained by evaluating (3.10b) and (3.10c) at $\alpha = \alpha^r$, $\mu = \mu^r$.

⁶It is of interest to note that the analog to the result in (3.13) fails to hold if the quadratic approximation described in step 2 is calculated about $(L_t, L_{t-1}, x_t, x_{t-1}, v_t) = (\tilde{L}, \tilde{L}, \tilde{x}, \tilde{x}, v_s) \neq (L_s, L_s, x_s, x_s, v_s)$. That is, in this case the solution to the LQ approximation problem is not a first order Taylor series expansion of the exact feedback rule, (3.9), about

$(L_{t-1}, x_{t-1}, v_t) = (\tilde{L}, \tilde{x}, v_s)$ An example will illustrate this point. Let $H = 1$, $\mu = 1$, $\sigma_\varepsilon^2 = 2$, $\rho = 1/2$, $\theta = 2$, $\alpha = .25$, $\beta = \frac{1}{1.05}$. Then, the exact solution to (3.1) is, from (3.9)

$$(a) \quad x_t = .238 \exp(v_t) L_{t-1}^{.75} x_{t-1}^{.25}, \quad v_t = .5v_{t-1} + 1.0 + \varepsilon_t,$$

$$(b) \quad L_t = .652.$$

In this case $L_s = .652$, $v_s = 2.0$, $x_s = 1.385$, and $Ex_t = 1.663$. Consider $(\tilde{L}, \tilde{x}, v_s) = (.652, 1.663, 2.0)$. Solving the LQ approximation problem with the modification that the quadratic expansion is taken about $(\tilde{L}_s, \tilde{x}_s, v_s)$ we get the following solution:

$$(c) \quad x_t = -4.5624 + .0340x_{t-1} + 2.3256L_{t-1} + 2.0502v_t$$

$$(d) \quad L_t = .6311 + .0031L_{t-1} + .0004x_{t-1} + .0078v_t.$$

The first order Taylor series expansion of (a) and (b) about

$$(L_{t-1}, x_{t-1}, v_t) = (.652, 1.663, 2.0) \text{ is}$$

$$(e) \quad x_t = -2.8998 + .2180x_{t-1} + 1.6674L_{t-1} + 1.4500v_t$$

$$(f) \quad L_t = .652.$$

Evidently (c) and (e) and (d) and (f) do not match up. These expressions would have matched up had the calculations been done on the basis of $(\tilde{L}, \tilde{x}, v_s) = (.652, 1.385, 2.0) = (L_s, x_s, v_s)$.

⁷From equation (3.10c), the steady state value of x_t when v_t is replaced by its unconditional mean is $x_s = [\exp(\mu/(1-\rho))B]^{(1/(1-\alpha))}$. For $\sigma_\varepsilon^2 > 0$ this is strictly less than Ex_t , according to (4.5). The two are equal from $\sigma_\varepsilon^2 = 0$.

⁸See Hannan (1970, Chapter IV).

⁹One way of calculating this is as follows. The covariance generating function (see Sargent (1979, Chapter XI)) of $\ln Y_t$, $g_{\ln Y}(z)$,

$$g_{\ln Y}(z) = \sum_{k=-\infty}^{+\infty} \sigma_{\ln Y}^2(k) z^k = \frac{\sigma_\varepsilon^2}{(1 - \rho z)(1 - \alpha z)(1 - \rho z^{-1})(1 - \alpha z^{-1})},$$

where $\sigma_{\ln Y}^2(k) \equiv E[\ln Y_t - E \ln Y_t][\ln Y_{t-k} - E \ln Y_{t-k}]$. Then (see e.g., Sargent (1979, p.232)),

$$(+)\quad \sigma_{\ln Y}^2(0) = \frac{1}{2\pi i} \int_C g_{\ln Y}(z) \frac{dz}{z},$$

where C is the unit circle and the integration is carried out in a counter-clockwise direction. The function $g_{\ln Y}(z) \frac{1}{z}$ is analytic on the unit circle. It is analytic everywhere inside C except at the points $z = \alpha$ and $z = \rho$. By the residue theorem (see, e.g., Churchill (1976, p. 172)) the expression to the right of the equality in (+) equals $R_\alpha + R_\rho$, where R_α and R_ρ are the residues of $g_{\ln Y}(z) \frac{1}{z}$ at the points $z = \alpha$ and $z = \rho$, respectively. These are defined as follows:

$$R_\alpha = \lim_{z \rightarrow \alpha} (z - \alpha) g_{\ln Y}(z) \frac{1}{z}, \quad R_\rho = \lim_{z \rightarrow \rho} (z - \rho) g_{\ln Y}(z) \frac{1}{z}.$$

Evaluating R_α and R_ρ and taking their sum yields (5.4).

¹⁰The result appears to be consistent with conclusions reached by Zellner and Geisel (1968) in a framework related to the one we consider. They examine several single period stochastic optimization problems and find that replacing the objective function by a quadratic approximation has little effect on the optimal setting of the control variable, unless there is a pronounced asymmetry in the objective function. The objective function used in this paper, (3.4), does not, for example, exhibit significant asymmetries relative to variations in L_t and x_t .

BIBLIOGRAPHY

- Anderson, T. W., 1971, The Statistical Analysis of Time Series (John Wiley and Sons, New York).
- Bartle, Robert G., 1964, The Elements of Real Analysis (John Wiley and Sons, New York).
- Benveniste, L. M., and J. A. Scheinkman, 1979, "On the Differentiability of the Value Function in Dynamic Models of Economics," Econometrica 47: 727-732.
- Bertsekas, Dimitri P., 1976, Dynamic Programming and Stochastic Control (Academic Press, New York).
- Chow, Gregory C., 1975, Analysis and Control of Dynamic Economic Systems (John Wiley and Sons, New York).
- _____, 1979, "Estimation of Rational Expectations Models, Econometrics Research Program Research Memorandum No. 252, Princeton University.
- Christiano, Lawrence J., and Edward C. Prescott, 1982, "Notes on Stochastic Optimization in Log Linear Models", mimeo, August.
- Churchill, Ruel V., James W. Brown, and Roger F. Verhey, 1976, Complex Variables and Applications (McGraw-Hill, New York, 3rd edition).
- Danthine, Jean-Pierre, and John B. Donaldson, 1981, "Certainty Planning in an Uncertain World: A Reconsideration," Review of Economic Studies XLVIII: 507-510.
- Granger, C.W.J., and Paul Newbold, 1977, Forecasting Economic Time Series (Academic Press, New York).
- Hannan, E. J. , 1970, Multiple Time Series (John Wiley and Sons, New York).
- Hansen, Lars P., 1982, "Large Sample Properties of Generalized Method of Moments Estimators," Econometrica.

- Hansen, Lars P., and Thomas J. Sargent, 1980, "Formulating and Estimating Dynamic Linear Rational Expectations Models," Journal of Economic Dynamics and Control 2: 7-46.
- _____, 1981, "Linear Rational Expectations Models for Dynamically Interrelated Variables, in R. Lucas and T. Sargent, editors, Rational Expectations and Econometric Practice (University of Minnesota Press, Minneapolis).
- Hansen, Lars P., and Kenneth J. Singleton, 1982, "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," Econometrica forthcoming.
- Holt, Charles C., Franco Modigliani, John F. Muth, and Herbert Simon, 1960, Planning Production, Inventories, and Work Force (Englewood Cliffs, N. J.).
- Kennan, John, 1979, "The Estimation of Partial Adjustment Models with Rational Expectations," Econometrica 47: 1441-1455.
- Kushner, Harold, 1971, Introduction to Stochastic Control (Holt, Rinehart and Winston, Inc., New York).
- Kydland, Finn, and Edward C. Prescott, 1980, "A Competitive Theory of Fluctuations and the Feasibility and Desirability of Stabilization Policy," in Stanley Fischer, ed., Rational Expectations and Economic Policy (University of Chicago Press, Chicago, Illinois).
- _____, 1982, "Time to Build and Aggregate Fluctuations," forthcoming, Econometrica.
- Kwakernaak, Huibert, and Raphael Sivan, 1972, Linear Optimal Control Systems (Wiley, New York).
- Long, John B., Jr., and Charles I. Plosser, 1982, "Real Business Cycles," forthcoming, Journal of Political Economy.

- Luenberger, D. G., 1979, Introduction to Dynamic Systems (John Wiley and Sons, New York).
- Maddala, G. S., 1977, Econometrics (McGraw-Hill, New York).
- Mariano, Robert S., 1978, "Commodity Market Modeling: Methodological Issues and Control Theory Applications," in F. Gerard Adams and Jere R. Behrman, eds., Econometric Modeling of World Commodity Policy (Lexington Books, Lexington, Mass.)
- Merton, R., 1975, "An Asymptotic Theory of Growth Under Uncertainty," Review of Economic Studies 42: 375-393.
- Nerlove, Marc, David M. Grether, and José L. Carvalho, 1979, Analysis of Economic Time Series (Academic Press, New York).
- Noble, Ben, 1969, Applied Linear Algebra (Prentice-Hall, Inc., New Jersey).
- Sargent, Thomas J., 1979, Macroeconomic Theory (Academic Press, New York).
- _____, 1980, Lecture Notes on "Linear Prediction and Control," manuscript.
- _____, 1978, "Estimation of Dynamic Labor Demand Schedules under Rational Expectations," Journal of Political Economy, vol. 86, no. 6.
- Taylor, John B., 1980, "Output and Price Stability: An International Comparison," Journal of Economic Dynamics and Control 2: 109-132.
- Theil, H., 1971. Textbook of Econometrics (John Wiley and Sons, New York).
- Zellner, Arnold, and Martin S. Geisel, 1968, "Sensitivity of Controls to Uncertainty and Form of the Criterion Function," in D. G. Watts, ed., The Future of Statistics (Academic Press, New York).